MONOPOLY, VALUE, AND INVESTMENT

Some Problems in the Theory of Monopoly

* * * * *

(Thesis submitted for the degree of Doctor of Philosophy in the Australian National University)

CANBERRA, A.C.T., 30 August, 1956.

R. O. HIESER, M.Ec.
"Monopoly appears to be deeply rooted in the nature of the capitalist system: free competition, as an assumption, may be useful in the first stage of certain investigations, but as a description of the normal state of a capitalist economy it is a myth."

M. KALECKI

*Essays in the Theory of Economic Fluctuations*
CONTENTS

Introduction

PART I - TOWARDS A SYNTHESIS IN VALUE THEORY

Ch. I The Fundamental Difficulty of Monopolistic Competition Theory ... Page 1
II The Nature of Cross-elasticity of Demand ... ... ... ... ... 21
III The Large Group: The Transition from the Competitive Group to Pure Monopoly ... ... ... ... ... 49
IV The Small Group: What is Oligopoly? 118
V Elements of a Synthesis ... ... ... 166

PART II - THE MEASUREMENT OF AND CHANGES IN THE DEGREE OF MONOPOLY POWER

VI Monopoly and Its Measurement ... ... 181
VII The Degree of Monopoly Power ... ... 202
VIII The Degree of Monopoly Power (Cont'd.) 231
IX Changes in the Degree of Monopoly ... 268

PART III - MONOPOLY AND THE RATE OF INVESTMENT

X Monopoly and Investment: The Problem 308
XI The Effect of Monopoly on the Rate of Investment ... ... ... ... 323

Bibliography
INTRODUCTION

This Introduction is by way of explanation and apology. First, let us treat of the explanation.

In the following pages will be found what at first sight must appear an odd assortment of topics, an assortment not easily reconciled in terms of traditional pedagogical classifications. True, they have one thing in common; they all deal with some aspect of monopoly theory. But they also have a contextual unity and perhaps it will help to see the matters dealt with here in perspective, if we briefly outline the context in which they have arisen.

Ultimately, the writer is concerned with the effects of monopoly upon economic growth, with determining whether monopoly represents the "hardening of capitalism's arteries" as suggested by Marxists and secular stagnationists, or whether monopoly gives capitalism new direction, "planning" and stability as
argued by others,\(^1\) whether it tends to neutralize itself by the development of "countervailing power",\(^2\) or whether it is merely unimportant one way or the other, as often seems to be implied.

This is, of course, a very ambitious task and one is very soon struck by the inadequacy of the tools at hand with which to tackle it. Fundamental to such a study are, of course, the theory of growth and the theory of monopoly. Unfortunately, both are unsatisfactory in many respects.

As far as the theory of growth is concerned, we have not advanced much beyond the stage of formal models which tell us little more than that, if the system is in equilibrium to begin with and everything grows proportionately, the system will continue along a path of equilibrium growth. Two aspects of growth theory in particular are inadequate. First, since

---

1. Cf., for example, a wartime comment by Sir Edgar Jones of the International Tinplate Cartel: "If Great Britain and the United States will take the lead and get the producers of each main product to prepare schemes for world export regulation, then general employment can be insured." - The New Republic, Special Supplement on "Cartels", March 27, 1944.

the "natural" growth of a capitalist economy tends to be cyclical, a much closer integration between the theory of growth and the theory of the trade cycle is needed. The idea that we can study a course of growth independent of the trade cycle, with cyclical fluctuations looked upon merely as deviations from a trend, is, it seems to me, fundamentally wrong. On this question, I believe Marx's position as stated by Dobb to be the correct one: "Marx clearly regarded crises, not as incidental departures from a predetermined equilibrium, not as fickle wanderings from an established path of development to which there would be a submissive return, but rather as themselves a dominant form of movement which forged and shaped the development of capitalist society. To study crises was ipso facto to study the dynamics of the system ..."³.

The second weakness, common to both growth and trade cycle theory, is the treatment of Autonomous Investment. To explain growth in terms of an ad hoc addition of a steady rate of Autonomous Investment, as has been common recently, hardly adds to our knowledge of what causes growth. Clearly, we need

³. Political Economy and Capitalism, London, 1937, p.80,
to know a lot more about the determinants of investment before we can make much headway with explaining the process of economic growth.

The theory of monopoly is, of course, more highly developed than that of growth. However, there are important weaknesses here too. For example, we find bewildering differences of opinion about the definition of oligopoly. For some economists, oligopoly is exceptional in the economy, for another it is the rule. Again, the theory of the firm is in fragments and what was once thought to exemplify monopoly is now said to be indicative of the workings of competition. 4

Secondly, if we are to talk of greater or lesser degrees of monopoly, we clearly must have at least a conceptual notion of how monopoly is to be measured. A number of such measures have been proposed from time to time, but none has commanded general acceptance.

Again, as Steindl has pointed out, "one may feel also that the whole theory of prices and imperfect competition is not linked up sufficiently closely with certain other parts of economic theory,

---

in particular the theory of investment. This latter study must, I feel, ultimately provide the bridge between monopoly and growth, if one is to be built.

In the following pages we are concerned, then, with the second group of problems, viz. those of the theory of monopoly. In Part I we try and pick up some of the fragments of the theory of the firm. In Part II we are concerned with the development of a measure of the degree of monopoly power and with a summary discussion of how the degree of monopoly may be expected to change. In Part III, a preliminary discussion of the effect of monopoly on investment is undertaken.

Finally, a few words of apology seem called for. Firstly, many of the matters dealt with in these pages are very abstract and, I am afraid, make difficult reading, with no immediate bearing on our more interesting larger problem. The only justification which may be pleaded is that a thorough examination of fundamentals may lead to a simpler structure which is theoretically sound and has direct applicability in the real world. However, it is not wished to

suggest that such a model has been found in these pages.

This brings me to the second head of apology. It will soon become obvious to the reader that I have posed many more questions than I have answered. I can only hope that the setting down of some logical difficulties and inconsistencies in received doctrine will assist in their ultimate remedy.
PART I

TOWARDS A SYNTHESIS

IN VALUE THEORY
I. THE FUNDAMENTAL DIFFICULTY OF

MONOPOLISTIC COMPETITION THEORY

Keynes' General Theory apart, perhaps no theoretical innovation has been accorded such a prompt and general acceptance as the theory of monopolistic (or imperfect) competition.¹ Foreshadowed by Sraffa's celebrated paper of 1926,² the publication, in 1933, of Chamberlin's Theory of Monopolistic Competition and Joan Robinson's Economics of Imperfect Competition resulted in almost immediate capitulation to the general assumption of monopoly in place of the competitive assumptions of traditional theory, "classical" and "neo-classical" alike.³ Two factors conspired to cause this. Firstly, the assumptions of Chamberlin and Mrs. Robinson were patently a better description of the real world as it appeared to common

¹. "For the historian of economic thought, the most revolutionary feature of monopolistic competition theories will probably be the unprecedented pace at which they conquered their audience" - R. Triffin, Monopolistic Competition and General Equilibrium Theory, Cambridge, Mass., 1941, p. 17.
³. "Classical" in the Marxian sense.
observation than those of traditional theory. And secondly, the economic malaise of the 'thirties - so at odds with traditional theoretical presumptions - had engendered a widespread dissatisfaction with that body of doctrine. Like the General Theory of three years later, but to a somewhat lesser extent, the theory of monopolistic competition met a very urgent need.

Yet, from the outset, the victory was not wholly undisputed. Side by side with the development of monopolistic competition theory, a "neo-Lausanne school" of value theorists was building upon the work of Walras and Pareto in the theory of consumers' behaviour and the theory of the firm within a framework of general equilibrium analysis. Led by Hicks and R.D.G. Allen, this school proceeded upon the assumption of perfect competition. In some sense, the work of this school has been looked upon as a system rival to that of monopolistic competition. This is underlined by the fact that, although an attempt has been made to


5. Cf. "A general abandonment of the assumption of perfect competition, a universal adoption of the assumption of monopoly, must have very destructive consequences for economic theory" - J.R. Hicks, Value and Capital, p. 83.
synthesize monopolistic competition doctrine and the theory of general equilibrium, no such attempt has been made to integrate monopolistic competition theory with the theory of consumers' behaviour. This is the more surprising when we recall that Marshall's theory of the firm was intimately linked with the theory of consumer demand, the one determining the supply curve and the other the demand curve for the "industry".

Although the relationship between the theory of consumers' behaviour and the theory of monopolistic competition will not be a principal object of study in the following pages, it is hoped to show in passing that the dichotomy between the two schools of analysis stems from the fact that both have chosen to ignore the basic problem which, once faced, would serve as a bridge between them. The important thing for us at the moment, however, is the fact that this parallelism exists.

Furthermore, unlike the General Theory, the initial success of monopolistic competition theory was not followed by a period of elaboration, refinement, and consolidation. Rather, it was very soon to be attacked from without, and rift by dissension within; so that

6. R. Triffin, op. cit.
today opinion ranges between qualified acceptance and outright rejection.

The main attack from without stemmed from the fact that business firms did not appear to behave in the manner to be inferred from the theory of monopolistic competition. This theory was based on the assumption of profit-maximization, within a framework of a given demand curve and a given cost curve. The profit-maximization solution implied that entrepreneurs should attempt to equate their marginal revenue with their marginal cost. Yet empirical studies revealed — and this was in line with common observation — that, in general, firms had but the haziest idea of their elasticity of demand and, hence, their marginal revenue. Moreover, the notion of marginal cost was virtually unknown to them, as, indeed, any check with cost accounting textbooks will soon reveal.

In the event, there has developed an anti-marginalist school of value theorists which has found a good deal of support in the name of "realism." For them, price formation is, in one sense or another, based upon average or total cost. The headquarters of this school was set in Oxford, and the chief protagonists of the anti-marginalist approach were Hall and Hitch and P.W.S.
Andrews; and a weighty volume of literature has developed around their writings. 7

That the adherents to the theory of monopolistic competition are sorely divided amongst themselves will be known to anyone familiar with the literature; and it is not proposed to detail the history of these differences at this stage, since they will form a large part of the discussion in the succeeding chapters. All I wish to do here is to advance the thesis that the internal dissensions besetting monopolistic competition theory are traceable back to certain fundamental difficulties inherent in that theory which have been with it from the outset and have not, as yet, been satisfactorily resolved.

It will be further argued that these difficulties, of a logical nature, are but manifestations of a more basic puzzle which is also at the root of the dichotomy between monopolistic competition theory and the theory

of consumers' behaviour which we noticed above. Moreover, once this fundamental problem is resolved the apparent inconsistency between monopolistic competition theory and the anti-marginalist position largely dissolves.

If these contentions are true, then, the way opens towards a synthesis in value theory, which is certainly sorely needed. To provide such a synthesis is, however, a much too ambitious programme to be undertaken here. In what follows, we shall be mainly concerned with monopolistic competition doctrine and our approach will be essentially exploratory. In particular, we shall attempt to find a way round the basic difficulty of monopolistic competition theory. For the rest, we shall be content to indicate the points at which our results are tangential to the hypotheses of other schools of thought. Insofar as we may find such points of contact which show the way towards a synthesis, this will, in some sense, constitute a theoretical verification of our critique of monopolistic competition analysis.

The basic problem with which the theory of monopolistic competition has had to grapple since its inception concerns the definition of a "commodity" or,
what amounts to the same thing, the definition or delineation of a "market"? How wide, or how narrow, should we define the area of a market to be? How good do substitutes have to be before they rank as the same or similar commodities selling in the same market? In short, when are commodities the same, similar, or different? This basic problem, in turn, poses two subsidiary questions: (1) How unlike (different) do commodities have to be before we may describe their sellers as "pure monopolists" or "isolated sellers"; and (2) How alike do products have to be before their prices become interdependent – the "oligopoly" problem? These problems have worried theorists from the beginning and, of course, a number of alternative solutions have been advanced from time to time. But attempts to formulate logical definitions which will take care of one aspect of the problem have inevitably opened up difficulties in other directions. Much of the following discussion in this Part will be devoted to showing why this is so. Before proceeding to this, however, we must set out in more detail the issues involved.
1. "Pure Monopoly" versus "Polypoly"

From the very first, Mrs. Robinson recognized the difficulty of defining "pure monopoly" - in the traditional Marshallian sense - in contradistinction to a seller in a group which arises as soon as we adopt the general assumption of monopoly. In the Marshallian system there was no such problem. Producers either sold the same commodity or different commodities. A "pure monopolist" was then simply and unambiguously defined as the only seller of a given commodity - literally a single seller. But in the scheme of Chamberlin and Mrs. Robinson this simple dichotomy between commodities which are the same and commodities which are different no longer obtains. We now have a theoretical scheme in which goods are, in general, only imperfect substitutes for one another. The difference between any two products is now one of degree, not of kind.

Mrs. Robinson drew attention to the logical dilemma implied in this change of assumptions. On the one hand, "every individual producer has the monopoly of his own output."8 On the other hand, "since every

article must have some rivals, and since in the last resort every article represents a use of money which is rival to every other, we should be compelled to say that no such thing as complete monopoly exists. *9

"Thus any attempt at a logical definition of a monopolist drives either monopoly or competition quite out of the field. *10

Mrs. Robinson's "solution" of this dilemma was to come down on the side of the first alternative, so that "we have only to take the word monopoly in its literal sense, a single seller, and the analysis of monopoly immediately swallows up the analysis of competition. *11

Faced with the same difficulty, Chamberlin adopted a different solution. Mrs. Robinson had put the puzzle thus: "It was tempting, under the old scheme, to arrange actual cases in a series of which pure monopoly would be the limit at one end and pure competition at the other, but a definition of pure monopoly which would correspond to the definition of pure competition was extremely hard to find. *12

9. op. cit., p. 4.
10. op. cit., p. 5.
11. op. cit., p. 5.
12. op. cit., p. 4.
is easy enough to find the limiting case at the competitive end of the scale. The limiting case occurs when the demand for the product of an individual producer is perfectly elastic. But what is the limiting case at the other end? ... We know what we mean by 'selling in a perfect market', but what is a perfectly imperfect market? Chamberlin's answer to this question runs as follows: "if we regard monopoly as the antithesis of competition, its extreme limit is reached only in the case of control of the supply of all economic goods, which might be called a case of pure monopoly in the sense that all competition of substitutes is excluded by definition." Then, the middle ground between the theoretical limits of pure competition at the one end and pure monopoly at the other is the domain of "monopolistic competition", the general case representing a blending of competitive and monopolistic elements. "To discard either competition or monopoly is to falsify the result ..."

The immediate objection to this "solution" is that the "limit" of pure monopoly as so conceived is not

13. op. cit., p. 5.
16. op. cit., p. 63.
analogous to the antipodal limit of pure competition. An actual real world situation resembling, or more correctly approaching, the definition of pure competition could occur in practice or, at any rate, could be imagined to occur in practice; but, as Kaldor has pointed out, the limiting case of "pure monopoly" in Chamberlin's sense — control of the supply of all economic goods — "not only does not exist, it is not even conceivable, since it would conflict with our basic assumptions about the nature of human wants."  

Certainly, to postulate such a situation would assume away the very problem which we have to solve — the problem of relative prices and outputs. In such a case, the revenue of the monopolist "could neither be decreased nor increased as a result of price changes because there would exist no alternative outlet for the income of buyers." Hence, the maximum profits criterion would establish outputs at the point of minimum total cost, i.e. at an infinitesimal level. In any case, there would be no incentive — i.e. profit incentive — to produce more than a fractional output of

17. R. Triffin, op. cit., p. 128n.
one commodity.

Triffin slightly generalizes the Chamberlinian concept of "pure monopoly" so that it defines the seller of a commodity which is neither complementary to, nor competitive with other goods, except for the universal competition of all goods for the consumer's dollar."18 But, again, such a firm would be faced with a constant outlay demand curve which, upon the assumption of maximization of profits, would result in an infinitesimal output by that firm. A similar definition of "pure monopoly" — called "absolute monopoly" — was earlier suggested by Sraffa.19

However, Triffin goes on to concede that "it must be admitted at once that the assumptions needed actually to realize monopoly in its pure form are quite fantastic."20 Because of this, the concept, once postulated, plays little further part either in the scheme of Chamberlin or in that of Triffin. By the same token, it solves nothing. We are still left with the problem of determining where to place the boundaries of the "group". If pure competition is looked upon as a group of sellers of an identical product, where, short of the quite academic concept of

20. R. Triffin, op. cit., p. 133.
"pure monopoly" advanced by Chamberlin or Triffin, do we consider the "group" to vanish as products become more and more differentiated. In his recent writings, Chamberlin's answer seems to be acceptance of the course advocated by Triffin — abandonment of the notion of the "group." He says:

"The individual firm is either isolated or related oligopolistically to others. The group has disappeared from the formulation here given, and with it the concept of 'large numbers', since the individual seller in this latter instance is again correctly described as 'isolated', even in the special case of pure competition."21

Thus, the wheel has come full circle. Leaving aside "oligopolistic interdependence" — a matter which Mrs. Robinson assumed away22 — Chamberlin now arrives at a position indistinguishable from that originally taken up by Mrs. Robinson. Oligopolists apart, all sellers

22. Cf. Mrs. Robinson, op. cit., p. 21: "The demand curve for the individual firm may be conceived to show the full effect upon the sales of that firm which results from any change in the price which it charges, whether it causes a change in the prices charged by the others or not." My emphasis.
are "isolated", even in the case of pure competition, which is precisely Mrs. Robinson's original position, a position which Chamberlin had taken the opportunity to scorn at an earlier date. The only difference seems to be that Chamberlin, having determined that the term "pure monopoly" should not have application to any situation of this world, now describes as an "isolated seller" what Mrs. Robinson calls a "pure monopolist". What's in a name! Certainly, if we insist upon a strictly literal criterion, the one term is scarcely more appropriate than the other.

But this is no solution of the basic difficulty at all. Abandonment of the distinction between a group of firms selling the "same" or "similar" commodities and a "pure monopolist" is to ignore a difference which we are able to recognize in practice, if not yet in theory. Further, the importance of the distinction resides in the fact, as Kaldor has it, that "in the case of polypoly there is (and in the case

23. Quoting Mrs. Robinson to the effect that "every individual producer has the monopoly of his own output - that is sufficiently obvious - and if a large number of them are selling in a perfect market the state of affairs exists which we are accustomed to describe as perfect competition," Chamberlin goes on to express his astonishment: "The individual seller, then, even under perfect competition, is a 'monopolist'!" Monopolistic Competition, Sixth Ed., p. 208. His emphasis.
of monopoly there is not) a group of firms whose prices jointly exert a large influence on the firm's sales.\textsuperscript{24} If, in the case of oligopoly, individual firms may influence the price-output policies of other firms, in the non-oligopolistic group equally significant pressure upon a firm's price-output policy may be exerted by the sum effect of the individual actions of the large number of other firms in the group. Indeed, in the limiting case of pure competition, this pressure is so pervasive that the individual firm can exercise no influence over its selling price. The firm must be purely passive and adaptive.

It is clear that, if we are to resolve the problem of distinguishing between an "isolated seller" or "pure monopolist" from a seller in a group, we need to refer to economic, not semantic, considerations. In earlier theory, a monopolist was a single seller of a commodity different in kind from any other commodity. This fact enabled the monopolist to exercise control over its price. By contrast, a competitive seller, producing only a small part of the total output of a given commodity, could exercise no significant influence upon price. This was the economic significance of the distinction between

monopoly and competition. It is, therefore, to beg the economic question altogether to place both these types of market situation in the same category. Yet, as we have seen, this is precisely what Chamberlin now asks us to do. For him, there is no fundamental distinction required to be drawn between an "isolated seller" for whose product there are no close substitutes and a seller in a purely competitive group for whose product there are many perfect substitutes.

Thus, we find that modern value theory - as distinct from rule of thumb - has not been able to provide us with any criterion by which we may distinguish what one would have thought were fundamental market categories, namely to distinguish between pure monopoly and pure competition. It is to this problem that we shall address ourselves in Chapter III of this Part, where an alternative solution will be put forward.

2. Oligopolistic Interdependence

Meanwhile, we must briefly consider the second problem of market classification which arises out of the assumptions of monopolistic competition - the
problem of "oligopoly". Here we find analogous difficulties to those discussed in the previous section. In the old scheme, an oligopolist was one who produced a sufficiently large share of market output that his price-output policy could significantly affect market price. With commodities different in kind, the determination of the degree of "oligopolistic interdependence" was, more or less, merely a matter of arithmetic. However, once we adopt the position that substitutability between commodities is only a matter of degree, we run into a new set of questions. In this situation, how do we define the "market"? Where should we set the boundaries of the market area? What do we mean by "market price"? And so on.

In short, we are now faced with the fact that, before we may begin to count the number of producers, we have to determine how alike must be the products which are to be counted. To define an oligopolistic situation, as indeed all market categories, we require two criteria: (1) a substitution criterion; and (2) a numbers criterion. In some sense, oligopoly arises where there are a few sellers of good substitutes. This leaves us with the fundamental problem of determining how good must be the substitutes
and how few the numbers.

Theory has never really solved this problem, although it has occasioned considerable discussion and a variety of solutions of greater or less value. Most discussion of the problem is vitiated from the start because of failure to recognize that two conditions are necessary to establish the existence of oligopoly. This fact, no doubt, lies behind the reminder, which R.L. Bishop feels bound to stress as late as 1952, that: "It is significant that oligopolistic interdependence is avoided both when there are many relatively close substitutes and when there is none." In other words, oligopoly may be absent either because of failure of the numbers condition or because of failure of the substitution condition. To concentrate attention on either one criterion to the exclusion of the other is bound to lead to an unsatisfactory result.

It would be premature at this stage to survey the variety of opinion which has grown up around the subject,

25. For example, those attempts made to define oligopoly simply in terms of significantly high cross-elasticities of demand.
although later we will be bound to discuss the main trends of thought in the field. This we will take up in Chapter IV where a detailed discussion of the problem of oligopoly will be undertaken. It is sufficient here to have indicated the nature of the problem.

3. *Towards a New Approach?*

To sum up, then, no-one is particularly happy about the present state of the theory of market relationships, least of all Professor Chamberlin who must be considered to have made the greatest investment in it. Indeed, Chamberlin has recently told us that "there is urgent need for examination of the variety of criteria by which areas intermediate between the firm and the whole economy may also be meaningfully defined."[^27] To this end, "the subject needs to be re-written in terms of that extremely useful concept which originated, I believe, with Mr. Kaldor of cross-elasticity of demand, rather than in terms of the number of sellers in a market."[^28]

The re-writing of the subject would, of course, be a very ambitious programme; and the writer hastens to disclaim any undertaking to attempt it here. However, it is hoped in this Part to essay some first steps in this direction. Unfortunately, however, a "chaos of disagreement" — to use the words of Bishop — has developed around the concept of cross-elasticity of demand, the basic tool which Professor Chamberlin prescribes for us. This, of course, merely reflects the fundamental contradictions in the theory of monopolistic competition itself. Each writer has twisted the notion of cross-elasticity to meet the requirements of his own pre-conceptions about, or his reading of, the theory of monopolistic competition.

Clearly, if cross-elasticities of demand are to be used to elucidate the nature of market relationships (and the author believes them to be very valuable in this respect), then we must have a firm idea of precisely what such cross-elasticities signify. Otherwise, any superstructure of analysis based upon them will be erected on shifting sands. Therefore, it is a necessary preliminary to our investigation of the basic types of market relationships to examine the nature of cross-elasticity of demand.

II. THE NATURE OF CROSS-ELASTICITY OF DEMAND

1. A Substitution or Numbers Criterion?

We concluded the previous chapter with the observation that over recent years a good deal of confusion and inconsistency has grown up around the concept of cross-elasticity of demand. This fact must have forced its attention on, and caused considerable distress to, all those who have attempted to follow the literature of value theory during this period. We are, therefore, all indebted to Professor Bishop for his valuable and timely article on the subject which he published at the end of 1952. In particular, Bishop has rendered us two important services. Firstly, he has provided an admirable survey of the variety of opinion surrounding the concept of cross-elasticity of demand and has brought into the open the deep cleavages

which exist. Secondly, he has recalled our attention to the fundamental fact that "the concepts in question are elasticities and cross-elasticities of demand." 3 In this connection he shows how many writers have attempted to qualify the concept of cross-elasticity by reference to supply factors, and directs our attention to "the profound anomalies that are implied by their half-breed supply-demand elasticities." 4 Unfortunately, however, when Bishop embarks on positive prescription, we find that the basic confusion remains embedded in his own work.

It is convenient to approach the problem of interpreting the meaning of cross-elasticity of demand by reference to the work of Triffin, who provides us with the first systematic application of cross-elasticities to the classification of market relationships. 5 Using the concept of cross-elasticity of demand of firm \( \lambda \) with respect to a change in the price of the product of another firm \( j \), Triffin distinguishes three general categories of market situation: 6

---

isolated selling (culminating in pure monopoly)
when the coefficient \( \frac{p_j \partial q_i}{q_i \partial p_j} \) = 0.7

(2) heterogeneous competition between firms, for which
the coefficient \( \frac{p_i \partial q_i}{q_i \partial p_j} \) takes a finite value.

(3) homogeneous competition between firms, for which
the coefficient \( \frac{p_i \partial q_i}{q_i \partial p_j} \) = \( \infty \).

To illustrate the method by which these definitions are derived, let us consider the argument by which a value of \( \infty \) is established for the case of homogeneous competition:

"First of all, let us take the case where the slightest cut in \( p_j \), \( p_i \) remaining unchanged, or the slightest rise in \( p_i \), \( p_j \) remaining unchanged, reduces to zero the sales (and revenue) of \( i \), driving all its customers toward firm \( j \). The value \( \infty \) then reached by the coefficient gives us an economic indication of the homogeneity of the goods sold by the two firms. In the traditional terminology, the two firms would

7. Triffin points out that the values 0 and \( \infty \) are to be conceived only as limiting cases; hence, correctly speaking, the coefficient will only approach these values.
be said to be selling the same commodity. Depending on the number of firms between which such a relation exists, the case would be classified as duopoly, oligopoly, pure competition. Let us call the general category homogeneous competition.  

At first sight, it may seem strange to find duopoly and oligopoly keeping company with pure competition; but, as Triffin rightly urges, the classification has its rationale in "the rigid connection between the prices of economically homogeneous goods ..." Clearly, then, the coefficient thus derived is intended to be a coefficient of substitution; it is designed to measure the degree of substitutability between the products of firm i and j.

To fully describe market situations, however, we need two criteria: (1) a substitution criterion; and (2) a numbers criterion. So, having described the degree of substitutability between products in terms of price cross-elasticity of demand, Triffin provides us with a second coefficient, what Bishop has called a "quantity cross-elasticity", by means of which the existence of "oligopolistic interdependence" may be

8. Triffin, op. cit., p. 103.
determined. This is his numbers criterion.

Consideration of Triffin's numbers coefficient more appropriately belongs to Chapter IV, where we shall be concerned with the problem of oligopoly in some detail. In this chapter, we shall confine our attention to the validity of price cross-elasticity as a substitution criterion. In particular, we shall show that, subject to a qualification in respect of the direction of the price change, Triffin's use of price cross-elasticity as his substitution criterion is unexceptionable. However, in the next chapter attention will be directed to certain implications of this which Triffin failed to notice.

In his use of price cross-elasticity, Triffin, with most writers, proceeds on the assumption that cross-elasticity of demand will be the same whichever the direction of the price change. However, I believe that most readers will accept, without quotation of chapter and verse, that invariably discussion has proceeded, either explicitly or tacitly, in terms of a price-cut. For the rest, it seems to have been automatically assumed that the same characteristics would attach to a rise in price. However, a price-cut involves the reduction of one price in relation to all others, whereas a price-rise implies the reduction of all other prices in relation to one. Where substitution in all directions is not continuous
(see the latter part of footnote 27 below), the asymmetry between a price-cut and a price-rise will not generally hold, a matter which will be fully discussed in Chapter III which follows.

Therefore, for the purpose of the present discussion, we define price cross-elasticity in terms of a cut in price, i.e. as the ratio of the percentage loss of output by any firm $j$ to the initiating percentage reduction in price by another firm $i$, all other prices being assumed unchanged. Then, $E_{ij} = \frac{p_i \delta q_j}{q_j \delta p_i}$, where $p_i$ is the price of firm $i$ and $q_j$ is the output of firm $j$, and where $\delta p_i$ is considered in the negative sense. This cross-elasticity is, then, the same as Triffin's first (substitution) coefficient, except that it specifically excludes the case of a rise in $p_i$. It is also the coefficient which most people have in mind when they speak of cross-elasticity of demand without qualification.

It should be noticed that this definition of cross-elasticity of demand makes no reference to the supply restrictions which Bishop rightly deplores. Moreover, it is clear that this cross-elasticity, freed from supply qualifications, is a coefficient of substitution. Indeed, this was precisely the purpose of its original introduction by Kaldor.10 The aim was to provide a more satisfactory

criterion of substitutability than that of physical description of the products. It was to meet the difficulty, as Triffin puts it, that "the competition may be keener between Ford and Rogers Peet than between Ford and Rolls-Royce." And certainly this conception of cross-elasticity has persisted in the textbooks, at least in certain chapters of them. Thus, for example, Stigler tells us: "The cross-elasticities provide a convenient index of the readiness with which consumers substitute one commodity for another." Yet, almost from the start attempts were made to make price cross-elasticity serve as a numbers criterion. Moreover, Professor Chamberlin has, at one stage at least, lent his authority to this position. As Professor Bishop points out: "Chamberlin, by contrast with Triffin, thinks that cross-elasticity with respect to price, $E_{ij}$ or $E_{ji}$, distinguishes merely (1) non-oligopolistic or 'isolated' selling from (2) oligopolistic or 'non-isolated' selling, according as this coefficient is (1) zero or (2) greater than zero." "To Chamberlin, this type of cross-elasticity significantly different from zero is an unfailing sign of oligopoly..." Other writers have sided with Triffin on

the one hand, or with Chamberlin on the other, or have ranged—rather like cross-elasticities—over all points in between. 15

But perhaps the extent of the confusion can best be illustrated by reference to the writings of Mr. Kaldor. In his "Market Imperfection and Excess Capacity" of 1935, he says on the subject of market classification: "'monopoly' and 'perfect competition' appear as the two limiting cases, where the 'cross-elasticities' are zero or infinite, respectively; and there can be little doubt that the large majority of industrial producers in the real world are faced with imperfect markets in this sense." 16 Clearly, Kaldor here uses cross-elasticity of demand as a substitution criterion; and it is precisely this conception of price cross-elasticity which has come down to Triffin. However, some years later, in reviewing Triffin's Monopolistic Competition and General Equilibrium Theory, Kaldor has this to say: "His principle for classifying market situations (the 'nature of external interdependence') is the nature of the cross-elasticities of demand (the elasticity of sales of one firm with respect to the change in price by some other firm). Under pure competition, these cross-


elasticities must be either infinite or zero (according as the 'other firm' is situated in the same industry or some other industry); under pure monopoly, they must all be zero; under imperfect competition they must have a positive finite value - in conditions of 'polypoly' they are small, under oligopoly they are large. So far I am fully in agreement...

A brief examination of this statement shows that, practically throughout, Kaldor is treating cross-elasticity of demand as a substitution criterion, with exactly the same sense as in his earlier writings on the subject. Yet so pervasive has become the association of cross-elasticity with numbers that, towards the end, he suddenly switches over to use cross-elasticity as a criterion of numbers: "in conditions of 'polypoly' they are small, under oligopoly they are large." Notice, moreover, the implications of this. For pure competition cross-elasticities are infinite, for oligopoly they are high; for pure monopoly they are zero; for polypoly they are small. Thus, it seems, the


18. Actually, Kaldor is here reading into Triffin his own pre-conceptions, because Triffin does not use price cross-elasticity, but a quite different coefficient, to distinguish between his "circular" and "atomistic" categories. This, by the way, is a further illustration of the strength of the pre-conception that price cross-elasticity is a function of numbers.
greater the "oligopolistic interdependence", the nearer is the approach to pure competition; contrawise, the slightest imperfection in an otherwise purely competitive market results, not in polypoly as one might have expected, but in oligopoly, unless one is prepared to believe, as indeed some writers have inferred, that the introduction of the slightest degree of imperfection into the market causes cross-elasticity to skip suddenly from $\infty$ to zero.\(^{19}\)

Kaldor here is trying to make price cross-elasticity serve as a substitution criterion and a numbers criterion. But, as we have pointed out above, the classifications of markets is a two-dimensional affair. It requires two independent criteria to define a given situation. One coefficient can no more define a market relationship than one co-ordinate can define a point in a plane. Cross-elasticity of demand, then, must be either a coefficient of substitution or a coefficient of numbers. It cannot be both. This follows immediately from the fact, which I take

---

19. Cf. "Even if we...interpret the coefficient as having a value of infinity for pure competition, its value skips discontinuously to zero with the slightest departure in terms of heterogeneity; it can only be made to proceed through a stage of finite values by operating on the scale of numbers." - Chamberlin quoted by Bishop, *loc. cit.*, p.787. Bishop goes on to add: "In an earlier draft of this paper, I must confess that I represented $E_{ij}$ as infinite under pure competition but necessarily small under differentiated competition."
it no-one would deny, that we can have an infinite rate of substitution associated with small numbers, a situation of pure oligopoly, or an infinitesimal rate of substitution associated with large numbers, a situation we are accustomed to call, somewhat paradoxically, "pure monopoly".

Although cross-elasticity may not be a measure of both substitutability and numbers, it may, of course, not be a unique measure of either. And it is just this conclusion at which Professor Bishop arrives. He says: "The truth of the matter, of course, is that the value of the $E_{ji}$ depends on both the scale of numbers and the scale of product homogeneity-heterogeneity; and consequently it is not a sure clue to either separately." 20 While this displays a commendable spirit of compromise, the conclusion is not correct. Indeed, if it were correct, the concept of cross-elasticity of demand, not being a unique measure of anything, would have little purpose to serve and might be rejected out of hand.

Professor Bishop does, however, make it serve in his system of market classification; but he is only able to do this by considering its value in relation to the value of own-elasticity of demand. Thus, he postulates as his "numbers equivalent" the relationship $n_i - 1 = -\frac{E_{ii}}{E_{ji}}$.

This works out superficially - insofar as it does give us a numerical criterion of numbers, weighted for asymmetries - only because Bishop makes an equal and opposite error in his interpretation of the significance of own-elasticity of demand. Nevertheless, the confusion involved necessarily prevents his carrying the problem of market classification very far forward. Since the matter is of fundamental importance to our latter analysis, we must now turn to consider the exact nature of the relationship between cross-elasticity and own-elasticity of demand.

2. Cross-Elasticity and Own-Elasticity of Demand.

The cross-elasticity of demand $E_{ji}$ is a measure of the rate at which $j$ customers will move to product $i$ in response to a reduction in the price $i$. From the indifference function of each individual $j$ customer, we may deduce the rate at which each such customer will substitute product $i$ for $j$. By summation, in the manner of the Marshallian demand schedule, we may construct a cross-demand schedule ($q_j$ plotted against $p_i$) for the product $j$ with respect to the price of product $i$. This will give us the rate at which $j$ customers as a whole will substitute product $i$ for product $j$.

Cross-elasticity of demand for an individual consumer, which we may denote by $E$, is given by $E_{ji} = k_i \sigma - k_i \bar{Y}_j$, where $k_i$ is the fraction of income spent on $i$; $\sigma$ is the elasticity of substitution between $i$ and $j$, and $\bar{Y}_j$ is the income elasticity of demand for product $j$. (See J.R. Hicks and R.D.G. Allen, "A Reconsideration of the Theory of Value", Economica (N.S.), May 1934, I, pp.201-2.) Then, if we neglect the income effect as we
But, whereas cross-elasticity of demand is a function of the rate at which consumers will substitute the one commodity for another, own-elasticity of demand depends, not only on the rate of substitution between individual products, but also upon the number of substitutes available. After all, as Bishop himself stresses, own-elasticity of demand is a measure of "total substitutability", i.e. the substitutability between the "own" product and the collectivity of products which make up its competitive field. It

Actually, this is not the whole story, because in conditions of pure monopolistic competition, if I may coin an expression, the process of substitution will not generally be continuous, e.g. as in the case of substitution of one brand of toothpaste for another. Rather, in this situation, there would be a critical price ratio at which total expenditure on the one product (abstracting from income effects) would be shifted to the other. In terms of indifference curves, this case would be represented by a curve of infinite elasticity of substitution, but with a slope (marginal rate of substitution?) different from unity. Here, resort must be had to indirect methods to translate consumer preferences into cross-elasticities. Incidentally, it will be noticed that these cross-elasticities (finite) are more appropriate indices of substitutability than the elasticities of substitution which are infinite. However, this is a digression. (The matter will be discussed in detail in Chapter III.) My purpose here is only to show that cross-elasticity is directly related to the theory of consumer choice which, of course, is essentially a study of substitution. After all, the concept of cross-elasticity was born in the London School which, at the same time, was also nurturing the neo-Paretian analysis.
is surely obvious that any change in the variety or number of products within this competitive field will affect the value of own-elasticity of demand.

Let us pursue this question of the relationship between cross-elasticity and own-elasticity of demand further. For this purpose, we need to make a slight change in our definition of elasticity and cross-elasticity. Usually, elasticities are expressed as a ratio between a percentage price change and a percentage output change. However, under differentiated competition, units of output, as between firms, are not strictly commensurable. Therefore, instead of output, we now insert sales, i.e. output in value terms; and to signify the change we substitute in our elasticities $Q_1$ for $q_1$. We further suppose that the total value of sales lost by any firm $j$ in response to a unilateral price-cut by $i$ will accrue as additional sales to $i$. (This assumption is equivalent to neglecting the income effect of the price change.) Then, on this basis, own-elasticity of demand may be readily expressed in terms of the array of cross-elasticities:

$$ -E_{ii} = \frac{Q_1 E_{ij}}{Q_1} + \frac{Q_1 E_{ki}}{Q_1} + \cdots + \frac{Q_1 E_{ni}}{Q_1} + \frac{Q_1 E_{oj}}{Q_1} + \frac{Q_1 E_{pi}}{Q_1} + \cdots + \frac{Q_1 E_{sj}}{Q_1} $$

where the $E_{ij}$ denote cross-elasticities of the first order.
of importance and the $E'_{o1}$ denote those of second-order
significance.  

For convenience, we assume: (1) that the first-order
cross-elasticities are all equal, i.e. the $E_{ji} = E_{ki}$, and
that the second-order cross-elasticities are all zero; and
(2) that all products are substitutes, i.e. none are com-
plementary, so that we may take the arithmetical value of
elasticities without ambiguity, and so speak of larger or
smaller elasticities without regard to sign. Then, on these
assumptions, we may rewrite (1) in the simpler form:

$$E_{ii} = \frac{Q_j + Q_k + \cdots + Q_n}{Q_i} E_{ji} \quad (2)$$

or

$$E_{ii} = \frac{\text{Non-}i \text{ Output}}{\text{Output of } i} E_{ji} \quad (3)$$

Now, if the concept of numbers means anything in relation
to a group, it expresses a relationship between firm output
and total output. If numbers are large, the firm's share

---

22. Cf. Stigler, op. cit., p. 235: "The group may be defined
as all firms whose cross-elasticities are greater than
some constant. This constant can best be chosen in any
particular problem by ranking the cross-elasticities and
considering a group to end where a considerable gap appears
in the array of cross-elasticities." Interestingly
enough, most economists would accept this procedure. But
many do not seem to realize that in accepting this tech-
nique, based as it is on the notion of a gap in the chain
of substitutes, they are accepting cross-elasticity as a
valid substitution criterion.
in group output is small; and conversely if numbers are few. Then it is easy to see from (3) that any change in firm i's share of group output will, other things remaining the same, involve a corresponding change in own-elasticity of demand, $E_{ii}$. Take, for example, a decrease in numbers, i.e. an increase in i's share of the market. This must necessarily involve a decreased elasticity of demand for firm i, since, ex hypothesis, the size of the market from which sales may be attracted has contracted in relation to i's own output. Although own-elasticity of demand depends on numbers in this fashion, cross-elasticity does not. This is so because the sales lost by any firm j in response to a price-cut by another firm i are related to the output of firm j. Hence, so long as the proportions between j customers of various degrees of sensitivity are unchanged, any increase (or decrease) in j's share of the market will affect the numerator and denominator of the cross-elasticity ratio in the same proportion.

The reader will not have failed to notice that our equation (3) expresses the same relationship as Bishop's so-called "numbers equivalent" which he writes $-E_{ii} = (n_i^{-1})E_{ji}$. Also, it will be clear that the essential difference between Bishop's interpretation of this relationship and that given here is that, whereas Bishop treats cross-elasticity of demand as the dependent variable when $n_i$ changes, for us it is own-elasticity which is dependent. Indeed, Bishop's
argument always proceeds on the basis: Given own-elasticity of demand, then cross-elasticity will be large or small according as numbers are small or large. But, it is fair to ask, by whom or what is own-elasticity of demand given—apart from by the textbooks—except in relation to the closeness and the number of the individual products available as substitutes, the "total substitutability" as Bishop describes it. It is cross-elasticity which measures the closeness of the substitutes and is therefore the natural substitution criterion. Own-elasticity is then determined by the order of the cross-elasticities and the number of them.

This argument is best demonstrated by considering the case of homogeneous competition. In this case, cross-elasticity of demand will be the same whether there are two firms (pure duopoly) or 2,001 firms (pure competition), since cross-elasticity in both situations will be equal to the reciprocal of the percentage price-cut. But own-elasticity of demand for the pure competitor will be much higher than for the pure duopolist. Although, of course, both elasticities may be made as large as we like, if we make the price-cut small enough, elasticity of demand under pure competition will always be much the higher for any given price-cut, however small. Indeed, if the firms in each

23. Since for homogeneous competition \( \frac{\partial Q_j}{Q_j} \) will always be equal to unity.
group were symmetrical, own-elasticity in the pure competition case would be just 2,000 times that in the duopoly case.\textsuperscript{24} It follows as a corollary that, with any given price-cut, we can always make own-elasticity of demand, but not cross-elasticity, as large as we like by making our firm's share of total output small enough, i.e. by making numbers large enough.

At the other end of the scale of substitutability, our interpretation of cross-elasticity is also the natural one. By contrast, Bishop's conception of the relationship between cross-elasticity and own-elasticity leads to paradoxical results when applied to "pure monopoly". In this case, own-elasticity of demand will be finite, while cross-elasticities will be very small or zero. Hence, their ratio - Bishop's "numbers equivalent" - will be very large indeed. And this Bishop frankly confesses, for he says that in this case the numbers equivalent "must be more or less uniformly high with respect to all other firms, without exception."\textsuperscript{25} At best, this can hardly be said to be a happy choice of terms. Triffin, on the other hand, would describe this situation as one in which the cross-elasticities of demand (measuring

\textsuperscript{24} It will be remembered that we have excluded supply conditions from our definition of cross-elasticity. Hence, a price-cut by a purely competitive firm must be considered, conceptually, to capture the whole market.

\textsuperscript{25} Op. cit., p.800.
the substitutability of products) are more or less uniformly small. Surely this is a more apposite description. This becomes obvious as soon as we put the question: Is pure monopoly non-oligopolistic because of the absence of close substitutes (failure of the substitution condition) or because of large numbers (failure of the numbers condition)?

It should be stressed that nothing in the foregoing argument depends on Mrs. Robinson's notion of products "coming closer together" or of new firms "coming in between" old firms, a conception which has been the subject of debate from time to time.26 Mrs. Robinson, it will be recalled, argued that:

...if the new firms were set up, so to speak, in between the old firms (either geographically or in respect of special qualities which appeal in various degrees to different customers)...the difference, from the point of view of buyers, between any one firm and the next would thus be reduced, the customers of each firm would become more indifferent, and the elasticity of demand would be increased.27

This is a valid argument.28 But it is important to recognize that there are two effects involved here: not only an

26. In particular, during the 1938 debate between Chamberlin and Kaldor in the Quarterly Journal of Economics.


28. I.e., given Mrs. Robinson's condition: "If the new firms were set up in between the old firms." The fact that an increase in the numbers of firms may involve merely an increase in area and not in density - a point made much of by Chamberlin - is irrelevant to the particular argument in question.
increase in numbers per se (the numbers effect), but also a change in the distance between individual products, as is indeed stated in the italicized phrase (the substitution effect).

The conceptual distinction between the substitution and numbers effects becomes clear when we consider that the notion of products "coming closer together" in no way depends on increasing numbers. It could equally well come about, for example, as a result of an all-around reduction in transport costs or of a technically inspired move towards greater standardization of existing products. In these cases, numbers would remain unchanged; but cross-elasticities, and hence own-elasticity, of demand would increase. On the other hand, an increase in numbers need not logically imply any change in the degree of substitutability between individual products, as for example when the entry of new firms does not alter the "distance" between any two products, as indeed will always be the case with homogeneous competition.

However, in the real world of differentiated products, a change in numbers will almost certainly involve some rearrangement of consumers' preferences; and hence some

---

29. Strictly speaking, an increase (say) in the number of firms will not of itself alter consumers' preferences as between the old products. Rather it will establish between the old firms and the new firms, and between the new firms themselves, a new set of cross-elasticities, which may be larger or smaller than existing cross-elasticities. (In Mrs. Robinson's example they will, of course, be larger.) Nevertheless, the change in numbers will imply some redistribution of consumers between firms and this will, in general, affect the pre-existing cross-elasticities also.
change in cross-elasticities of demand. This effect will usually be in the same direction as, but remain conceptually distinct from, the numbers effect. The important fact is that cross-elasticities change, not in virtue of the change in numbers as such, but because of the change in substitutability which will usually be associated with a change in the product-structure of a group.

It is from confusion of these two effects, it seems to me, that has arisen the "third force" view, as represented by Bishop, for example, that cross-elasticity is a function both of product substitutability and numbers. The clear implication of this view is that the effect of a change in numbers upon cross-elasticity cannot be theoretically determined, but is an empirical question only to be answered by reference to the facts of any given situation. But this approach is theoretically misleading. In so far as a change in the number of firms (products) changes the opportunities for substitution and the pattern of consumer attachments, it changes the pattern of cross-elasticities. But the important point is that it changes cross-elasticities via changes in substitutability or, if you like, via changes in the complex of consumers' preferences which theory has almost universally taken as data. However, there is no essential connection between such changes and changes in numbers in the sense used throughout this paper and by Bishop. This latter conception of numbers relates
to the ratio of firm to group output, and changes in this ratio may clearly occur independently of any change in consumers' scales of preference. 30

3. The Origin of the Confusion of Cross-Elasticity with Numbers.

Since the matter is of fundamental importance for formal analysis, perhaps it is well to attempt to pinpoint the source of the long-standing confusion of cross-elasticity with numbers. In the old scheme — when commodities were different in kind, not merely in degree — numbers were determined as a simple arithmetic relationship between the capacity of any one firm and the total market supply of the commodity in question. Thus, if a firm's capacity were small compared with total output, the situation would be one of pure competition, because any additional supply it could add to the market would, when "spread over" that market, affect other suppliers only to a negligible degree. On the other hand, if a firm's potential capacity were large in relation to total supply, oligopolistic implications would arise.

30. Perhaps the distinction is best made clear by an illustration. Imagine that a sectional increase in i customers' incomes results in a doubling of purchases from firm i, the outputs of all other firms in the group and the whole complex of consumer preferences remaining unchanged. Own-elasticity of firm i will approximately halve. By contrast, the cross-elasticities will not be affected, except in so far as the alteration in income shifts the margin at which substitution takes place. Since, when we speak of substitutability, we mean substitutability at the margin, there is clearly a sense in which we may take cross-elasticity, but cannot take own-elasticity, as data given by the total map of consumers' preferences.
This conception was adapted by Chamberlin to the general case of monopolistic competition. Thus, in a formulation which he still finds "unobjectionable", he defined the "large group" as follows:

Specifically, we assume for the present that any adjustment of price or of "product" by a single producer spreads its influence over so many of his competitors that the impact felt by any one is negligible and does not lead to any readjustment of his own situation.

It is but a short step from here to those "half-breed supply-demand cross-elasticities" which Bishop denounces, a step, incidentally, that Chamberlin himself seems to have taken in his more recent work. It would take me too far from the present purpose to discuss here the validity of Chamberlin's formula as a numbers criterion. But I do wish to dispute its relevance to the concept of cross-elasticity of demand.


33. For example, "Measuring the Degree of Monopoly and Competition", a contribution to Monopoly and Competition and Their Regulation, E.H. Chamberlin, Ed. (London, 1954), esp. Pt. 4, pp. 262-67. More recently, however, he has retreated from this position and reverts to his original definition of isolation which I have quoted in the text above. He now says, "It should be noted that zero cross-elasticities, though sufficient, are not necessary to isolation in this sense; and also that in the vexed case of pure competition, however one may rule as to the value of the cross-elasticity coefficient, a single seller is isolated in the sense described." (Comment on Bishop's article, op. cit., p. 913.) In short, he now admits there is no necessary relation between his large-small numbers criterion and cross-elasticity of demand.
Let us consider Chamberlin's formula in terms of a price-cut. Now it is clear that we must invoke some sort of supply restrictions if we are to make any sense at all of his conception of the effect of a price-cut being "spread over" a given number of rival firms. This is so because, theoretically, we can make the increment of sales to the price-cutting firm, and therefore the amount to be taken from rivals, as large (small) as we like if we make the price-cut large (small) enough. This follows immediately from the fact that, ex hypothesi, the relationship between a price change and a quantity change under conditions of monopolistic competition is always a finite one. Therefore, the increment of sales accruing to a price-cutting firm, and hence the amount taken from each rival, cannot be determined, unless appeal is had to some consideration of supply.

Now it is possible to argue that the amount by which a "large group" seller may increase sales is limited by supply factors. But, the smaller is a firm's potential capacity, the smaller need be the price-cut to take up this capacity; and since cross-elasticity of demand is the ratio of the quantity change to the price change, its value will not be affected by any considerations of supply.

Rather, in conditions of monopolistic competition, if any firm i cuts price, it will pull sales from firm j at the

---

34. Bishop is overgenerous in saying that attempts to qualify cross-elasticities with supply restrictions have been confined to homogeneous competition. *Op. cit.*, p. 781.
rate given by the cross-elasticity of demand $E_{ji}$; and the value of $E_{ji}$ depends on the rate at which $j$ customers will exchange product $i$ for product $j$. Also, firm $i$ will pull sales from firm $k$ at the rate $E_{ki}$; and so on. Now it cannot be emphasized too strongly that, income effects apart, the movement of buyers from $j$ to $i$ is quite independent of the movement of buyers from $k$ to $i$, from $l$ to $i$, and so on. Otherwise, why does not the volume of sales gained from a price-cut "spread over" all other firms? Or why does the spread fall more heavily upon one firm than upon another? We cannot deduce the answer to these questions from own-elasticity of demand. In other words, own-elasticity of demand fails as a substitution criterion. But we do know that, for any given price-cut, the better the substitute, the heavier will be the sales loss. And it is precisely this of which cross-elasticity of demand is a measure.

It is perfectly true, of course, that, with given rates of substitution, the greater the number of competitors, "the less the price reduction necessary to secure any desired increase in the sales of $A$...And...the less the effect of any given increase in the sales of $A$ on the sales of rival products..."\(^{35}\) However, since the effect of (say) greater numbers is (1) to reduce the price-cut required to achieve

\(^{35}\) G.J. Stigler, \textit{op. cit.}, p.235.
any predetermined result; and (2) to reduce the amount which it is necessary to take from any individual competitor, the cross-elasticities of demand are unaffected, because the numerator (percentage sales loss) and the denominator (percentage price-cut) are changed in the same proportion. Hence, cross-elasticity of demand is independent of numbers.

Since the idea of cross-elasticity as a measure of substitutability seems simple and straightforward enough, or at least so it has appeared to a good many writers, it may well be asked why its general adoption has been so strenuously resisted. The answer, I believe, is to be found in the implications which stem from its acceptance.

The argument advanced in this paper is essentially the same as that put forward by Kaldor in the 1938 controversy with Chamberlin. During that debate, it will be recalled, Kaldor regarded as "crucial" the proposition that "a shift of the demand curve to the left will increase the elasticity of demand at the equilibrium level of output." And since such a shift to the left is associated with an increase in the number of firms, the "curves become more and more elastic with an increase in numbers." Chamberlin rejected this proposition and reiterated his opposition to the idea that

37. Loc. cit.
"differentiation of the product (in my sense) may be exorcised by the process of merely increasing the number of firms."

Chamberlin was perfectly correct in asserting the latter, as he was wrong in rejecting the former, for it is clear that, no matter how small may become the output of a firm as the result of increasing numbers, there may still be a proportion, and even a constant proportion, of customers who will be prepared to continue paying something more for the product of that particular firm than for any of its substitutes.

Product differentiation remains; and upwards elasticity of demand, i.e. elasticity in respect of a price-rise, does not necessarily increase as output contracts in relation to group output. Indeed, as output contracts the firm may be left with only a hard core of its most loyal customers. Elasticity to a price-raise may actually decrease.

Thus, we arrive at the position where, as a firm's output contracts in relation to group output, its downward elasticity, $E_d$, increases (vide Kaldor), while its upward elasticity, $E_u$, may be constant or even decreasing (vide Chamberlin). And both points of view are correct. The seeming contradiction, which has persisted from the 1938 debate to Bishop's work in 1952, is only a contradiction so long as we cling to the pre-conception, which has no real basis except in the economist's

penchant for simple stable equilibria, that $E^d$ should approximate to $E^u$, i.e. that the demand curve under differentiated competition, should necessarily be smooth. This is a matter we will meet again in the chapter which follows.
III. THE LARGE GROUP: THE TRANSITION FROM THE
COMPETITIVE GROUP TO PURE MONOPOLY*

1. The Group and the Concept of Numbers

In the previous chapter, we concluded that Triffin's
use of cross-elasticity of demand as a substitution
criterion is, subject to the qualification stated there,
unexceptionable. At this point, we must depart from
Triffin, for in this chapter we shall be concerned with
the group and with numbers. Triffin, on the other
hand, does not discuss the group at all, except insofar
as was necessary to reject it. "In the general pure
theory of value", he says, "the group and the industry
are useless concepts."¹ "Moreover, when this concept
is abandoned, the criterion of numbers is left in
mid-air."²

Our purpose in this chapter is, then, to rescue the
concept of numbers from mid-air. The reason why this
is necessary has been outlined briefly in Chapter I.
Without some concept of the group, and the associated

* This Chapter is an expansion and development of an
argument previously published under the title, "A
Kinked Demand Curve for Monopolistic Competition*;
Economic Record, May, 1953.
1. R. Triffin, op. cit., p. 89.
2. R. Triffin, op. cit., p. 100.
notion of numbers, we have no analytical way of
distinguishing "pure monopoly" or "isolated selling"
from sellers in a group and, ultimately, from the
pure group type which we call pure competition. The
difference in the degree of control over price - the
significant economic distinction - which distinguishes
pure competition from "pure monopoly" can only be
explained in terms of groups and numbers.

Nevertheless, it is not wished to deny that there
is a good deal of substance in criticisms, including
Triffin's, which have been levelled against traditional
definitions of the group. In earlier theories, the
group is implicitly delineated in the definition of a
"commodity". A group is merely a collection of firms
selling the same (or a similar) commodity. Numbers
are determined by a simple process of counting the
firms falling within the group. Once, however, we
recognize the essential heterogeneity of commodities,
this schema falls to the ground. At first sight this
may not indeed be obvious. Why, it may be asked, is
it not possible, if we wish to consider groups of firms
and to describe them in number terms, to define the
group as a collection of firms linked by significant
cross-elasticities? Cross-elasticity being a criterion
of substitutability between products, we would in this way demarcate a group of more or less good substitutes. Could we not then determine the number of firms by simple counting, weighting firms according to output if it is desired to take into account the relative importance of individual firms?

This notion of bounding the group at the point where a gap in the chain of substitutes appears has, of course, often been proposed; and, as a rule of thumb in fairly clear-cut cases, this method serves well enough. Indeed, we used it in Chapter II for the purpose of illustrating the relationship between cross-elasticity and own-elasticity of demand. But theoretically it will not do, for three main reasons. The first and obvious one is that in many cases there may be no clear and unambiguous break in the chain of substitution. Secondly, in the group of the Marshallian tradition, there resides a factor (a property attaching to the "commodity") which is common to all firms in the group. This common factor is
derived from the substitutability in use of the products in question. But cross-elasticities of demand cannot define a group in this fashion. They cannot determine the existence of a common factor resident in a number of commodities, because they represent a relation which exists between two products taken in isolation. Thus, product A may be significantly related to product B, which in turn may be related to product C; but there may be no relationship at all between the products A and C.

It may be objected that this could easily be resolved by considering the cross-elasticity between A and C. But, even if this were significantly large, it could quite well reflect an altogether different demand characteristic from that existing between B and C. A quite distinct sub-set of customers could be involved. In other words, C may be in a group relationship with A because of a quite different

3. Cf., for example: *For some purposes it may be best to regard Chinese and Indian teas, or even Souchong and Pekoe teas, as different commodities; and to have a separate demand schedule for each of them. While for other purposes it may be best to group together commodities as distinct as beef or mutton, or even as tea and coffee, and to have a single list to represent the demand for the two combined ...* - Marshall, Principles, 8th Ed. p.100n.
substitution-in-use characteristics from those operative in its relationship with B. C may well be, as it were, selling in two distinct "markets" simultaneously. Perhaps this point can be made clear with an illustration. Let us revert to Triffin's striking, if somewhat hyperbolical, example which we quoted in the previous chapter, viz. that "the competition may be keener between Ford and Rogers Peet (a tailoring firm) than between Ford and Rolls-Royce." It is clear that any substitutability between Ford and Rogers Peet will represent a quite different substitution in use - "conspicuous consumption", perhaps - from that which would be involved in substitution between Ford and Rolls-Royce - transport, presumably. Almost certainly the customers who would move in the one direction in response to a price-change would be quite different from those who would move in the other direction. We have, as it were, a situation of triangular chain relationship of consumer preferences. The classical conception of the group, in which every firm within the group is related to every other firm and all firms are selling, or potentially selling, to the same group of customers, must be distinguished from

chain-relationship, which may, of course, be circular, giving something of the appearance of a group. 5

The third reason why cross-elasticities are, by themselves, theoretically inadequate to define the group is the problem of deciding what value of such cross-elasticities should be considered significant. To say that any value different from zero is significant is to rule out at once the possibility (as a real world case, at any rate) of "pure monopoly", as Kaldor has demonstrated. 6 On the other hand, to fix upon some specific positive value as being significant would be purely arbitrary. At this point, our real world cases would jump suddenly from firms considered as competitive within a "large group" context to firms each of which is conceived as a "single seller".

Again, on this basis, the boundary of any particular group and the number of firms to be included in it would be quite arbitrary. We would be drawing a line where none exists in the continuum of reality. Fundamentally, of course, our difficulty arises from

5. The full significance of this particular argument will, it is hoped, become clear later in this chapter.
6. N. Kaldor: "Professor Chamberlin on Monopolistic and Imperfect Competition", loc. cit., p. 526, as quoted in Chapter I, p. 11.
the fact that, strictly speaking, we can only count like things. This is, indeed, the basic premise of Triffin's thesis. However, Triffin's way out—abandonment of the group and the concept of numbers—7—is no solution. It turns away from one difficulty to run at once into another or, more correctly, a different facet of the same fundamental puzzle. Our immediate task, then, is to explore an alternative way around the difficulty. In so doing we will be forced to face implications for the theory of price formation going much beyond the problem of mere classification.

2. The Two Conceptions of Numbers

Earlier, we have insisted that market relationships must be described in terms of two variables, viz.

---

7. In fact, Triffin did not, and could not, wholly abandon the idea of numbers, as he had to face the problem of oligopoly. In this connection, he postulated a coefficient of "circularity" to isolate the oligopoly case. But this coefficient is purely definitional. At no stage does Triffin postulate the conditions of demand or supply which determine the value taken by his circularity coefficient. It is only this omission that he is able to avoid bringing numbers into his scheme, as small numbers, in some sense, are one of the conditions of supply necessary to a situation of oligopoly. (For further discussion of this matter, see Chapter IV following.)
(1) substitutability, and (2) numbers. The theoretical consequence of this is that we can derive two distinct conceptions of numbers, according as we hold the one or other variable constant while allowing the second to change.

Let us take the case of 1,000 firms selling homogeneous products, a situation we are accustomed to call "pure competition". Now gradually allow the substitutability of products (i.e. cross-elasticities) to fall. While substitutability is finite but still high, we have 1,000 sellers of imperfect substitutes, a situation of "polypoly" in traditional terminology. Allow substitutability to fall further so that cross-elasticities approach zero. We then have 1,000 "isolated sellers" in Triffin's terminology. At some stage as substitutability decreases, we must arrive at a point where — in traditional terms — the "group" ceases to exist. Put in another way, we may say that we will approach a position where the intra-group cross-elasticities are not significantly higher than cross-elasticities with some firms which were originally outside the group. There is now, if you like, no marked gap in the chain of substitutes. It follows that, as intra-group cross-elasticities become smaller and
smaller, more and more firms will be encountered with which cross-elasticities are of the same order as those now operative between our original firms. If we look upon "pure monopoly" as some sort of limit as cross-elasticities approach zero, then the number of firms of the same order of interdependence (i.e. insignificant interdependence) becomes very large indeed.

This, then, is one conception of numbers. It is, indeed, the conception of numbers represented by Bishop's "numbers equivalent" which, we saw in the previous chapter, he writes \( n_i - 1 = -\frac{E_{ii}}{E_{ji}} \); and, therefore, as we approach "pure monopoly" (\( E_{ii} \) finite, \( E_{ji} \) infinitesimal) \( n \) becomes very large. The change in numbers from "pure competition" to "pure monopoly" in terms of this conception may be formalistically represented as in Figure 1.

**FIGURE 1**

The Group

---

"Pure Monopoly"
We arrive at the other, and more familiar,
conception of numbers when, beginning as before with
1,000 sellers of homogeneous products, we now hold
substitutability constant, while varying numbers. Then,
if we proceed to "knock out" firms - say by amalgamating
them in pairs - sooner or later we will be left with
only a few firms which we call "pure oligopoly" and
finally with just one, "pure monopoly". Formalistically,
this conception of numbers can be represented as in
Figure 2.

Clearly, it is this conception of numbers which is
appropriate to the older value theory of the tradition
of Marshall and the Lausanne School and followers, for
whom commodities are either the same or are different.

The difficulty with both these conceptions of
number; once we adopt the relativist position of
monopolistic competition when commodities are merely
different in degree, is that each can, in these conditions, only reflect a partial aspect of the complexity of the real world, since here we will find variation in substitution and numbers simultaneously. In other words, we may expect to find every manner of combination between the scale of numbers and the scale of substitutability. What we want, if we are to describe the transition from the group to the isolated seller, is some form of numbers index which will be able to embrace simultaneously variations in substitutability and in numbers (in the arithmetic sense).

As an approach to establishing an appropriate index, it will be helpful to look again at the two conceptions of numbers which we have outlined. It is helpful to reconsider these as density concepts. Any group-number scheme can, of course, be looked at in this way. Numbers must relate to an area; and it is the "group" which defines the area.

In terms of the second conception of numbers, our 1,000 sellers of homogeneous products may be described as 1,000 sellers located at a point. With pure oligopoly we have a few sellers at a point, with pure monopoly one and only one. Each point would then represent a unique "market". The trouble, however, is that as soon as we allow the slightest degree of product
differentiation to appear sellers will cease to be concentrated at a point. Each seller will now have a unique point, although such points may lie very close together. If we stick to the old conception of density (or numbers), such sellers of slightly differentiated products would each have to be described as "isolated sellers" or "pure monopolists". This is the essence of Triffin's position, one which we have seen Professor Chamberlin has latterly taken up.

The alternative conception of density is to count sellers in relation to an area. However, once we switch to such a notion of density, we are back with our old problem of defining the area which is relevant, in short, of defining the "group". So far, every criterion proposed for de-limiting such an area has depended, either directly or indirectly, on the idea of a gap in the chain of substitutes – on there being a concentration of firms within an area with a density greater than that of surrounding areas. But as we approach "pure monopoly", or as products become less and less substitutable, it becomes harder and harder to de-limit any area of concentration of firms. The area will get larger and larger, and with it the numbers within it. This is the conception of density involved in Bishop's numbers formula.
Neither of these conceptions of density, we have seen, is adequate in a world of differentiated products. And the reason goes directly back to the dilemma which Mrs. Robinson pointed out to us at the outset. Insofar as each seller has a monopoly of its own output, it may be represented as a single seller at a point. Insofar as every product competes, as an alternative use of money, with every other product, the firm selling it may be considered as a member of a "group". In short, every firm may be considered, at one and the same time, as both an individual seller and as a member of a collectivity. Indeed, both aspects of a firm's existence must be taken into account. To take one or the other of these aspects to the exclusion of the other, is to take only a partial view of a firm's relationship with its environment. Necessarily, the result will be to plug a gap on one side only to open another on the other side. What we require, then, is a concept of density which embodies both aspects of a firm's being.

3. An Alternative Conception of Density

The conception of density (or numbers) to be
proposed as an alternative to those we have examined is closely related to the traditional idea of numbers, in that it will define numbers in relation to a point. This avoids any arbitrary demarcation of areas as such. How then do we treat the problem raised by the "group" of differentiated competition, where firms are concentrated into a small area, yet where no two firms are concentrated at the one point? We shall take account of this by describing the situation as if such firms were concentrated at a single point. In what follows, we shall propose a reduction formula which will give us a density "number" representing an index of the effective "group force" operating upon any given firm. In the case of homogeneous competition, where firms are in fact concentrated at a particular point, this density number will be equal to the arithmetic number of firms so concentrated. In the case of "pure monopoly", it will be equal to unity. Such an index will allow us

---

8. There are numerous analogous conceptual procedures in scientific theory. For example, in pure mechanics, the mass of a body is treated as if it were all concentrated at the centre of gravity.
to describe the transition from pure competition to pure monopoly as a continuous function.

As an approach to establishing such a formula, let us recall the reason why, as we have argued, cross-elasticity of demand as defined in Chapter II is inadequate to describe the "group". This was so, we said, because cross-elasticity of demand represents a relationship between two firms considered in isolation. It was, therefore, unable to tell us anything about group characteristics. Let us consider this a little further.

When any firm i cuts price, then a stimulus is applied to buyers of the product j to shift expenditure away from j towards product i. Income effects apart, there is for each such customer one — and only one — avenue of substitution opened up as a result of the price-cut. However, consider the position if firm i raises price. There is now applied a stimulus to the customers of i to move some or all their expenditure in the direction of all other products or any one of them, since the price of i now compares less favourably with every other price. Whereas, in the former case, the
customers, at least insofar as their expenditure on product \( j \) is concerned, have only to consider their relative preferences as between the two products \( i \) and \( j \), in the latter case, \( i \) customers must consider the whole system of preferences as between \( i \) and all other products. If we were to ring around the two firms \( i \) and \( j \), separating them from all other products, this would in no way affect the value of the cross-elasticity \( E_{ji} \) defined by a price-cut. But it is otherwise in the case of the cross-elasticity (which we will denote \( e_{ji} \)) defined in respect of a price-rise, for by isolating the two firms we would be cutting off avenues of substitution which would otherwise be opened.

Let us return to consider the case of homogeneous competition. If price is cut by any firm \( i \), then, conceptually \( i \) will gain the total output of the related firm \( j \), however small the price-cut. If the price-cut is made small enough, the coefficient \( E_{ji} \) approaches \( \infty \). This happens, as Triffin has shown us, whether the situation is one of duopoly, oligopoly, or pure competition, so long as \( i \) and \( j \) are perfect substitutes. Now consider the value of \( e_{ji} \) when the price of \( i \) is raised, all other prices being unchanged. Now firm \( i \) will lose its total output, but this will only accrue wholly to \( j \) in the case of pure duopoly. In the case
of oligopoly, its loss will be distributed at random over the few rival firms, in the case of pure competition it will be distributed over a large number of competitive firms. The cross-elasticity defined by a rise in price has, in short, a numbers element, and it is no doubt the reason, or one reason, why the confusion of cross-elasticity with numbers which we discussed in Chapter II has proved so intractable. As mentioned previously it has been the almost invariable practice to identify the two.

An exception, however, is to be found in the well-known article of Hall and Hitch. In that paper, the authors explicitly distinguish the cross-elasticity defined by a rise in price, our \( e_{ji} \), as the one relevant to the determination of numbers. Moreover, "there are two factors which ... tend to make these cross-elasticities small. One is the smallness of the proportion of consumers (or potential consumers) for whom the elasticity of substitution is high ... The second factor is the range and evenness of 'scatter' of the affected consumers among the products of other firms."  

It is important to notice that this cross-elasticity, $e_{ji}$, is relative to the substitution cross-elasticity, $E_{ji}$; for it is, in part, determined by the substitutability of the products $i$ and $j$. Thus, as Hall and Hitch tell us, the value of $e_{ji}$ varies directly as the substitutability between products and inversely with the number of similarly related firms. Therefore, if we divide $e_{ji}$ by the substitution cross-elasticity $E_{ji}$, we may eliminate the substitutability component, leaving only the numbers element. However, since the relationship between $e_{ji}$ and numbers is an inverse one, we must take the reciprocal ratio as our criterion of numbers. Then we have $N = E_{ji}/e_{ji}$ as an index of numbers or, if you like, density.

Now the significant thing about this coefficient is that it will reflect changing numbers in both the senses discussed in Section 2 above. It will take account of movement along either scale separately and hence, if both are operative, jointly. Beginning as before with our 1,000 firms selling homogeneous products, we find that the value of $N$ will approach
if outputs are symmetrical. Thus, if \( i \) cuts price, it will gain the whole of \( j \)'s output. (It will be remembered that we have rejected any supply conditions upon cross-elasticity.) On the other hand, if \( i \) raises price it will lose its total output and this will be distributed at random between 999 other firms. Firm \( j \) can only expect to gain \( \frac{1}{999} \)th of \( i \)'s output, as against the 100 per cent of \( j \)'s output \( i \) would get as a result of cutting its price. It is worth noticing also that our coefficient correctly reflects asymmetry between firms' outputs. Thus, in our example, if we were to double the output of \( i \), leaving the other 999 firms as before, the coefficient would be halved. A price-cut by \( i \) would still involve \( j \) in 100 per cent loss of sales, leaving the numerator unchanged. However, \( j \) could now expect to get twice the extra sales from a rise in \( i \)'s price, as \( i \) has now twice the sales to lose. The denominator is therefore doubled and the numbers coefficient halved, reflecting the increased importance of \( i \) in the market.

Let us now "knock out" some of our 1,000 firms in the manner described earlier in connection with the

\[11. \] It need hardly be pointed out that cross-elasticity equation will always reflect numbers exclusive of firm \( i \).
second conception of numbers which we described, \( N \) will fall \textit{pari passu}. When we are left with only two firms, pure duopoly, our coefficient will be equal to unity. Thus, as we operate on the scale of numbers with substitutability held constant, \( N \) moves in correspondence with the arithmetic number of firms.

Again, if we operate on the scale of substitutability with the arithmetic number of firms held constant, in accordance with the first conception of numbers, the coefficient \( N \) will approach unity as we let substitutability get very small, that is as we approach "pure monopoly" or "isolated selling". This is so because, for those customers substituting marginally between the products of two isolated sellers \( i \) and \( j \), there is no difference, so far as the distribution of their expenditure between \( i \) and \( j \) is concerned, between a cut in the price \( p_i \) (which gives us \( E_{ji} \)) and a rise in \( p_i \) (which gives us \( e_{ji} \)). The concept of marginal substitution is here used to distinguish the situation from one involving integral substitution which would usually be expected to occur between good substitutes in use, whether perfect or not. In the one case the consumer will be buying the two products simultaneously - which in some sense to be discussed later makes them "different" commodities - whereas in the other
consumption will normally be mutually exclusive, the commodity actually being used depending upon relative prices. The distinction is, I believe, logically important and will be developed at greater length in Section 5 below.

The fact that \( N \rightarrow 1 \) in both the case of duopoly and "pure monopoly" highlights one of the basic properties of our numbers coefficient. Triffin has rightly stressed the need to cast the theory of the firm in terms of "external interdependence". But if we are going to describe the interdependence of firms, we need to begin, not with the one-firm "atom" of Professor Chamberlin, but rather - if we may persist with his somewhat shaky metaphor - with a "molecule" of two firms. And it is precisely upon this that our conception of numbers or density is based. This two-firm starting point is additionally valuable, since it also provides us with the basic link of chain relationship. There is one further point worth noticing. Adoption of such an approach makes the theory of the "external interdependence" of the firm symmetrical with recent developments in the

theory of demand where, instead of beginning with a consumer's utility function in respect of one commodity, we begin with his substitution (indifference) function between two commodities.

We have, then, two basic "molecules", each of which constitutes a two-firm relationship. In the duopoly case, the firms are selling the same or a "similar" commodity. In the case of "pure monopoly", they are selling "different" commodities. It remains, then, to distinguish between them, as it does to distinguish the oligopolistic group from the non-oligopolistic. Both these ends are achieved, of course, once we determine a criterion of "oligopolistic interdependence" by which we may isolate the existence of oligopoly. This, however, is our task in the next chapter, and it would take us much too far from the present theme to develop the matter at this juncture. For the time being, then, we anticipate the appropriate criterion and assume that we have isolated the oligopoly case.

When oligopolistic interdependence is eliminated, the coefficient \( N \) enables us to trace the continuous transition from pure competition to "pure monopoly", It becomes a measure of the group force exerted by the collectivity of related firms upon any given firm. This
group pressure upon the firm comes about as the result of the asymmetrical relation of the firm which only supplies a fraction of total output. It arises from the submersion of the actions of the one firm by the actions of the many. The degree of such pressure, measured by N, varies continuously as the asymmetry of the traditional group disappears, so that with duopoly, chain relationship, or "pure monopoly", when N is equal to unity, the firm in question is no longer considered to be a member of a group. Only in the case of "pure monopoly", however, is the firm's price-output policy, in some sense, "independent".

Since the terms "pure monopoly", "single seller" and "isolated seller" imply, or have acquired the implication of, a degree of isolation from the general economic matrix which, in fact, no firm can enjoy in practice, and to emphasize the special sense of our definition, it is proposed to call a firm in this category an independent seller, a description the appropriateness of which it is hoped will become apparent later. An independent seller is defined, then, by an N which approaches unity, on the one hand, and the absence of any oligopolistic effect, to be
determined later, on the other. Any departure of $N$ from unity brings into play some "group effect" which reaches a maximum in the large-number homogeneous group which we call pure competition, when $N$ will be equal to the arithmetic number of firms in competition with the firm under consideration.

4. The Reflex Kinked Demand Curve*

In this section, it is hoped to show that the description of the transition from pure competition to independent selling which we have elaborated above constitutes something more than mere academic refinement. For, where $N$ is significantly large, a numbers factor appears which, it will be argued, must fundamentally affect the pricing behaviour of the firms concerned.

Let us momentarily assume, merely for purposes of algebraic simplification, that all firms which are related have outputs of the same size. Then it is easy to see from our definitions of the two cross-elasticities, $E_{ji}$ and $e_{ji}$, that:

(a) the ordinary elasticity of demand of firm $i$ with

* i.e. reflex to the origin. This description follows the usage of Emfroyson, "A Note on Kinked Demand Curves," *American Economic Review*, Vol. XXXIII, March, 1943.
in respect of a cut in price, denoted $E_{ii}$, is equal to $\sum x_{ji}$, where the sum is taken over all firms $j$ which are related to $i$.\(^{13}\)

(b) the ordinary elasticity of demand of firm $i$ in respect of a rise in price, denoted $e_{ii}$, is equal to $\sum e_{ji}$.

Now the condition that the demand (average revenue) curve facing firm $i$ be smooth may be expressed in terms of the derivative function. Alternatively, and for our purposes more conveniently, it may be expressed in terms of the derivative of the total revenue function, i.e. in terms of the marginal revenue curve. Thus, the demand curve will be smooth, if the ratio $\frac{E_{ii} - 1}{E_{ii}}$, as we approach any point $P$ from below, approaches the value of the ratio $\frac{e_{ii} - 1}{e_{ii}}$, as we approach $P$ from above. This condition will be satisfied, then, if either (i) both $E_{ii}$ and $e_{ii}$ are very large, i.e. as we approach homogeneous

13. Cross-elasticities can, of course, only be summed in this simple fashion in virtue of our temporary assumption that the $q_i = q_j$. With asymmetrical outputs, a weighted sum after the manner of equation (1) of Chapter II (p. ) is required.
competition; or (ii) $E_{ij} \rightarrow e_{ij}$.

From the previous discussion it will readily be seen that this latter condition will hold only if, for a given small $\phi_i$, $E_{ji} \rightarrow e_{ji}$, or, in other words, if $N \rightarrow 1$. In short, a smooth demand curve will occur only in the case of independent selling. By contrast, in the general case of heterogeneous competition a group relationship, a kinked demand curve as in Figure 3 emerges.\(^{14}\)

\section*{FIGURE 3}

![Diagram showing a kinked demand curve]

\(^{14}\) Perhaps it is the nature of this curve which explains one of the large paradoxes of the monopolistic competition debate. On the one hand, Kaldor has told us that monopolistic and imperfect competition amount to the same thing, since there is no monopoly in either. On the other hand, it would seem, if I understand him correctly, that Andrews' complaint is that there is no competition in either.
Clearly, this result derives from the asymmetrical relationship of the individual firm vis-a-vis the collectivity which we call the "group". It is, indeed, a generalization of the proposition that: if a firm in a large group cuts price, it will attract marginal or sensitive customers from a large number of rival firms; if, on the other hand, the firm raises price, it will lose its own marginal customers only. That is to say, the downwards elasticity facing (say) firm i is a function of the reactions and preferences of the non-i customers, whereas the upwards elasticity is a function of the reactions of the i customers. Since the i customers are few compared with the non-i customers, there is no reason at all why we should expect these two distinct functions to define a smooth demand curve. Quite the contrary!

We may now abandon the assumption of symmetrical outputs and express our numbers or density coefficient in terms of own-elasticity of demand. If we let $z$ be the numbers coefficient so defined, then $z = \frac{E_{ii}}{e_{ii}}$. And, hence, we may now re-define an independent seller as a seller faced with
a smooth demand curve such that $\gamma \rightarrow 1$ and not subject to oligopolistic pressure, which we rule out according to the criteria to be developed in Chapter IV. It may be noticed that this definition excludes the "smooth" (horizontal) curve of the firm under pure competition.\(^{15}\) For, under pure competition, although both $E_{ii}$ and $e_{ii}$ may be made to approach $\infty$, if we make $\frac{\partial p_i}{\partial q_j}$ small enough, $E_{ii}$ will always be large compared with $e_{ii}$; and the ratio $E_{ii}/e_{ii}$ still correctly describes the number of sellers involved.

Finally, it should be pointed out that the conclusion of this section is not novel; for it is implied in Marshall's notion of "marginal mobility". Marshall, who was not unaware of the imperfections to be found in real markets, believed that, so long as there were a few sharp housewives, a few sensitive buyers attached to each of a large number of firms, a competitive result, at least upon the assumption of

15. One of the least satisfactory features of Chamberlin's more recent position, which we discussed in Chapter I, is his treatment of the pure competition case. Thus, *the individual seller in this latter instance (large numbers) is again correctly described as 'isolated', even in the special case of pure competition.* Monopolistic Competition Revisited, loc. cit., p. 362. My italics.
free entry and no collusion, would follow from the competitive bidding for the custom of these sensitive buyers. This would happen, Marshall believed, even if the majority of customers were very immobile, i.e. if the demand-elasticities upwards were small.\(^\text{16}\)

Also, after having been driven by certain persistent logical contradictions in the theory of monopolistic competition to the conclusions of this section,\(^\text{17}\) my attention has been drawn to more explicit statements of the consequences of marginal mobility by A. J. Nichol and Alfred Nicols.\(^\text{18}\) These writers have placed particular stress on the existence throughout most imperfect markets of unattached buyers for whose custom sellers may be expected to compete. It is easy to verify that the presence in any significant number of such unattached buyers will result in a discontinuity in the demand curve facing the individual firm. If these buyers are literally "unattached" — in

\(^{16}\) I am originally indebted to Maurice Dobb for this point. See *Political Economy and Capitalism* (1937), p. 196n.

\(^{17}\) First stated in my paper on "A Kinked Demand Curve for Monopolistic Competition", Economic Record, May, 1953, 19-34.

the sense that an infinitesimal price-cut will be sufficient to move them - the demand curve would include a horizontal segment immediately below the prevailing price. On this basis, Nichol and later Nicols have attempted to lead a movement towards "the rehabilitation of pure competition". The present writer, however, prefers to state the situation in the more general form of a kinked demand curve, because the general case does not depend on buyers being literally unattached. For even where buyers are willing to pay a finite premium for the enjoyment of one commodity rather than another, these buyers will become "marginal" or "indifferent" with the establishment of certain price ratios, i.e. when the price-cut is just sufficient to wipe out the premium involved. Since in the real world price changes must always be finite and since it is clearly more appropriate to consider finite changes when concerned with monopolistic competition, it seems preferable to describe any discontinuity in the demand

curve in terms of the effect of a series of small
finite price changes in both directions. Further
insight into the nature of the situation, so described,
may be obtained by returning to the fundamental question
raised by monopolistic competition theory: What is a
commodity?

5. What Is a Commodity?

Once it is agreed that, in the case of homogeneous
competition, the ratio $E_{ii}/e_{ii}$ is equal to the arithmetic
number of sellers in competition with firm $i$, and that,
in the case of independent selling, $E_{ii}/e_{ii}$ is equal
to unity, it becomes clear that, in the continuum of
reality, the ratio may take any of a continuous series
of values between these limits. The idea that, with
the slightest differentiation of products, the ratio
jumps suddenly to unity— a fundamental assumption of
monopolistic competition theory—is repugnant to
commonsense. Rather, it may be expected, the ratio
will be the greater, the nearer the approach to pure
competition; it will be the less, the closer the
approach to independent selling.
We must now explore the nature of the "middle ground" between our two limits in more detail. Preliminary to this end, let us first look more closely at our limiting cases.

In some sense, homogeneous competition occurs between firms selling the same commodity; independent selling occurs when a firm is selling a commodity different from any other commodity. The final judgment as to whether or not commodities are the same or different rests, of course, with the consumer. Let us, therefore, look at our two limiting cases in terms of the theory of consumer preferences as developed by the neo-Lausanne school led by Hicks and R.D.G. Allen. This analysis distinguishes three types of relationship existing between any two products X and Y. The indifference curves representing these three cases are set out in Figure 4.

**FIGURE 4**
Curve oo' represents the case where substitutability between the commodities is zero. Consumption is independent of the relative prices of the two commodities concerned. This type of relationship will, of course, exist between a good many pairs of commodities in the real world. However, it cannot obtain between any one commodity, say X, and all other commodities. Otherwise, the seller of X would be a "pure monopolist" à la Triffin whose demand-curve would be of unit elasticity, with profit-maximization achieved with an infinitesimal output. From the point of view of the firm, then, this relationship is irrelevant. Firm Y may just as well belong to the economy of Mars as far as firm X is concerned. We shall therefore confine our attention to the situations described by aa' and bb'.

Curve aa' represents the case where commodities are identical (in the eyes of the consumer). It is a straight line - elasticity of substitution equal to $\infty$ - with a slope of $45^\circ$. Substitution in this case may be described as integral and instantaneous.
In other words, the slightest deviation of the price-line from a slope of $45^\circ$ will result in the concentration of total consumer expenditure at one end of the curve or the other, according as which price is the lower.

Curve bb' represents the case where substitution between commodities is marginal or continuous. A change in the ratio of prices in this case results in an increment in the consumption of the commodity whose price has fallen (relatively) and a decrement in the consumption of the commodity whose price has risen. The important thing to notice is that both commodities are consumed simultaneously. This implies that the two commodities have different use-values, so that, in some sense, they are different commodities.

Of course, the indifference curve analysis was born in the period of the general assumption of perfect competition. Hicks and Allen and their followers have continued in this tradition and have refused to adopt the Chamberlin-Robinson introduction of general monopolistic assumptions. And the crux of the perfect competition model is that commodities are either the same or different. Curve aa' represents the case where they are the same; curve bb' the case where they are
different. These two categories were exhaustive in traditional theory, general equilibrium and particular equilibrium theory alike.

It may be objected that such "different" commodities are only relatively different, since the rate of substitution between pairs of goods may be expected to differ from one pair to another. But in this respect the distinction is no different from that made, without question, in many other branches of theory. For example, the rate of substitution between "different" factors of production will not generally be the same. It is hardly likely that capital will substitute with labour at the same rate as with land, and so on. The logical basis for describing them as different factors lies in the fact that, in varying proportions, they are used simultaneously or, if you like, in combination.

We may now verify the proposition, made in Section 3 above, that where any firm i is related to other firms only in the manner of bb', i.e. where substitution of i's product is continuous (or zero) in every direction, the ratio \( N = \frac{E_{ji}}{e_{ji}} \) will be equal to unity. In Chapter II, we saw that, ignoring

income effects, cross-elasticity of demand in respect of the individual consumer, \( E_{ji} \), is given by \( E_{ji} = k_i \sigma \), where \( k_i \) is the fraction of expenditure spent on \( i \) and \( \sigma \) is the elasticity of substitution between \( i \) and \( j \). Then cross-elasticity for the firm is given by the weighted sum of these individual cross-elasticities. Thus:

\[
E_{ji} = \sum \frac{q_{jn}}{Q_j} k_i \sigma_n
\]

where \( q_{jn} \) is the outlay on \( j \) of the \( n \)th customer and \( Q_j \) is total consumer outlay on product \( j \), i.e. \( Q_j = \sum q_{jn} \). Since with an indifference curve of the type bb' the elasticity of substitution will be approximately the same for movement in either direction from any point \( P \), it follows that cross-elasticity of demand, given by (1), will also be approximately the same in respect of movement in either direction. Hence, in the limit, \( E_{ji} \) will approach \( e_{ji} \).

Let us now turn to consider differentiated or, in some sense, similar products. The products X and Y are differentiated when consumers are prepared to pay
something extra for the product X in preference to the product Y, or vice versa. In short, in the pure case, when all of a series of similar products are differentiated in the minds of consumers, preferences are ordered or ranked, so that a certain preference ratio exists between any two of the products. Thus, a consumer may have the following scale of preferences:

<table>
<thead>
<tr>
<th>Product</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>96</td>
</tr>
<tr>
<td>C</td>
<td>90</td>
</tr>
<tr>
<td>D</td>
<td>88</td>
</tr>
<tr>
<td>E</td>
<td>86</td>
</tr>
<tr>
<td>F</td>
<td>81</td>
</tr>
<tr>
<td>G</td>
<td>80</td>
</tr>
</tbody>
</table>

The preference ratio between A and C is then 10:9.

In other words, if the price of A is 10/- and the price of C is 10/-, this consumer will buy A. If the price of C is cut to 9/6, the consumer will still buy A. If the price falls to 9/-, the consumer will be on the margin of indifference between A and C. Once the price of C falls below 9/-, the price of A remaining at 10/-, this consumer would switch his total expenditure, abstracting from income effects, to product C. In terms of indifference curves, this case
would be represented as in Figure 5. The curve would be a straight line along the line of indifference, i.e. with a slope of $10/9$.

Here, we may describe the situation as one involving finitely-lagged integral substitution, as distinct from the instantaneous integral substitution of homogeneous competition. It is in these terms, then, that we describe the "middle ground" of heterogeneous or differentiated competition in which the products are neither the same nor different.

We have now proposed an answer to the basic question: What is a commodity? Commodities are
the same when they are related in the manner of aa' 
(homogeneous competition); they are similar when
related as in dd' (heterogeneous competition); they
are different when related as in bb' (independent
selling). Alternatively, we may describe the
categories as follows:—

(1) **Homogeneous competition** — when substitution is
integral and instantaneous.

(2) **Heterogeneous competition** — when substitution is
integral but finitely-lagged.

(3) **Independent selling** — when substitution is
continuous (or zero) in every direction.

These are, of course, "pure" categories; in the
real world they will blend into each other. Thus,
between (1) and (2) we will get hybrid cases where for
some consumers commodities are perfect substitutes (as
in aa'), while for others they will be imperfect
substitutes (as in dd'). Homogeneous competition in
which all related products are perfect substitutes for
all consumers is clearly a limiting case. Likewise,
we may expect to get hybrid types between heterogeneous
competition and independent selling. It is not hard
to imagine situations in which, under the impetus of changing price ratios, some consumers will substitute integrally between products such as tea and coffee, butter and margarine, Scotch and Australian whisky, whereas other consumers would react by buying a little less of the one and a little more of the other, i.e. substitute marginally. Thus, pure independent selling also appears as a limiting case. However, unlike the limiting case of "pure monopoly" as postulated by Chamberlin or Triffin, our lower limiting case - independent selling - is conceivable as an actual market situation. Not only is it conceivable, it is likely.

We have shown that for symmetrical homogeneous competition our numbers coefficient, \( \gamma \), will be equal to the arithmetic number of firms in competition with the firm under consideration. We have also shown that for independent selling \( \gamma \) will be equal to unity. It remains to show, in the case of heterogeneous or differentiated competition where a group of firms is involved, that \( 1 < \gamma < n \), where \( n \) is the arithmetic number of firms related to the firm being considered. nearer It will be unity, the nearer the approach to independent selling; it will be closer to \( n \), the closer the approach to homogeneous competition.
Consider the pure case of heterogeneous competition, as defined by dd', associated with large numbers. From the discussion above, it will be seen that two fundamental characteristics distinguish it from the case of pure independent selling:

(1) Consumption of "similar" commodities is mutually exclusive. This involves the division of consumers into two distinct categories: those buying from any firm i, on the one hand, and those buying from the large number of producers of similar products on the other. By definition, the number of i customers will be small compared with the non-i. In other words, firm i may potentially attract many more customers than it can lose. Usually, the "law of large numbers" is invoked to justify the ordinary economic assumptions of continuity; but no matter how many individual consumers may be assumed to be purchasing the particular range of "similar" products in question, this cannot affect the basic asymmetry resident in the "large group" case.

(2) Under the stimulus of finite price changes, consumers become "indifferent" between "similar"
commodities. As the price-line passes through the line of indifference, the consumer concerned is involved in integral substitution in the one direction or the other.

These two factors, taken together, make it inevitable that the demand-curve for a firm related in this way to a group of other firms be discontinuous. The situation is asymmetrical and discrete; and we cannot apply to it the logic of the symmetrically continuous. For the demand-curve to be smooth under such circumstances, customers would have to leave i, in the face of a rise in i's price, at n times the rate at which customers would leave j, if i cut price, where n is the ratio of non-i to i customers or, if the sales of each firm are the same, where n is the number of firms. Such a situation cannot be envisaged in terms of any realistic assumptions we can make about the nature of consumers' preferences. 22

Alternatively, the condition for a continuous demand function for the individual firm i in this context may be stated as follows: the curve will be smooth only if any firm i will attract the same number

22. See Appendix "A" to this chapter.
of customers from another related firm z (i.e. related in the manner of dd'), whether firm i cuts price alone, on the one hand, or multilaterally with firms j, k, ... y on the other. The unreality of such a condition is immediately obvious.

Consider the situation in which firm i, which is related to firms j, k, ... z, unilaterally cuts price by, say, 10 per cent, so that all other prices are unchanged. Let us consider the effect of such a price-cut on the movement of customers from z to i. Those customers who will be induced to move may be divided into three exclusive categories:

(1) Those for whom product i is a unique second preference and for whom the operative price-cut is sufficient to invoke their second preference but no others. In other words, they would not be induced to move by an equivalent price-cut by any of the other firms j, k, ... y.

(2) Those for whom product i is an equal second preference with one or more other product j, k ... y and who would move indifferently to such equal second preferences in the face of a 10 per cent price-cut.
(3) Those for whom the given price-cut is sufficient to invoke lower preferences than the second, say the $r^{th}$, and for whom product $i$ is ranked as the $r^{th}$ preference or higher.

Now it is clear that only in case (1) will the movement of customers from $z$ to $i$ be independent of whether other firms $j$, $k$, ... $y$ cut price simultaneously with $i$, and it is equally clear that this case can hardly be the general one for all customers. In cases (2) and (3), firm $i$ must share some of the customers it could attract from $z$ by unilateral price-cutting with firms $j$, $k$, ... $y$, if those firms also cut price by an equivalent amount. Let us call the customers in category (1) "$i$-specific" customers. (They are specific in the sense that, for the given price-cut, they will move to $i$ and $i$ alone.) Let us also suppose that the proportion of such customers is $k$, so that the percentage in categories (2) and (3) is $1-k$.

Now consider the situation when firms $i$ and $j$ cut price together. The movement of customers from $z$ to $i$ cannot now be described by the cross-elasticity $E_{zi}$, since this has been defined on the assumption that
all prices other than i's are held constant. A new cross-elasticity, \( E'_{zi} \), must be defined to describe the situation now being considered. \( E'_{zi} \) will differ from \( E_{zi} \) on account of the behaviour of customers in categories (2) and (3). All these customers will not go to i, if j also cuts price. If we assume that consumers' preferences are randomized, i.e. if we assume that any order of consumers' preferences is equally likely, then —

(a) For those customers in category (2) who are indifferent, as second preferences, between i and j, it will be equally likely that they will move to j as to i, so that each firm may expect to get one-half of such customers.

(b) In respect of those customers in category (3), it will be equally likely that customers rank j higher than i as it is that they rank i higher than j, so that again both firms may expect to get one-half of such customers.

Since there will be no effect upon the movement of i-specific customers, it is easy to see that when i and j cut price together that —

\[
E'_{zi} = kE_{zi} + \frac{1-k}{2}E_{zi}
\]
Similarly, if firms i, j, and k cut price together —

\[ E_{zi}'' = kE_{zi} + \frac{1-k}{3} E_{zi} \]

while if all the n firms related to z cut price equivalently and simultaneously —

\[ E_{zi}^n = kE_{zi} + \frac{1-k}{n} E_{zi} \]

\[ = \frac{(n-1)k + 1}{n} E_{zi} \]

Now the cross-elasticity \( E_{zi}^n \) represents the movement of customers from z to i, when all related firms cut price together. It is equivalent, therefore, to the cross-elasticity defined by a unilateral rise in the price of z, i.e. to \( e_{iz} \). Then, on the simplifying symmetry assumption that \( E_{zi} = E_{iz} \) or \( e_{zi} = e_{iz} \), we may now express the numbers coefficient developed earlier in this chapter in the form —

\[ N_{zi} = \frac{E_{zi}}{e_{zi}} = \frac{E_{zi}}{E_{zi}^n} = \frac{n}{(n-1)k + 1} \]

For homogeneous competition, k is equal to 0 and, therefore, \( N = n \); that is to say, our numbers coefficient takes the arithmetical value of the number of firms related to the firm under consideration, a result which we established earlier. On the other hand,
if $k$ approaches unity, $N_{zi}$ will also approach unity and the demand-curve would be smooth. As has been pointed out, the condition that $k \to 1$ requires that virtually all customers fall into category (1), a situation which is inconceivable where customers' preferences are of the form of $dd'$. This fact will be further demonstrated in Appendix "A" which follows. For the rest, so long as there are, for any given price-cut, customers falling in the categories (2) and (3), the downward elasticity in respect of such a price-cut will be significantly different from the upward elasticity defined by a price-cut of equivalent order.

It may be objected that no-one ever believed that the two elasticities would be equal. However, I think it will be accepted that the general presumption has been that, within relatively small ranges of price-change, the one would approximate the other. Certainly, the idea that downward elasticity may be 2, 3, or even $n$ times greater than upward elasticity has not been generally envisaged, and this has led the theory of monopolistic competition into a number of logical difficulties which we have noticed in this and previous chapters. The significance and importance facing this implication of monopolistic competition assumptions will be further taken up in
Chapter V. At the moment, we must pause to consider two fundamental matters involved in the argument of this chapter now that the broad outlines of that argument have been stated.

6. Our Concept of Density Reconsidered

Two matters require further consideration. First, the discontinuous demand curve postulated above must appear as somewhat disembodied, as floating in mid-air as it were, until we attempt to fix an origin from which our elasticities and cross-elasticities are measured in the one direction or the other. Second, some consideration must be given to the size of the price-changes which we use to define our elasticities and, hence, our numbers coefficient.

In the cases of literally homogeneous competition and of independent selling there is no difficulty in this connection. Since, in the former case, an infinitesimal price-change is sufficient to produce the total effect and, in the latter case, substitution is continuous, we may think, as has been traditional, in terms of infinitesimal price-changes. Also, we may take, as is also usual, the given market prices as the origin from which our elasticities and cross-elasticities are calculated.
However, this will clearly not do in the case of heterogeneous competition, where products substitute as in \(dd'\). For example, an infinitesimal price-change may, in these circumstances, produce a cross-elasticity of zero, although a finite change in price results in a cross-elasticity which is significantly greater than zero. Perhaps the best way to approach the issues involved is by analogy.

Imagine any one firm of a large group to experiment with its price, beginning at a figure very much higher than that at which rival products are selling. Small reductions in price at this level are likely to provide negligible increases in sales. Demand at this level is inelastic. However, as further reductions in price are made so that the firm's price begins to approach the general level of prices for similar commodities, the firm will approach a region of greater competitive "density". As it moves its price through this region of competitive density, the elasticity of demand will increase sharply. The demand curve will be refracted, as it were. As price falls well below the general level of prices of related products, it again enters a region of lower competitive density, where only customers of the highest immobility may be
affected by further reductions in price. In this region, the demand curve once again becomes inelastic. Thus, the kinked demand curve postulated in this chapter may be looked upon as a refraction of the demand curve as price moves through an area of greater competitive density. Clearly, the location of this region will depend upon the pattern of prices of similar products. The manner of this refraction of the demand curve may be shown graphically as in Figure 5A.

After all, it is the existence of just this region of higher competitive density which distinguishes the "large group" of monopolistic competition from "pure monopoly". As previously quoted, Kaldor has
drawn the basic distinction between the two market situations in just these terms: "in the case of polypoly there is (and in the case of monopoly there is not) a group of firms whose prices jointly exert a large influence on the firm's sales."\(^\text{23}\)

In practice, of course, this region in which prices of similar commodities exert an overwhelming influence on the elasticity of demand need not necessarily be sharply bounded. By and large, however, the greater the density of surrounding products, i.e. the greater the number of competitive products and the closer they are to the particular firm, the sharper will be the refractory effect.

In the limit, with pure competition, the refraction of the demand curve will take place within the bounds of a single line, the line of market price. Above that line, the firms loses the whole of its own sales; below it, the firm theoretically gains the whole world, i.e. total market output. As competition becomes less and less dense, the refractory effect will become less and less marked. In the uniformly thin competitive field of the independent seller, it will disappear.

Having thus pictured the "large group" demand curve in terms of the refraction analogy, we are now able to consider the two ambiguities of our analysis mentioned earlier, viz. an origin from which to measure the elasticities involved in our conception of numbers - if you like, the centre of gravity for our coefficient of density - and the size of the price-changes for which they are defined.

In the case of pure competition, the first point is easily taken. Clearly, the origin will be taken at the going market price which will be the same for all firms. A similar procedure would be in order where imperfect competitors are selling at a uniform price. But how should we deal with the situation in which imperfect substitutes are selling at different prices? The taking of some sort of weighted average of such prices suggests itself as a solution. (Incidentally, Kalecki has based his new value analysis upon this notion of a weighted average price for the "industry".) However, such a procedure does involve identification of those products whose prices are to enter into the average, since the average of all other prices would be

meaningless in this connection. The solution suggested here is that we should take as origin the weighted average price of all "similar" products, i.e. products which consumers predominantly substitute integrally, as in $d'd'$. In the case of independent selling in which substitution is continuous, there is no reason to depart from traditional practice of taking the firm's own price, wherever it may be, and calculate elasticities in terms of infinitesimal price movements.

Again, in the case of heterogeneous competition with its discrete system of consumer preferences, the notion of point elasticity of demand - of infinitesimal price movements - is also inappropriate. What we need is an arc elasticity related to the finite nature of the system. With integral substitution we will have discrete systems of ordered preferences - see Appendix "A" below - and, in some sense, the size of the price-changes required to bring out the nature of a given situation should be related to the size of the price-changes required to bring the pattern of consumers' preferences into play, i.e. should be related to the intervals between successive preferences. Since these intervals will vary from consumer to consumer, again
an average suggests itself. However, not only will the height of preference schedules vary from consumer to consumer, but the intervals between successive preferences of the same consumer will, in general, be different. In other words, the slope of $dd'$ will not be the same as between successive pairs of products. As the preferences of individuals will tend to be bunched with respect to some commodities and dispersed with respect to others according to the individual's situation and tastes, the modal $dd'$ would probably serve our purpose best. Then, for the individual consumer, we take as the required price-change that price-change necessary move the price-line from a slope of 45° (representing the origin of equal average price for each commodity) to the modal $dd'$. For the determination of arc elasticities in respect of the firm, we then take the weighted average of the price-changes so determined for individual consumers.

To sum up: to unambiguously define our numbers coefficient for heterogeneous competition, we calculate arc elasticities of demand for the individual firm in either direction from the weighted average price of all "similar" commodities and in
respect of finite price movements representing the weighted average of consumers' modal rates of substitution as given by the slopes of the indifference curves dd'.

(103)
APPENDIX "A"

to

Chapter III.

-----

DISCRETE CONSUMER PREFERENCES AND THE SHAPE
OF THE DEMAND CURVE

As we noticed in Chapter I, little attempt has been
made to integrate the theory of monopolistic competition
with the theory of consumers' behaviour. Moreover, we
have seen that there is very good reason for this. The
traditional theory of consumers' behaviour has been
developed in terms of maps of consumer preferences in
which substitution between products is assumed to be
continuous. Of course, it has always been recognized
that in individual cases there may be discontinuities
in the substitution process, but this has usually been
held to disappear in the aggregate. Thus Hicks tells us:

"This convenient assumption of continuity does, of
course, always falsify the situation a little (or
sometimes more than a little) as far as the individual
consumer is concerned. But if our study of the
individual consumer is only a step towards the study
of a group of consumers on the market, these
falsifications can be trusted to disappear when the
individual demands are aggregated." (1)

1. J.R. Hicks, Value and Capital, p. 11n.
By and large, the monopolistic competition theorists, while aware of the discrete nature of consumers' preferences between imperfect substitutes, have tended, implicitly, to rest their analysis upon a similar notion. While individual substitution may be integral and jumpy, this discontinuity could be expected to disappear in the aggregate, rather as a discrete statistical distribution may be approximated by a continuous distribution when the population is large.

In fact, the analogy does not hold because, where substitution is integral and hence the consumption of products is mutually-exclusive, we have, from the point of view of any firm i, two distinct populations — the i-customers and the non-i. If we have a large-number-of-firms group, then the non-i population will be very much greater than the i population. In these circumstances, no increase in the number of customers in the market or in the fineness of adjustment of their preferences can alter this basic asymmetry. By contrast, consumers buying different commodities (as in bb') will, of course, figure in both populations at the same time.
The widespread existence of consumer substitution according to a discrete scale of preferences raises many problems for the theory of consumer behaviour, if that theory is to be general. Here, however, we are only interested in the more limited question of the relationship between discrete systems of preferences and the demand curve for the individual product.

Perhaps the best way to bring out the implications of discrete scales of consumer preferences is to consider the conditions which are required if such a system of preferences is to generate a symmetrical demand function for the individual firm. Alfred Nicols has stated the conditions for a smooth falling demand curve, such as is assumed by the theory of monopolistic competition, as follows:

"As long as all the customers of each seller act by elaborate preference scales, with each scale roughly the same height, a seller might find this to be true. But once it is admitted (1) that not all the customers of each seller need be attached, and (2) that even where they are attached, the various premiums they are willing to offer for their favoured products, rather than go without them, differ, the demand curve for the individual seller becomes far more complicated than the existing theory of monopolistic competition holds."

2. A. Nicols, loc. cit., p. 119.
We may illustrate these conditions as follows. Suppose there are 5 products A, B, C, D and E within a group sufficiently well-defined that we may neglect substitution with extra-group products, at least for modest price-changes. There are 120 customers in the market at the going price which, for convenience, we may assume to be the same for all 5 products. Each customer purchases an equal amount — say one unit of product per period. Then, if customers' preference schedules are given at random, i.e. if we assume that each order of preferences is equally likely, there will be 5! ways in which preferences may be ordered. These are set out in Table 1 below. Since we have assumed there to be 120 customers and that each combination of preferences is equally likely, each possible combination will be represented by one customer. It is clear that there will be 24 customers attached to each of the firms A, B, C, D and E.

Further suppose that all customers are of equal sensitivity or, if you like, that all preferences are equally spaced such that every customer will move to a lower preference if, but only if, there is, say, a 5 cent difference in price. The going price, taken as origin, is assumed to be \$1.
We are now prepared to investigate the distribution of customers under the stimulus of price-changes. Suppose that firm A cuts its price by 5 cents or 5 percent. Then, this price-cut will attract all customers of B, C, D and E for whom A is a second preference, since we have assumed that a price differential of 5 cents is sufficient to invoke second preferences. From Table 1 it is clear that A will gain 6 customers (or units of sale) from each of the firms B, C, D and E - 24 in all. Elasticity for the price-cut is therefore $100\%/5\% = 20$. If A raises price by 5 cents, it will lose all its 24 customers, 6 to each firm. Hence, elasticity of demand for the price-rise will also be equal to 20. Thus, we find an apparent symmetry between upward and downward elasticity of demand for firm A.

The unreality of the case, however, is immediately apparent. For any price-rise less than 5 cents, A loses no sales at all. For a price-rise of 5 cents, it loses 100 per cent of its sales. Clearly, the unreality stems from the assumption that consumers' preferences are evenly spaced. We must now drop this assumption. We now consider a market in which 1,200 customers are operating. Again, preferences are given at random, there being no good reason why one sequence of preferences
should be more popular than another. However, we now assume that there is a difference in the degree of consumers' sensitivity, so that, taking a rectangular distribution, we assume that 120 customers will move to the next lowest preference under the stimulus of a 1 cent price differential, 120 under the stimulus of 2 cents, ... and 120 under the stimulus of 10 cents. We can then imagine 10 schedules similar to Table 1, with Schedule 1 having preferences spaced by 1 cent, Schedule 2 spaced by 2 cents, ... Schedule 10 spaced by 10 cents.

It is clear that, with all products originally priced at £1, there will be 240 customers attached to each of the 5 firms A, B, C, D and E. Then, as firm A reduces price the movement of customers will be as set out in Table 2. Firm A's elasticity of demand will be as follows:

<table>
<thead>
<tr>
<th>Price Differential</th>
<th>Elasticity Upwards</th>
<th>Elasticity Downwards</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 cents</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>4 *</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>6 *</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>8 *</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>10 *</td>
<td>10</td>
<td>21</td>
</tr>
</tbody>
</table>

Thus, for small finite price-changes, A's downward
elasticity of demand is approximately double its upward elasticity. Different shapes of this demand curve may be obtained by making different assumptions about the distribution of customers of varying degrees of sensitivity. In particular, the asymmetry increases as we increase the number of firms and their closeness together. On any reasonable assumption about the distribution of customers' preferences between the various alternatives, the demand curve derived does not look very much like that traditionally assumed.

Let us briefly consider the reason for the asymmetry which appears once all customers are not of the same degree of sensitivity, i.e., when all preferences of all customers are not equally-spaced. Each successive price-rise exhausts the customers of each corresponding schedule of sensitivity. But successive price-cuts do not do this. As price falls, a firm does not only attract the second preferences from each successive schedule, but also the third, fourth, or nth preferences of preceding schedules.

However, even this second example, in which we have allowed a spread of customers' sensitivity, is not very realistic. In the example given, all A's customers will have left in the face of a price-rise of 10 cents. All customers will have left B, C, D and E for A when A's price has fallen by 40 cents. Thus, it takes 4 times
longer to exhaust the potential customers, i.e., those who may ultimately be attracted from competitors, than it does to exhaust one's own customers. This may not at first sight appear unrealistic. However, the unreality soon becomes apparent once we increase the number of firms (products). Thus, if there are 20 firms it will take - irrespective of the particular distribution of consumer sensitivity we assume - just 19 times longer to exhaust potential customers by price-cutting than to lose the whole of one's own customers by raising of price. With consumer distributed as assumed in our example, a price rise of 10 cents will still exhaust all A's customers. On the other hand, with 20 products, a price-cut of 100 cents - reducing price to zero - will still leave customers attached to other products. In short, if A becomes a free good, there will still be a significant number of customers buying B, C, D and E at the price of 100 cents. Product A could hardly be said to be competitive with B, C, D and E so far as these customers are concerned. Indeed, their consumption of B, C, D or E would be independent of the price of A.

Thus, we find that, if the preferences of each customer are ranked and equally-spaced, however one may rule as to the distribution of customers between various degrees of sensitivity, it will take just n times longer
to exhaust potential customers than to exhaust one's own, where \( n \) is the number of competitive products. It should be noticed that this means that on the average customers will leave A at the same rate at which they are attracted to A, roughly the assumption of the smooth demand curve usually posited in monopolistic competition theory. Nevertheless, the rate at which customers move to A will be very uneven, there being a distinct hump in the region of the smaller price-cuts, a thinning out as larger and larger price-cuts are involved. As we have seen, this humping in the lower stages is due to the fact that at the same time as a price-cut is invoking the second preferences of less sensitive customers, it will also be invoking the third, fourth, or nth preferences of more sensitive customers.

Thus, when substitution between products is integral, we find two extreme asymmetries:

(1) when products are homogeneous, we find that any given price-cut, however small, will attract potential customers at \( n \) times the rate at which customers will be lost for the same price-rise; and

(2) when products are differentiated, such that all preference of all customers are ranked or ordered, with an equal absolute interval between preferences, it will require a price-cut \( n \) times greater to exhaust potential customers than the price-rise
required for any individual firm to lose the whole of its sales.

Clearly, most firms of the real world will lie between these extremes. Moreover, the closer the substitutes (i.e. the nearer the approach to homogeneous competition) and the greater the number of firms, the greater will be the asymmetry between upwards and downwards elasticities.

Finally, it is easy to verify that once some customers become indifferent\(^3\) between some commodities, whether as first, second, or subsequent preferences, the asymmetry becomes decidedly more marked. In the case of first preferences the reason is easy enough to see. The existence of some customers who are indifferent between some commodities means that there is a sub-set of the market which acts within an area of homogeneous products, with the resulting asymmetry which arises out of homogeneous competition. In the case of lower preferences, the reason for the asymmetry derives from the fact that, if (say) there are rth preferences which are distributed at random between all firms or some firms, only a fraction will be attached to any particular firm. A price-rise of the order required to invoke rth preferences will involve the individual firm only in the

\(^3\) Indifferent in the sense that customers are not prepared to pay more for the one product than the other.
loss of its own fraction of the customers concerned; whereas a price-cut of the same order would attract to the firm all other such customers who will be equally distributed among the other firms in the group.

Again, we have assumed in our example a rectangular distribution of customers of varying degrees of sensitivity, i.e. we have assumed that consumer sensitivity to a 1 cent price-change is as equally likely as to a 5 cent change or a 10 cent change. Once we assume anything in the nature of a normal distribution of consumer sensitivity, with a few very sensitive customers, a few very insensitive ones, and with most ranged around some modal degree of sensitivity, the humping of the demand curve for medium-sized finite price-cuts becomes more marked, even when, as in our example, only 5 products are involved.

Finally, this humping in the movement of consumers, involving as it does an asymmetry between upward and downward arc elasticities of demand, corresponds to what, in Chapter III, we have called "refraction" of the demand curve as price passes through a region of greater "competitive density". The significance of this asymmetry which arises when
the demand curve passes through a phase in which elasticity is rapidly increasing in this manner is, of course, that the marginal revenue function may no longer be taken as monotonic decreasing, a fact which must affect the strategies, as well as the condition for reaching an equilibrium, of a firm operating in such a context. This is a matter which will be taken up again in chapter V.
### TABLE I

The Possible Combinations of Ordered Preferences with Five Products; and the Distribution Between Firms When Prices Are Equal.

<table>
<thead>
<tr>
<th>Firm A</th>
<th>Firm B</th>
<th>Firm C</th>
<th>Firm D</th>
<th>Firm E</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDE</td>
<td>BACDE</td>
<td>CABDE</td>
<td>DABCE</td>
<td>EABCD</td>
</tr>
<tr>
<td>ABCED</td>
<td>BAGED</td>
<td>CBEAD</td>
<td>DABEC</td>
<td>EABDC</td>
</tr>
<tr>
<td>ABDEE</td>
<td>BACED</td>
<td>CABED</td>
<td>DABEC</td>
<td>EACBD</td>
</tr>
<tr>
<td>ABDEC</td>
<td>BADEC</td>
<td>CADEB</td>
<td>DACEB</td>
<td>EACBD</td>
</tr>
<tr>
<td>ABEDC</td>
<td>BAECD</td>
<td>CAEBD</td>
<td>DAEBC</td>
<td>EADBC</td>
</tr>
<tr>
<td>ABECD</td>
<td>BAEDC</td>
<td>CAEAB</td>
<td>DAECB</td>
<td>EADCB</td>
</tr>
<tr>
<td>ACBDE</td>
<td>BCADE</td>
<td>CBADE</td>
<td>DBACE</td>
<td>EBACD</td>
</tr>
<tr>
<td>ACBED</td>
<td>BCDAE</td>
<td>CBAED</td>
<td>DBAEC</td>
<td>EBADC</td>
</tr>
<tr>
<td>ACDEB</td>
<td>BCDEA</td>
<td>CBDEA</td>
<td>DBCEA</td>
<td>EBCDA</td>
</tr>
<tr>
<td>ACEDB</td>
<td>BCEDA</td>
<td>CBEDA</td>
<td>DBEAC</td>
<td>EBDAC</td>
</tr>
<tr>
<td>ACEBD</td>
<td>BCEAD</td>
<td>CBEDA</td>
<td>DBECA</td>
<td>EBDCA</td>
</tr>
<tr>
<td>ACEDC</td>
<td>BCEDA</td>
<td>CBEAD</td>
<td>DGEAB</td>
<td>EBDCA</td>
</tr>
<tr>
<td>AEDBC</td>
<td>BEADC</td>
<td>CEABD</td>
<td>DEABC</td>
<td>EDABC</td>
</tr>
<tr>
<td>AEBDC</td>
<td>BEACD</td>
<td>CEADB</td>
<td>DEACB</td>
<td>EDACB</td>
</tr>
<tr>
<td>AECBD</td>
<td>BECAD</td>
<td>CEBAD</td>
<td>DEBAC</td>
<td>EDBAD</td>
</tr>
<tr>
<td>AECDB</td>
<td>BECDA</td>
<td>CEBAE</td>
<td>DEBCA</td>
<td>EDBDA</td>
</tr>
<tr>
<td>AEBCD</td>
<td>BEADC</td>
<td>CEDAB</td>
<td>DEGAB</td>
<td>EDCAB</td>
</tr>
<tr>
<td>AEDCB</td>
<td>BEGCA</td>
<td>CDEBA</td>
<td>DEGAB</td>
<td>EDCBA</td>
</tr>
<tr>
<td>AEDCB</td>
<td>BEGCA</td>
<td>CEDEB</td>
<td>DEGAB</td>
<td>EDCBA</td>
</tr>
</tbody>
</table>
### Table 2

Demand Schedule for Firm A: 10 Schedules of Sensitivity

<table>
<thead>
<tr>
<th>Price (Cents)</th>
<th>Sales (Firm A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>108</td>
<td>48</td>
</tr>
<tr>
<td>106</td>
<td>96</td>
</tr>
<tr>
<td>104</td>
<td>144</td>
</tr>
<tr>
<td>102</td>
<td>192</td>
</tr>
<tr>
<td>100</td>
<td>240</td>
</tr>
<tr>
<td>98</td>
<td>312</td>
</tr>
<tr>
<td>96</td>
<td>432</td>
</tr>
<tr>
<td>94</td>
<td>528</td>
</tr>
<tr>
<td>92</td>
<td>624</td>
</tr>
<tr>
<td>90</td>
<td>720</td>
</tr>
<tr>
<td>88</td>
<td>792</td>
</tr>
<tr>
<td>86</td>
<td>816</td>
</tr>
<tr>
<td>84</td>
<td>888</td>
</tr>
<tr>
<td>82</td>
<td>936</td>
</tr>
<tr>
<td>80</td>
<td>984</td>
</tr>
<tr>
<td>78</td>
<td>1008</td>
</tr>
<tr>
<td>76</td>
<td>1056</td>
</tr>
<tr>
<td>74</td>
<td>1056</td>
</tr>
<tr>
<td>72</td>
<td>1104</td>
</tr>
<tr>
<td>70</td>
<td>1128</td>
</tr>
<tr>
<td>68</td>
<td>1152</td>
</tr>
<tr>
<td>66</td>
<td>1152</td>
</tr>
<tr>
<td>64</td>
<td>1176</td>
</tr>
<tr>
<td>62</td>
<td>1176</td>
</tr>
<tr>
<td>60</td>
<td>1200</td>
</tr>
</tbody>
</table>
IV. THE SMALL GROUP: WHAT IS OLIGOPOLY?

1. The Problem

In the last chapter we put aside until now the problem of providing a criterion of "oligopolistic interdependence" which is, of course, necessary to complete our description of markets. Unfortunately, this subject is also an area of profound confusion which is probably best demonstrated in terms of the attitudes engendered towards oligopoly as a result of the theoretical muddle.

Thus, in the course of a review of Bain's Pricing, Distribution, and Employment, Stigler wrote: "Bain believes that the economy is overwhelmingly oligopolistic ... My own tentative estimates ... suggest that ... competition is three or four times as important as all other forms of market organization combined."¹ Here, then, we have crystallized for us an impression one cannot escape in any survey of value discussions over the past decade or so. For one

economist oligopoly is the exception; for the next it is the rule. Interestingly enough, this fundamental split emerges even in the camp of the "empiricists". Thus, for Hall and Hitch oligopoly is virtually ubiquitous; "... the typical case is that of monopolistic competition with an admixture, which is usually large, of oligopoly."\(^2\) Clearly it is this aspect of the Hall-Hitch thesis which disturbs Andrews so much and, in large part, accounts for his pains to stress the competitive nature of industry.\(^3\) In fact, it is this, one suspects, which constitutes the real substance of the distinction between a price based on "full cost" and a price determined by average prime cost plus a "costing margin".

It hardly seems necessary to stress the important implications which flow from one's judgment about the market-structure, and in particular the pervasiveness of oligopoly, of the economy. It will influence one's views about the adaptability of the

---

The reaction of prices — as opposed to output — to an increase in Effective Demand, the role of anti-trust policies, and so on.

Therefore, if the widely divergent opinions held by economists about the importance of oligopoly in the economy represented different professional judgments upon the same set of facts, the claim of economics to scientific status would be practically nil. Fortunately, however, it is possible to put a more charitable interpretation upon the matter. Could it not be that these writers are, in fact, really talking about different things when they speak of "oligopoly"? This is the central question to which we address ourselves in this chapter.

Ever since the publication of Chamberlin's Theory of Monopolistic Competition there have, indeed, been two distinct conceptions of oligopoly running through value theory. The first (and older) sense is that involving the literal notion of "fewness". But there is a second conception of oligopoly which is general in modern value analysis. In this case, oligopoly is defined in terms of the transmissibility of price changes. Thus, as Triffin has it, "the crux
of the distinction between large and small numbers is really whether or not a price move by one seller induces the other sellers to follow suit ... \(^4\) It is in this sense that Chamberlin and Hall and Hitch speak of oligopoly.

In the old scheme – where commodities were different in kind, not merely in degree – the connection between "fewness" and a firm's influence on price was clear enough. If a firm's potential output was small compared with total market supply, then any price-output decision of that firm could only have a negligible effect upon market price. On the other hand, if its potential output were large, oligopolistic implications would arise. But in the modern scheme the concept of the "market" and of "market price" dissolves. Moreover, in the general case, a firm can only vary its output by first changing price. Hence, oligopoly came to defined in terms of the effect of a price change by one seller on the price policy of other sellers. By and large, the problem has remained at this definitional stage. Little has been done to work out the conditions of demand and supply which would give rise to this interdependence of prices beyond leaving

\(^4\) H. Triffin, op. cit., p. 30.
us with a vague idea that, on the demand side, the products should be more or less "good* substitutes, while, on the supply side, the firms should be - loosely - "few" in number. Our primary aim in this chapter is to investigate how good must be the substitutes and how few the numbers, if prices are to be oligopolistically interdependent.

2. From Chamberlin to Triffin

Let us return to Chamberlin's criterion for distinguishing the large from the small group which we quoted in Chapter II:

"Specifically, we assume for the present that any adjustment of price or of 'product' by a single producer spreads its influence over so many of his competitors that the impact felt by any one is negligible and does not lead to any readjustment of his own situation."(5)

Although this definition remains broad enough to suggest an affinity with the older criterion postulated in respect of homogeneous products, it leaves us rather uneasy. In particular, it overlooks the fact, which we noticed in Chapter II, that, in the generality of monopolistic competition in which a finite price-cut is required to attract a finite volume of sales from other firms, the size of the impact upon competitors will not be independent of the size of the price-change involved.

5. op. cit., p. 83.
This seems to have been in Triffin's mind when he re-states Chamberlin's criterion symbolically. In Chamberlin's terms, oligopolistic interdependence between firms i and j will be absent, he says, if 
\[ \frac{\partial R_i}{\partial p_j} \text{ is of insignificant size,} \]
where \( \partial R_i \) is the revenue lost by i in response to a price-cut of \( \partial p_j \) by j. The impact upon i is now specifically related to the size of the price change involved. This ratio, Triffin claims, is an adequate criterion of numbers, given Chamberlin's symmetry assumptions.

Unfortunately, however, this more precise formulation of Chamberlin's criterion merely results in a coefficient of substitution. The ratio will be large when a small price-cut by i is sufficient to take a large amount of revenue from i, i.e. when i and j are good substitutes. It will be small when it requires a substantial price-cut by j to attract a few sales from i, i.e. when i and j are poor substitutes. And this is precisely the position to which Triffin is led. First, he makes two minor modifications to the coefficient: (1) he expresses

the impact upon $i$ in terms of output instead of revenue, merely for the purpose of bringing symmetry into his own exposition; and (2) he expresses the price-changes and the quantity-changes as relatives, i.e. in elasticity form, to allow for asymmetrical relationships. The coefficient then becomes of the form $p_j \frac{dq_1}{q_1 \, dp_j}$ which is the ordinary form of price cross-elasticity of demand which we discussed at length in Chapter II.

Having derived this coefficient in terms of Chamberlin's numbers criterion, Triffin immediately proceeds to use it as his substitution criterion. With it, he is able to distinguish the categories of homogeneous competition, heterogeneous competition and isolated selling. It tells us nothing of numbers at all, for Triffin himself agrees that each of the first two of his categories may be either oligopolistic or non-oligopolistic. Indeed, he introduces another coefficient to distinguish these sub-categories. Nothing could better demonstrate the amazing confusion that has grown up around the concept of cross-elasticity than the fact that Triffin can, without apparent challenge, derive a
coefficient specifically as a numbers criterion and use it at once as a substitution criterion. This he does, not as an aberration on a minor point nor as an error in the footnotes, but at the very heart of his system. Not only does he get away with it, but it is adopted by others. It seems to satisfy both of the major camps. For those who consider cross-elasticity essentially as a numbers criterion there is the explicit derivation from Chamberlin's numbers condition, not to mention a numbers coefficient of Triffin's own which looks very much like cross-elasticity of demand; for those who believe that price cross-elasticity is a substitution criterion there is Triffin's open use of it as such. So far as a numbers criterion is concerned, then, Triffin's development of the Chamberlinian criterion is a dead end.

However, as mentioned, Triffin himself postulates a numbers criterion of his own which he writes $q_i \frac{\partial p_j}{p_j \partial q_i}$. This looks, at first glance, very much like the reciprocal of the ordinary price-cross-elasticity of demand which he uses as his substitution criterion. It cannot be this, of course, for it could tell us no more than its reciprocal. The only other interpretation
is to take it as a coefficient determined by the reverse causal sequence: as the ratio of the percentage price-change induced in firm j as a result of a given percentage change in the output of i. However, a firm can, generally, only increase output by first altering price. Hence, the coefficient must boil down to a measure of the proportionate effect upon the price of j induced by a given variation in the offer price of i. In other words, Triffin does not advance beyond his original statement that "the crux of the distinction between large and small numbers is really whether or not a price move by one seller induces the other sellers to follow suit..." Triffin does not get past definitions. He does not tell us the conditions of demand (substitutability) or the conditions of supply (number of firms) which are likely to give rise to price interdependence.

3. The Determinants of Price Interdependence

We are interested in the effect of a price-change by firm i on the price policy of another firm j. Now it is clear that the initial impact upon firm j will be in the form of a change in j's output. Firm i

7. op. cit., p. 30.
cannot directly force a change in j's price. It can only do so in virtue of the effect its price policy has on firm j's output. Let us call the impact on j's output resulting from a reduction of i's price \( I_{ji} \) which we define as being equal to \( \Delta q_j/q_j \). (The impact on j's output is best expressed as a relative, since a change in output which may be a fleabite to a large firm may be of considerable significance to a small one.) Thus —

\[
I_{ji} = \frac{\Delta q_j}{q_j}
\]

(1)

and since

\[
\frac{\Delta q_j}{q_j} = \frac{\Delta p_i}{p_i} E_{ji}
\]

(2)

therefore

\[
I_{ji} = \frac{\Delta p_i}{p_i} E_{ji}
\]

(3)

We see then that the impact upon the output of firm j occasioned by a price-cut by firm i depends on:

1. The degree of substitutability between the products i and j as reflected by the cross-elasticity of demand \( E_{ji} \); and
2. The size of the price-cut invoked by firm i.
Moreover, this conclusion is in accord with commonsense. Perhaps it is most clearly seen when products are identical physically and only differentiated spatially. For example, if we suppose that consumers are prepared to travel one mile for a shilling, then clearly if a firm cuts price by a shilling it will eat into the sales of other firms within a mile radius of the price-cutting firm; if it cuts by two shillings it will eat into the sales of firms within a radius of two miles; and so on. The impact of the one firm on any other clearly depends on the size of the price-cut and the distance between the firms.

We can demonstrate the same conclusion by reference to the relationship $dd'$ which we postulated in Chapter III for "similar" products. In such a situation, it is clear that any particular customer will transfer his allegiance if the price-line moves sufficiently, i.e. if relative prices are changed sufficiently. And the rate at which individual customers will change over will depend upon the degree by which their indifference line deviates from the line of perfect substitution $aa'$. In short, the impact of a price-cut by one firm on the sales of another product will depend on the size of the price-cut invoked and the closeness of the substitutes.
The influence of the degree of substitutability upon interdependence is more or less recognized by economists, although not without misgivings in some quarters. Most economists would agree that two sellers, one of beer and one of bricks, could be correctly described as "isolated". They would also agree, no doubt, that two adjacent sellers of beer are likely to constitute a duopoly. Clearly, the distinction drawn depends solely on the degree of substitutability in the two cases.

The misgivings which sometimes arise about saying straight-out that oligopoly pre-supposes a high degree of substitutability between the individual products concerned stem from the fact that pure competition involves perfect substitutability; yet pure competition has been traditionally regarded as the very antithesis of oligopoly. Certainly for pure competition the \( E_{ij} \rightarrow \infty \). And it is true that this expresses "the rigid connection between the prices of economically homogeneous goods ...". But the fear that this implies a condition of oligopoly is groundless; for, as Bishop has pointed out, "the purely competitive firm is essentially a quantity-decider rather than a price-quoter. Hence the extreme interdependence as to price that is implied by infinite cross-elasticities

8. R. Triffin, op. cit., p. 140.
has no oligopolistic consequences, in the light of the fact that the purely competitive firm has no significant price decision in the first place."³⁹

On the other hand, the fact that, in the general case of monopolistic competition, the impact of a price-cut of one firm on the sales of another will depend on the size of the price-cut involved has not generally been recognized explicitly. Thus, to take a case at random, Papendreou asks: "How large must the number of firms in an industry be in order that it may fall in the category of a large-number-of-firms industry? ... The number must be so large that firms cannot influence one another's volume of sales by cutting price."¹⁰ However, in the general case of monopolistic competition, we can, theoretically, always make the volume of sales transferred by a price-cut as small (large) as we like, if we make the price-cut small (large) enough.

Two conclusions follow from these facts. Firstly, a "quantity" criterion is the appropriate one in the case of literally homogeneous goods, since their producers are not in a position to make finite changes in offer prices. Secondly, a "quantity" criterion is not applicable in the

general case where substitutes are imperfect, however slightly, because the quantities involved will depend upon the size of the price-changes with which they are associated.

In the case of homogeneous products, the question of the presence of oligopoly must be posed in the form: will a variation in the quantity supplied by any one firm significantly affect market price? Symbolically, if $E_m$ is market elasticity of demand, then

$$E_m = \frac{p_m \Delta q_m}{q_m \Delta p_m}$$

or

$$\frac{\Delta p_m}{p_m} = \frac{\Delta q_m}{q_m E_m}$$

Then, if the addition to market supply by any one firm is $\Delta q_i$ -

$$\frac{\Delta p_m}{p_m} = \frac{\Delta q_i}{q_m E_m}$$

(4)

Therefore, with a given elasticity of market demand, the question whether $\Delta p_m/p_m$ is significantly different from zero depends upon the ratio of firm $i$'s potential capacity to total market supply. It follows as a corollary that some limitation upon the capacity of the individual firm is an essential condition for the existence of pure
competition. Kaldor was surely wrong when he argued, in opposition to Triffin, that pure competition is compatible with constant costs.\footnote{11}

In the general case of monopolistic competition, the concept of market price and market supply dissolves. Moreover, the individual firm can only finitely add to supply by first cutting its price by a finite amount. Hence, the oligopoly criterion poses itself in the form: will a price-cut by any one individual firm $i$ significantly affect the price of any other individual firm $j$? However, as we have seen above, the impact of a price-cut by $i$ on the position of $j$ depends on the size of the price-cut invoked by $i$. This fact, as we have seen, has been the source of a good deal of logical confusion. Triffin attempted to take account of it by taking the ratio between $j$'s loss of revenue and $i$'s initiating price-cut, but this led him directly to a substitution coefficient. Clearly, the correct statement of the criterion should be: will a price-cut by $i$ induce firm $j$ to also cut price by an amount significantly large in relation to $i$'s initial price-cut? This formulation avoids the difficulties discussed above. In the first place, any price move by one seller will theoretically involve some readjustment in the position of all related sellers, however small.

\footnote{11. See Triffin, op. cit., pp. 155-157; and Kaldor's review of Triffin, Economica, Nov. 1942, pp. 409-12.}
However, so long as j's price reaction is negligible compared with i's initial price-cut, firm i can afford to neglect it. At the other extreme, if j reacts by cutting price in the same proportion as i, so that in the outcome the ratio of the two prices remains the same, then the effect of i's price move is completely nullified vis-a-vis j. Let us consider the matter in detail.

We take r firms linked by cross-elasticities, E_{ji}, of the same order. Cross-elasticities with any other firms are zero or so small as to be quite negligible. We wish to find the effect of a price-cut by one firm i upon the price of any related firm j. For the present, we make the further simplifying assumptions:

(1) that all firms act unilaterally in attempting to maximize their profits, i.e. they individually attempt to equate their marginal revenue and marginal cost;

(2) that the demand curve for firms is linear, at least over the relevant range, i.e. that sales are transferred from one firm to another under the impetus of a price-differential at a constant rate;

(3) that marginal cost is constant over the relevant range.

Then the equation of firm j's demand curve in the initial situation is of the form: 
\[ q = k - mp \]  
\[ \text{where } m = q^{0} \frac{E_{ji}^{0}}{p^{0}}, \text{ the upper suffix denoting values at the initial point, and where } k \text{ is a constant.} \]

Now let firm i cut its price by an amount \( \delta p_{i} \), so that \( j \) is involved in a loss of sales \( L = q^{0} E_{ji}^{0} \delta p_{i}/p_{i} \), where \( E_{ji}^{0} \) is the initial value of the cross-elasticity existing between \( i \) and \( j \). Then the demand curve of firm \( j \) will shift downwards and its equation becomes:

\[ q = (k - L) - mp \]

Multiplying through by \( q \) gives

\[ pq = \left(\frac{k - L}{m}\right)q - \frac{q^{2}}{m} \]

The equation of \( j \)'s new marginal revenue curve is therefore:

\[ \frac{\partial (pq)}{\partial q} = \left(\frac{k - L}{m}\right) - \frac{2q}{m} \]

Firm \( j \) will now attempt to expand output to \( q' \) where its new marginal revenue is equal to its marginal cost (which we have assumed to be constant). This new output will be given by:

\[ \frac{k - L}{m} - \frac{2q'}{m} = c \]

where \( c \) is marginal cost. Then

\[ q' = \frac{1}{2}(k - L - mc) \]
Substituting for \( q \) in equation (6) we get:

\[
p' = \frac{r}{m} \left( \frac{k - \mathbf{L}}{m} + c \right) \tag{11}
\]

We are now in a position to calculate the immediate change in \( j \)'s price induced by the price-cut by \( i \). From (5) we have \( k = q^o + m^o \), so that

\[
p' = \frac{r}{m} \left( \frac{q^o}{m} + p^o - \frac{L}{m} + c \right) \tag{12}
\]

Moreover, we know that \( m = q^o E^o / p^o \) and \( L = q^o E^o_j i dp_i / p_i \) and hence

\[
p' = \frac{m}{E_{ji}} \left( p^o + p^o \frac{E^o_{ji}}{E_{ji}} dp_i / p_i + c \right) \tag{13}
\]

Now in virtue of our assumption that firms attempt to maximize their profits by unilaterally equating marginal cost and marginal revenue, we may write \( p^o / E^o_{ji} = p^o - c \) and hence:

\[
p' = \frac{m}{E_{ji}} \left( 2p^o - p^o \frac{E^o_{ji}}{E_{ji}} dp_i / p_i \right) \tag{14}
\]

and

\[
p^o - p' = dp_j = \frac{m}{E_{ji}} \left( 2p^o \frac{E^o_{ji}}{E_{ji}} dp_i / p_i \right) \tag{15}
\]
Expressing the relative change induced in j's price, in the first instance, in terms of the relative price-cut initiated by i, which ratio we call \( D_1 \), we get:

\[
D_1 = \frac{\frac{\partial p_j}{\partial p_i}}{\frac{p_j}{p_i}} \times \frac{E^0_{jj}}{E^0_{jj}} \tag{16}
\]

However, having cut its price in this proportion, firm j will find that it has not achieved its aim of an output of \( q' \). Since we are assuming a symmetrical situation in which all firms are unilaterally attempting to equate marginal cost with marginal revenue, the firms k, l, m ... will also react in a manner similar to j, with the result that j will only have succeeded in increasing its output to the extent to which it has been able to regain sales from i. This it will do in the proportion which its price-cut matches i's. (When j's percentage price-cut is equal to that of i, j should have won back all customers lost to it, since the ratio of their prices will now have been restored vis-à-vis the initial situation.) Let the sales regained from i by the first price-cut by j be \( R_1 \) so that \( R_1 = \frac{\partial p_j}{\partial p_i} \times L \), \( L \) being j's original loss of sales to i. Then, from (16),

\[
R_1 = D_1 L.
\]

In the new situation, j's demand curve will be:
where \( k' = q^o - L + R_1 + mp' \). Since \( j \) will have been disappointed in its attempt to restore an equilibrium output (because related firms have moved with it), it will, so long as it is concerned only with its own marginal cost and marginal revenue position, try to attain an equilibrium (maximum) position by resort to a further price-cut. Proceeding as before, we obtain the new equilibrium price:

\[
p'' = \frac{1}{2}(k'_m + a).
\]

Since \( k' = q^o - L + R_1 + mp' \),

\[
p'' = \frac{1}{2}(\frac{q^o}{m} - \frac{L}{m} + \frac{R_1}{m} + p' + a).
\]

Further, since \( R_1 = D_1L \) and \( \frac{q^o}{m} = \frac{P^o}{E_{jj}} = p^o - c \)

\[
p'' = \frac{1}{2}(p^o + p' - \frac{L}{m} \cdot 1 - D_1)
\]

Moreover, \( \frac{L}{m} = 2p^oD_1 \frac{dp_1}{p_1} \), so that

\[
p'' = p' - p^oD_1(1 - D_1)\frac{dp_1}{p_1} + \frac{1}{2}dp_1,
\]

where \( \delta p_1 = p^o - p' \). Hence

\[
p' - p''/p^o = \frac{\delta p_2}{p^o} = D_1(1 - D_1)\frac{dp_1}{p_1} - \frac{1}{2}\delta p_1/p^o
\]
so that

\[ d_2 = \frac{\frac{\partial p_j}{\partial x_j} - \frac{\partial p_j}{\partial x}}{p_j} = d_1(1 - d_1) - \frac{1}{2}d_1 \]

\[ = d_1(\frac{1}{2} - d_1) \]  \hspace{1cm} (18)

* * * * *

Again, firm j will fail to achieve its target output - although it will have got closer to it - because of the simultaneous unilateral action of other related firms. Repeating the above argument, it can be seen that firm j will now wish to invoke a further price-cut in an effort to attain an equilibrium (unilaterally maximum) position. It can be shown that this third-round price-cut will be:

\[ d_j = d_1(\frac{1}{2} - d_1)^2 \]  \hspace{1cm} (19)

* * * * *

Thus, firm j will, theoretically, move to its new equilibrium position by a series of price-cuts of diminishing magnitude. The total proportionate price-cut which j will make in response to a price-cut by i may, therefore, be expressed in the form of the infinite series:

\[ D = D_1 + D_2 + D_3 + \ldots \]
This series must converge, of course; otherwise no new position of stable equilibrium would be attainable. In fact, the series (20) has a common ratio of \( \frac{4}{5} - D_1 \). Moreover, since \( D_1 = \frac{5E_{jj}}{E_{jj}} \) and \( E_{ji} \leq E_{jj} \), therefore \( D_1 \leq \frac{5}{4} \). If \( E_{ji} = E_{jj} \), as in a symmetrical duopoly situation in which firm \( j \) is related to firm \( i \) and to no other in any degree whatsoever, then \( D_1 = \frac{5}{4} \) and \( D = D_1 \). In other words, the first-round price-adjustment is all that is necessary to achieve a new equilibrium. In general, however, \( D_1 < \frac{5}{4} \) and the common ratio \( \frac{4}{5} - D_1 \) is positive and less than \( \frac{5}{4} \). Hence, generally, the series converges to the limit:

\[
D = \frac{D_1}{\frac{5}{4} + D_1}
\]

or

\[
D = \frac{1}{1 + E_{jj}/E_{ji}}
\]  

(21)

Moreover, since in the symmetrical situation assumed \( E_{jj}/E_{ji} = r - 1 \) (the relationship expressed in Bishop's "numbers equivalent" and discussed in Chapter II), therefore:

\[
D = \frac{1}{r}
\]  

(22)

\( r \) being the number of firms in the group.
Thus, we find that, on the assumptions we have adopted, the relative price-reaction of a firm which will be induced by a price-cut by a related firm varies inversely with the number of firms so related or, if you like, the number of firms in the "group". This is the link between numbers and the transmissability of price-changes inherent in Chamberlin's analysis. Going back to relationship (21), the precise significance of Bishop's "numbers equivalent" is brought out. It is in fact directly related to the Chamberlinian concept of numbers but, as Bishop himself has pointed out, has the advantage of being applicable to asymmetrical situations in which the mere counting of firms would be misleading. At this level and on these assumptions, it will also be seen that Bishop is correct, as against the many writers who have taken cross-elasticity of demand as a numbers criterion (high cross-elasticities imply oligopoly, low cross-elasticities the absence of it), in asserting that it is the ratio of cross-elasticity and own-elasticity which is significant in regard to Chamberlin's conception of numbers. Directly, however, it will be shown that at another level - on different assumptions - cross-elasticity does become the important factor. It is almost certainly this fact which has given rise to much of the confusion in this matter.
In fact, the analysis developed above does not take us very far, mainly because of the restrictive nature of some of the assumptions. However, before proceeding upon more general assumptions, i.e. before we relax our more fundamental conditions, it is worth pausing to consider the consequences of varying some of our minor assumptions.

Firstly, we assumed marginal cost to be constant over the relevant range of output changes. The algebra necessary to obtain a similar solution for variable marginal costs is much heavier even than that involved in the result obtained above and will not be undertaken here. However, it may be said in passing that, if marginal costs are rising in the relevant range, the price-reaction will be greater than for constant costs; if they are falling, the price-reaction will be smaller.

Secondly, in our analysis above, we assumed that second-order cross-elasticities were negligible and could be ignored. Insofar as the sum of all second-order cross-elasticities ("market" elasticity of demand) is significant, this will damp down the price-reaction by firms related to the price-cutting firm by cross-elasticities of first-order significance. The reason for this is fairly obvious. The more sales which intra-group firms can attract from extra-group firms as they successively cut
price in the search for a new equilibrium, the less they will have to cut their price in order to restore an equilibrium output, with marginal cost and marginal revenue equated.

More fundamental to the whole argument is the assumption that firms unilaterally seek to equate their marginal cost and marginal revenue. Indeed, it might be thought at first sight that this assumption implies the answer, since it is equivalent to assuming that firms act "non-oligopolistically". In fact, the argument is of the reductio ad absurdum type which purports to show:–

(1) that unilateral maximizing behaviour by a large number of firms represents consistent behaviour;

(2) that such action by a small number does not, because in this latter case individual firms will become aware that their own price moves will produce significant reactions in the price policies of competitors.

Nevertheless, such an argument cannot sustain the converse: that the absence of small numbers necessarily implies that firms will attempt unilateral maximization by equation of their individual marginal revenues with their marginal cost.
For this reason the argument so far leaves a large question mark in the air. Should we consider any firm which does not equate marginal cost and marginal revenue as "oligopolistic", *ex hypothesi*? In terms of the traditional concepts of Chamberlin and Mrs. Robinson, it would appear that the answer should be "yes"!

However, I think there are good reasons for not leaving the matter there -

(1) This delineation would cut across the Chamberlin-Triffin definition of oligopoly in terms of price interdependence. It is both conceivable and possible that a firm should not attempt to equate its marginal cost and marginal revenue, and yet should be sufficiently insulated from other firms not to be forced to match their price moves. (e.g. firms which are ignorant of the shape of their demand curve, or firms faced with a discontinuous demand curve such as that derived in Chapter III.)

(2) Oligopoly, with its literal connotation of "fewness", would be compatible with a substantial number of firms.

(3) All firms operating in such manner as posited by Andrews, for example, would be oligopolists, by definition.
Therefore, it becomes necessary to examine further the problem of the interdependence of the price-output policies of related firms. To this end, we shall now ask the question the other way round: under what conditions will a firm follow a price-cut in full (i.e. reduce its price in the same proportion as the firm initiating the price-cut) and for certain. We will retain our minor assumptions:

(a) that marginal cost may be taken as constant over relatively small changes in output; and

(b) that second-order cross-elasticities are negligible compared with first-order (intra-group) cross-elasticities.

However, we no longer posit the assumption that firms necessarily attempt to achieve a maximum position by unilateral action based upon their individual marginal revenues and marginal costs. Rather, we ask: under what circumstances will a firm alter its price-output policy, on whatever basis it may have reached its initial position, in response to a move by a competitor.

4. Full and Certain Price Retaliation

We shall begin with two firms i and j. We have to determine under what conditions a price move by firm i will be followed by an equi-proportional price move by
firm j. Now if firm i cuts price, firm j will, other things being equal, be certain to follow the price-cut if its net revenue, after adoption of this course, is greater than it would be if it stuck to its original price.

If firm j does not follow i's price-cut, its net revenue will be reduced by -

$$\partial q_j (p_j - m_j)$$

(A)

where $\partial q_j$ is the output lost to firm i and $m_j$ is j's marginal cost.

On the other hand, if firm j makes the same proportionate price-cut as that initiated by i, it will, it may be assumed, retain the customers which it would have otherwise lost to i, since the ratio of the prices will be unchanged. Then, the net revenue of j will be reduced to the extent -

$$q_j \partial p_j$$

(B)

where $\partial p_j$ is the retaliatory price-cut forced on firm j.

It is clear then that j will follow the price-cut, if (A) is greater than (B), i.e. if:

$$\partial q_j (p_j - m_j) > q_j \partial p_j$$

(23)

Moreover, we have assumed that $\partial p_j/p_j = \partial p_1/p_1$, so that
\[ \frac{\partial q_j}{\partial p_i} = \frac{p_j}{p_i} \frac{\partial p_i}{\partial p_i}. \]

Then substituting in (23) we get:

\[ \Delta q_j(p_j - m_j) > \frac{p_j}{\partial p_i} \frac{\partial p_i}{\partial p_i} \]

or

\[ \frac{p_j}{q_j \partial p_i} > \frac{p_j}{p_j - m_j} \tag{24} \]

Now the expression on the left-hand side is clearly the cross-elasticity of demand \( E_{ji} \), so that (24) may be re-written:

\[ E_{ji} > \left( \frac{p}{p - m} \right)_j \tag{25} \]

* * * *

We must now generalize this result by considering the effect of increasing the number of related firms. Let us now take four firms \( i, j, k \) and \( l \). Now if firm \( i \) cuts its price, there are eight possible results:

<table>
<thead>
<tr>
<th>Firms Cutting Price</th>
<th>Firms Maintaining Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>( jkl )</td>
</tr>
<tr>
<td>( ij )</td>
<td>( kl )</td>
</tr>
<tr>
<td>( ik )</td>
<td>( jl )</td>
</tr>
<tr>
<td>( il )</td>
<td>( jk )</td>
</tr>
<tr>
<td>( ijk )</td>
<td>( l )</td>
</tr>
<tr>
<td>( ijl )</td>
<td>( k )</td>
</tr>
<tr>
<td>( ikl )</td>
<td>( j )</td>
</tr>
<tr>
<td>( ijk )</td>
<td></td>
</tr>
</tbody>
</table>
The important difference between this case and that of only two sellers is that, in this situation, firm j will, when calculating whether or not to follow a price-cut by i, have to take account of the fact that, if it does cut price, it may not only forestall any transfer of sales to i, but may also attract sales from k and l, provided these firms do not also follow i's price-cut.\(^{12}\)

In the general case, there will be \(2^{n-1}\) alternatives, where \(n\) is the number of firms related to i. The task of calculating all these separate cases would, of course, be prohibitive - both in theory and practice - even if we were given all the relevant cross-elasticities. How then can we determine the basic determinants of j's reaction-pattern?

We could, of course, postulate symmetry assumptions, leaving it to the business-men to do the calculating. Then, if one firm follows, they all follow, since their calculations are identical. It is easy to establish the the condition for all firms following, whether they take the decision individually or collectively, is that of our inequality (25).

But this is, I feel, to lean too heavily upon symmetry assumptions. A more realistic approach suggests itself. Now it is clear that if firm j's worst position after following i is better than its best
position by not following, then \( j \) will be certain to follow. It is clear also that the worst position for \( j \) after cutting price is when all other firms related to \( i \) and \( j \) do likewise. On the other hand, \( j \)'s best position for maintaining price is when firms \( k, l \ldots n \) also maintain price. Then, it is easy to verify, on this basis, that \( j \) will be certain to follow, if:

\[
E_{ji} > \left( \frac{p}{p - m} \right)_j
\]

which is the condition of our inequality (25).

Of course, this in no way implies the converse. Some firms in some situations with cross-elasticities lower than that required by (26) may have strong reasons for following price-cuts fully and unhesitatingly. However, these borderline cases represent the genuine area of "oligopolistic uncertainty" which, in practice, is probably small. It will be noticed that the calculation required of firms is of the "greater or lesser" variety. Most firms in most cases will know whether their cross-relationship with other individual firms (products) is significantly large or negligibly small. It is only in the intermediate cases that uncertainty will arise.

5. Some Results

The first interesting aspect of the above result
is that the expression on the right-hand side of (26) would, if firms were unilaterally equating marginal cost and marginal revenue, be equal to the elasticity of demand, $E_{jj}$. The condition of the inequality, therefore, would require that the cross-elasticity, $E_{ji}$, be greater than the own-elasticity, $E_{jj}$, which, in general, is impossible. This is consistent with our earlier result that, when firms are so equating marginal cost and marginal revenue, the greatest extent by which a price-cut by one firm will be matched by another is one-half, in the case of duopoly. In short, full (proportionate) price-reaction will never occur if firms are equating marginal cost and marginal revenue unilaterally.

It is an obvious and interesting corollary of this fact that, where firms do match the price-moves of other firms fully (i.e. in the same proportion), then we have prima facie evidence that the firm is not unilaterally maximizing profits and that the situation is one of "oligopoly".

The second thing to notice about the above result is that, once we drop the assumption of unilateral action by firms, it is cross-elasticity of demand which becomes important in determining whether or not a price-cut by one firm will induce other firms to follow suit. It is almost certainly this fact that has led to the
widespread belief that cross-elasticity is, of itself, an adequate numbers criterion.

Moreover, in the light of the above discussion, a number of other propositions familiar in economic theory can be brought together in harmony. Thus, the fact that marginal cost — or rather the level of marginal cost — is an important determinant of price interdependence lends colour to Efroymson's thesis that the direction of price-interdependence may vary cyclically. ¹３ Thus, the normal "obtusely" kinked demand curve, normally associated with oligopoly, is likely to have strongest application in depressed business situations where there is a good deal of excess capacity. By contrast, the curve may become of the "reflex" type — when price increases are followed, price-cuts are not — as may occur at high levels of employment when output is pressing against capacity. In terms of our result (26) this may be re-expressed in terms of marginal cost. When marginal cost is low, relative to price, firms have a strong incentive to chase after low cost sales; when marginal costs are very high, firms may be prepared to follow price-moves of competitors.

upwards, deciding upon the price increase as the more important as against the possibility of losing some high-cost (marginal) sales to rivals who might not follow the movement upwards.

On the demand side, we are also able to shed some interesting light on a problem which has been the subject of some debate in economic literature. This case originally arose from Mr. Harrod's so-called "Law of Diminishing Elasticity of Demand, which means that, as output as a whole increases and individuals become more affluent, their sensitiveness to price differences declines." In virtue of this "law", Harrod suggests that markets are more perfect in the slump, less perfect in the boom. Monopoly power therefore varies directly with the level of output over the course of the trade cycle.

However, increasing sensitivity of customers to price differences can only be reflected in increasing cross-elasticities of demand — firms (products) come closer together. And increasing cross-elasticities tend to make for stronger price interdependence. Thus, we arrive at the paradoxical notion that, as the

"perfection" of the market increases (in the slump), so does the degree and certainty of "oligopolistic" price-interdependence tend to increase. Incidentally, this tendency will be reinforced if, during the slump, marginal costs fall relative to price as compared with the boom, when firms may be working near capacity outputs.

The tendency for price-interdependence to increase in the slump has been noticed by a number of writers. And because of this increased interdependence, it is also noticed that the area of agreements, both overt and tacit, is widened, and existing agreements are strengthened. Thus, discussing the Harrod thesis, Kalecki writes: "Probably more important than the effect of cyclical changes in market imperfection upon μ is the influence of 'tacit agreements' in a deep slump which may be classified as changes in the degree of oligopoly."

The important thing to notice that, in the light of our conclusions above, there is no necessary contradiction in the two points of view. Increased substitutability between products, far from being

Incompatible with increased price-interdependence, is the most likely factor to give rise to it. But the most interesting aspect of the above argument is its implication for those theses on price formation which depend upon some ad hoc or conventional formula. This can best be illustrated by reference to the work of Andrews. Andrews, it will be recalled, believes that the typical behaviour in manufacturing industry is to calculate price as average prime cost plus a "costing margin" which includes an allowance for overheads and a margin for profit. Also, business-men take average prime cost as a good approximation for marginal cost over the significant ranges of output. Then, if average prime cost (= marginal cost) amounts to 80 per cent of price, any cross-elasticity of demand greater than 5 must, in the light of (26) above, induce price-interdependence. If average prime cost is 50 per cent of price (as it is in many industries), then any cross-elasticity greater than 2 will involve price-interdependence. If, as Andrews always suggests, most manufacturers are in competition with close substitutes, then most manufacturers, on his thesis, are "oligopolists" in the Chamberlin-Triffin sense.

We may or may not agree with the appropriateness of this conception of "oligopoly". I am sure, however, that it would not evoke Mr. Andrews' approval. Yet it is the usual, if not the most consistently applied, view!

6. Oligopolistic Uncertainty.

Often, the problem of oligopoly is put forward in terms of "uncertainty". The basis of the well-known kinked oligopoly curve is usually described: the curve is elastic upwards because firms fear that, if they raise price, they will not be followed by rivals; it is inelastic downwards because firms fear, if they cut price, they will be followed.

The relationship (26) derived earlier purports to express the conditions under which a price-cut is certain to be followed. In the real world, there is bound to be many situations in which a firm has little illusion about what other firms will do. It can be sure that firms close enough will follow a price-cut; and that firms far enough away will not. Hence, it is not true to say that the essence of oligopoly is uncertainty. However, it is perfectly true that in between the two cases mentioned - the very close and the far away - there will be a wide area of uncertainty.
From the point of view of the firm, uncertainty may derive from two main sources: (i) uncertainty of knowledge about demand and cost; and (ii) uncertainty about the actions of rivals. Latterly, increasing emphasis has been placed on the former. Indeed, ignorance of the shape of the firm's demand curve and of the firm's marginal cost has been represented as a chief reason why it does not attempt to maximize profits by equating marginal cost with marginal revenue. This argument is not altogether convincing. On the one hand, it must be recognized that in a dynamic economy prices and costs do change considerably from time to time. Further, seldom are adjustments in the firm's position carried out instantly and simultaneously with the changes in basic data. Consequently, one would expect that an entrepreneur would, on the basis of experience in a changing world, form a fairly accurate picture of the shape of his demand curve. Again, it is difficult to explain the disinterest of cost accountants and businessmen in marginal cost in terms of their inability to calculate it. In any case, however one may rule as to the importance in practice of the argument, it is of

17. Hall and Hitch, loc. cit. (for example).
doubtful value theoretically. Lack of knowledge of this sort is a limitation which affects many facets of economic behaviour and is the reason that most economic propositions are stated as "tendencies". The problem is therefore a general one and not specifically one of "oligopoly".

More central to the oligopoly problem is the uncertainty surrounding the probable actions of rivals. In the earlier formulations of monopolistic competition theory this "poker game element" was given pride of place. The relationship (26) derived above purports to show that a firm is "certain" to follow a price-cut by another firm, if the condition of the inequality holds. The firm is certain to act in this way in the sense that, in terms of cost and revenue, it will be better off in the second position than in the first, i.e. maintaining price, irrespective of how other firms react. However, if this inequality does not hold - as indeed it never will if firms are equating marginal cost and marginal revenue - it will pay firms collectively not to follow the price-cut. In other words, it will pay any one firm to hold its price provided other related firms hold theirs. (This is not, of course, to rule out the possibility that in practice firms may sacrifice short-period gains and follow a price move in the interests of long-period gains.)
Thus, whether firms in this position will hold their prices in the face of price-cutting by one particular firm depends upon whether they are in agreement — either overtly or tacitly — or not. Clearly, the likelihood and reliability of such agreement will depend principally on the number of firms involved — the smaller the number, the more likely is concerted action. This is a fairly well-recognized phenomenon: it is easier to maintain discipline in a small group than in a crowd. Thus we arrive at a somewhat paradoxical result. Where firms are related in this second, looser manner, the probability of price retaliation becomes less, the fewer the number of firms concerned.

From the point of view of the economist — interested in the prediction of a firm's behaviour in various circumstances — there is a second sort of uncertainty, viz. uncertainty about the motives, at any particular time, of the firm itself. Mention has already been made of changes in price-output policies for "tactical" reasons. Once a firm is in a situation where it can do better by agreement or cooperation with rivals than by unilateral action, tactical considerations become of first importance. On the one hand, they will be directed to preserving agreement so long as agreement is tenable; on the other hand, they will be directed to obtaining the
best strategic position in the event of agreement breaking down, of war breaking out. As Rothschild has pointed out, the analytical starting point for this type of study is not Newtonian mechanics nor Darwinian biology, but Clausewitz's Principles of War. This involves, of course, the introduction of many factors other than cost and demand, factors with which the economist is not accustomed to deal, if prediction of the firm's behaviour is to be accurate and realistic.

Where there is a definite interdependent group, as distinct from pure monopoly, duopoly or chain relationship, and the group-effect is manifest in a curve of the nature derived in Chapter III, this uncertainty must be magnified, because with a discontinuous demand curve of this type the firm cannot attain any position of stable equilibrium by unilateral action. While in peace firms may seek haven by directing their attention to the upper segment of the curve, the lower segment remains an ever-present invitation to war. The implications of this type of curve for the theory of price formation will be discussed more fully in Chapter V which follows. It is sufficient

here to notice that this type of relationship will give rise to uncertainty of the kind usually associated with oligopoly, although a large number of firms may be involved.

7. Oligopoly and "Bigness".

Another large source of confusion about the oligopoly problem arises from the pre-conception that oligopoly is closely related to "bigness". On the one hand, we envisage an oligopolistic industry as consisting of a handful of entrenched industrial giants, protected from outside competition by technical economies of scale, a network of patent rights, financial dominion or exclusive access to vital raw materials. On the other hand, from a formal, analytical point of view, derived from the criterion of demand inter-relationship, we may classify two milk-bars on opposite sides of a street in the same category as two automobile manufacturers of empire-like proportions.

Clearly, the latter, for many purposes, are something quite outside the world of milk-bars. Giant oligopolists, unlike milk-bar proprietors, may make and unmake governments. As Rothschild has it, "the
oligopolistic struggle for position and security includes political action of all sorts right up to imperialism.\textsuperscript{19}

It has almost certainly been a mistake to use the term "oligopoly" to describe all situations other than those in which a firm acts upon the assumption that its demand curve is independent of the actions of other individual firms or in which a firm does not unilaterally attempt to equate its marginal revenue and marginal cost because of its consciousness of being a member of a group. Such situations involve a very much larger area than those in which the firms making up an industry are both large and few. Confusion will persist so long as the one term is used to describe a whole range of situations involving many radically different characteristics.

8. Conclusions

On the basis of the discussion of this and the previous chapter, we could make the following formal classification of market relationships:

\[\text{Classification of Market Relationships}\]

\textsuperscript{19} K.W. Rothschild, \textit{loc. cit.}, p. 319.
<table>
<thead>
<tr>
<th>Substitution Criterion \n</th>
<th>Numbers Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>E_{ji}</td>
<td>\gamma = E_{ii}/e_{ii}</td>
</tr>
<tr>
<td>Pure Competition</td>
<td>\rightarrow \infty</td>
</tr>
<tr>
<td>Large Group (Differntiated Competition)</td>
<td>High*</td>
</tr>
<tr>
<td>Small Group (Oligopoly)</td>
<td>High</td>
</tr>
<tr>
<td>Duopoly</td>
<td>High</td>
</tr>
<tr>
<td>Independent Selling (Pure Monopoly)</td>
<td>\rightarrow 0</td>
</tr>
</tbody>
</table>

This classification differs from an interesting alternative suggested by Professor Bishop in respect of both the substitution and numbers criteria.

Cross-elasticity is used for the substitution criterion in preference to own-elasticity as proposed by Bishop. The main advantage of cross-elasticity for this purpose is that it represents a better measure of the relationship between individual firms than is own-elasticity. It is especially superior for the description of oligopolistic situations because it is an important factor determining the

20. R.L. Bishop, loc. cit., p. *i.e. significantly different from zero.
impact which the price-output moves of one firm will have on another firm, a matter we discussed at length in Section 4 above.

In the matter of the numbers criterion, a case could be made out for the use of Bishop's "numbers equivalent" in conjunction with cross-elasticity as a substitution criterion. As was shown in Section 3 above, a high Bishopian "numbers equivalent" - i.e. a high ratio of own-elasticity to cross-elasticity - is a necessary condition for the absence of significant price-reactions by competitors. On the other hand, the coefficient of numbers or density derived in Chapter III brings out a fundamental characteristic of the firm's demand curve as we move from the extreme of pure competition, on the one hand, and pure monopoly or independent selling, on the other.

Formal classifications such as those proposed by Bishop and others from time to time, including the classification set out above, probably serve little real purpose, except perhaps an historical one. They do purport to show what conditions are necessary for the unilateral action along marginalist lines which is the behaviour traditionally assumed for the firm in the majority of instances.

If the analysis of this Part serves any significant
purpose it is to show that, apart from the independent seller and the quantity-decider of pure competition, the likelihood of situations in which firms will usually attempt to maximize profits by unilateral action to equate their individual marginal revenues with marginal costs is small, a result, incidentally, which would accord with the general findings of empirical studies and the practices of cost accountancy. Such a conclusion would leave a large area of the economy in which firms are not likely to act unilaterally and marginally, without thought to their group existence. The basic question then becomes: Should the whole of this area be described as "oligopolistic" merely for this reason? If not, what principles should be invoked to further sub-divide this area?

Certainly, attempts to delimit the area of "true" oligopoly by use of mechanical definitions such as that used by Chamberlin and Triffin, viz. that "the crux of the distinction between large and small numbers is really whether or not a price move by one seller induces the other sellers to follow suit ..." does not meet the case, although this may sometimes be one of the characteristics of an oligopolistic situation.
As we have seen, there is no simple connection between the fewness of firms and the interdependence of prices. Quite a large number of firms operating in the Andrews manner would be quite consistent with interdependence of their prices. On the other hand, the fewness of firms may actually assist the maintenance of prices by rivals in the face of price-cutting by one firm. In other words, sometimes failure to follow a price cut may be indicative of small numbers. Further, because of tactical and strategic considerations, a firm may resort to price retaliation in one set of circumstances, but not in another. In short, a firm may not follow a price move today, but may well do so tomorrow. More generally, one would expect a firm's behaviour to be radically different, depending upon whether it is at peace or war with its competitors.

Perhaps the fundamental weakness of most systems of market classifications is that they are usually cast in terms of the demand relationships existing between firms. But we cannot ignore the cost characteristics of the firms concerned. To do so, is to use only one blade of Marshall's scissors. We have already seen that even the matter of interdependence of prices is not independent of the
cost position of the firms involved. Again, the relationship between numbers and size which we briefly discussed earlier must largely be described in terms of economies of size related to a given market area. What we urgently require for a realistic description of markets, and in particular for the description of oligopolistic situations of various kinds, is a comparative anatomy of industrial structures based on objective considerations of demand and cost from which the probable behaviour of firms may be inferred. To attempt a description in terms of the behaviour of firms is to put the cart before the horse. This is easily enough said, of course. The trouble is that we still lack the tools for making such an objective description. Indeed, we may well conclude this chapter as it began, by asking: what is oligopoly?
V. ELEMENTS OF A SYNTHESIS

In this final chapter of this Part, a brief indication will be attempted of the signposts which point in the direction of a possible synthesis of the various fragmentations which make up present-day value theory.

1. The Theory of Consumers' Behaviour

One of the main purposes of Chapter III and the Appendix thereto was to show - I think conclusively - that to reconcile the theory of consumers' behaviour and the assumptions of monopolistic competition a new approach is required. Whereas, traditionally, this subject has been a field of application solely of the mathematics of the continuous, it is necessary, to handle the type of consumer behaviour posited by monopolistic competition theory, to introduce the mathematics of the discrete. If indeed consumers have ordered schedules of preferences as between "similar" commodities and substitute between them integrally under the impetus of price-changes, there is no other way out. Whether advance along these lines would merely increase the complexity of an already complex - and abstract - branch of theory, or whether it could result in a more realistic analysis that would provide
us with sorely needed additional tools, is an open question.

If the attempt to state a more general theory of consumers' behaviour were to lead, for example, to a more adequate development of community indifference curves or some analogous concept, the result might well be a gain in simplicity and applicability. A satisfactory solution of the many problems associated with such curves would go a long way towards bridging the gap between the theory of consumers' behaviour and the sort of demand characteristics assumed by monopolistic competition theory. Gross-elasticity of demand, and the interdependence of firms, could then be directly linked with the appropriate indifference curves and this would provide the required bridge between the two branches of theory. On the other hand, it may be necessary to develop a parallel theory of consumers' behaviour based upon the mathematics of permutations and combinations. In any event, however, the matter must ultimately rest very much with the mathematicians.

2. Non-marginalist Theories of Price Formation

It is, I feel, generally agreed that there is little difficulty in reconciling the practice of "full cost" or "average cost" pricing with the theory of oligopoly.
Under conditions of oligopoly, firms will not be expected to act unilaterally in normal circumstances, i.e. in times of peace, although they may do so in the event of war. Therefore, in greater or lesser degree, the emphasis shifts from the firm to the (oligopolistic) group. Since group elasticity of demand will generally be small, it will be in the interests of the firm to seek to maintain the greatest margin between cost and price either by open or tacit agreement. The ceiling on the price which may be safely charged will be set by the potential or actual threat to the group from new or outside competition. Then, since entry is a function of "full cost" - the new firms must cover total costs - existing firms, or more usually the price-leader among existing firms, will have every reason for deciding upon their price policy upon the basis of "full cost". Indeed, precisely the same conclusion emerges as soon as we attempt an analytical definition of "full cost".¹

The real difference of opinion does not, I think, lie in this direction. Rather the puzzle is to account for the generality of the practice of "full cost" price.

¹ See Chapter VII below.
formation (or variants thereof), to explain why a seller in a "non-oligopolistic group" does not attempt to equate marginal revenue with marginal cost. The problem is to reconcile the theory of the "large group" with the practice of "full cost" pricing.

There are two ways in which this difficulty may be answered. Firstly, it may be held that, in fact, oligopoly is quite general. This, for example, is the position taken by Hall and Hitch. For them, "...the typical case is that of monopolistic competition with an admixture, which is usually large, of oligopoly." However, this is not a complete answer unless we can also show why oligopoly should be so ubiquitous.

The second answer would be to concede that there are situations broadly corresponding to the assumptions of the "large group"; but to posit additional reasons why this does not lead to unilateral action resulting in the equation by firms of marginal cost and marginal revenue. This is broadly the line taken by Andrews.

The analysis which we have undertaken in this Part puts us in both camps, although, of course, for different reasons. On the one hand, we showed in Chapter IV that

literal oligopoly — in the sense of a "few" firms sharing the total "market" — is merely the leading species of a large genus. (Chain relationship, for example, is another such sub-group which may be clearly distinguished.)

However, more fundamental than this is that our analysis of Chapter III helps explain why oligopolistic forms should be so widespread, why markets should, generally, tend to break up into smaller areas. Usually, the explanation for this tendency runs in terms of costs only. Economies of scale are such that in many instances there is only room for a few firms producing at economic cost in each industry. Moreover, it is almost certainly from this argument that the identification of "fewness" with "bigness" has arisen. By contrast, we can offer a supplementary explanation of the growth of oligopolistic forms from the demand side. We have seen that, as a small element of product differentiation is introduced into a large group, the demand curve will become discontinuous, with the downward elasticity greater than the upward. One consequence of this is that firms will tend to concentrate upon the one segment of the curve or the other. Some firms will look to the upper segment of the curve, where it will cater for the less sensitive demand at a higher price.
Other firms will direct attention to the downward segment representing the lower-priced mass demand. (Often, of course, the one firm will operate in both directions by marketing a speciality line and a mass-appeal line at a lower price.) Thus, it may be seen that, once product differentiation enters upon the scene, further differentiation is likely to follow. Product differentiation, in short, tends to beget further product differentiation. As a result, the market for any "product" will tend to be broken-up into smaller areas. This process is likely to continue so long as there is a significant asymmetry on the demand side, i.e. so long as there is a "large number" relationship between firms. Advertising is also likely to be an integral part of the process of break-up. The net result of such trends will be, of course, to split the market into smaller and smaller sub-sets of firms which are more or less insulated one from the other. The end product, therefore, is almost certain to involve "oligopolistic" relationships.

However, the process outlined is probably not continuous and is likely to develop by fits and starts. It would probably involve a fairly long period in which to work itself out and would be continually subject to shifts induced by spontaneous changes in consumer behaviour or the effects of advertising. Hence, a demand
curve such as that postulated in Chapter III may be found in practice. Although this situation is essentially a transitional stage — although of possibly lengthy duration — it is interesting to consider the consequences of such a curve for the theory of price determination.

Speaking of a "reflex" curve of this type — which, however, he posutlates for quite different reasons3 — Efroymson says: "the entrepreneur may be pictured moving in a succession of leaps toward the profit maximizing solution; but each time, just as he arrives at the new solution, the profit maximizing solution moves swiftly away, probably in the direction from which he jumped."4 This assumes, of course, that firms are acting unilaterally in their search for a maximum position — by no means a necessary state of affairs. But even if this were so the picture is too static! In practice, there is likely to be, at any time, a constellation of other factors in the situation which will make one or the other segment of the curve more attractive. For example, expectations about the

3. Efroymson’s reflex curve is posited as an inverse kinky oligopoly curve — where rivals follow price-rises but not price-cuts.
future course of prices will be important. Again, considerations of long-period strategy are likely to exercise a large influence - the old problem of short-period versus long-period maximization. Moreover, in the face of general changes in basic data as distinct from those changes specific to one firm, the pressures upon firms are likely to be impulsive for all in the same direction.

In the real world, stability in the price-output relationships of firms in such a situation would most likely occur as a result of tacit agreement or the operation of trade convention. Operation along the lower segment of the demand curve will be ruled out as "ungentlemanly" or as "unfair competition". By contrast, the convention or agreement that each firm should look to the upper segment of its demand curve vis-a-vis its policy towards rivals would result in higher prices all round. And there is some evidence that this is indeed how firms normally act. Thus, D.C. Hague has, as a result of a Hall-Hitch type of examination of business behaviour, concluded among other things that: "the main occupation of most entrepreneurs is keeping the margin between these average cost and price as wide as possible at any given level of output."  

This is, I believe, the real essence of Kahn's "second degree of collusion" - of Professor Fellner's esprit de corps, which it would pay the individual to ignore but which 'yet he does not violate because he places above individual advantage good standing in the group, or the interests of the group.' But this second degree collusion derives its force not from any altruistic esprit de corps or pure patriotism towards the group (although appeal is almost certainly made to such notions as part of the system of sanctions), but simply because all firms happen to be in the same boat. (If one or a few firms get the upper hand at any time, war is very likely to break out.) Competitors will certainly nibble at each other's markets insofar as they can get away with it. The whole system of non-price-competition, secret rebates, selling expenditures, quality variation, and so forth, bears witness to this fact. But this will only generate into open price warfare in exceptional circumstances. The price weapon is much too overt and has too many unpleasant implications for normal use.

Moreover, it is easy to see why such emphasis is placed by firms - and particularly price-leaders - on

---

the notion — however formulated — of "full cost" and why this conception is the chief preoccupation of cost accountants. By and large, the lower segment of the individual firm's demand curve will also approximate — in elasticity — to the demand curve facing potential entrants into the industry. There are, therefore, very definite limits to the height to which price may be lifted by agreement or convention. Andrews is almost certainly correct when he asserts that long before the highest price which could be gained from exploitation of individual consumer attachments and immobility is reached firms collectively will be threatened by an inrush of new competition. In other words, above a certain price — "full cost" in some sense — the demand curve for firms is likely to become very elastic to further price increases. "In the long run, then, demand is very sensitive to differences in prices, and even a well-established market will give no protection against the competition of those who are able to quote a lower price for the same quality product with the same level of associated services. The business men thinks in terms of a right level of price, to pass which would mean that he would have the ground

7. See Chapter VII below.
cut from under his feet and lose his market through the encroachment of other businesses ... \textsuperscript{8}

We may conclude, then, that the normal state of affairs for firms in the situation described by the curve of Chapter III will be one of neutral equilibrium sustained by tacit agreement or convention. However, it should perhaps be pointed out that this collusion does not necessarily imply monopoly exploitation. In general, it should be regarded as in the nature of a non-aggression pact rather than a monopoly alliance. It serves to put a floor under price, to prevent "cutthroat competition". This is, as I understand it, an important element in Andrews' position. But it would, I feel, be equally wrong to go to the opposite extreme, as Andrews tends to do, and to conclude that price will be essentially a competitive one. Clearly, much will depend upon the vulnerability of the market to entry by "outsiders" (which, incidentally, may be existing firms, e.g. foreign competitors, as well as potential entrants). It is hard to agree with Mr. Andrews' inference that entry will usually be sufficiently free to ensure a competitive result. In fact, a good deal of the energies of "insiders" is devoted to erecting

barriers against the encroachment of "outsiders". It is difficult to believe that these efforts are altogether ineffectual.

3. The Classification of Market Relationships

A few brief comments may now be made on the overall subject of market classification.

Firstly, the argument of this Part has provided us with alternative definitions of "pure monopoly" which are free from the objections discussed in Chapter I. Thus, we may define as a "pure monopolist" a firm with a smooth demand curve of finite elasticity and not subject to "oligopolistic interdependence" in respect of any other firm. Oligopolistic interdependence may be ruled out in terms of small cross-elasticities of demand with all other firms. (Although the existence of high cross-elasticities is not a sufficient condition to establish interdependence, the absence of them is a sufficient condition to rule interdependence out.) More precisely, we may rule out interdependence by the condition that cross-elasticity should be small in relation to own-elasticity which, for the firm, must be greater than unity. This is essentially Bishop's position in the matter of providing a numbers criterion, although this
way of stating it avoids some of the implications which are suggested in his formulation. In this case, it is not so much the number of firms between whom the particular relationship exists which is important, but rather the fact that there is a great "distance* between any two of them.

As an alternative, we may take our definition back one stage further and define a "pure monopolist" in terms of the substitution characteristics of his product. Then, pure monopoly exists where substitution between the own-product and other products is continuous in all directions, as in the case with the \( bb' \) curves of Chapter III.

It should be noticed that such definitions of "pure monopoly* are ideal ones in the sense that the definition of "pure competition* is an ideal one. In the real world, firms will only approach the one or the other. However, firms may approach our definition of "pure monopoly* as close as it likes without involving us in the logical difficulties noticed in connection with traditional definitions of pure monopoly, especially the Chamberlin-Triffin conception of pure monopoly as a limit in which the elasticity of demand for the "pure monopolist* approaches unity.
Once we are able to formulate a satisfactory definition of "pure monopoly" we are able to provide the other bound to the middle ground of market relationships variously described as imperfect or monopolistic competition. We are able to answer Mrs. Robinson's question: "But what is the limiting case at the other end? ... We know what we mean by 'selling in a perfect market', but what is a perfectly imperfect market?" Once having answered this question, we are then able to yield to Mrs. Robinson's temptation "to arrange actual cases in a series of which pure monopoly would be the limit at one end and pure competition at the other." However, it should be pointed out again here that the transition from the one limit to the other is not simply a linear movement along one scale. The transition will take place both along the scale of numbers and the scale of substitutability; and, in practice, there will be an infinite number of combinations possible between the two variables.

Finally, what can we say about the "middle ground" of imperfect or monopolistic competition. Very little! But one thing does seem obvious: the simple classification of this large field into "large group" cases and "small group" cases is quite inadequate and
highly misleading. In particular, the idea that the cases may be clearly separated and that distinctly different pricing behaviour may be expected in either camp is almost certainly wrong; and this in turn represents an almost insuperable barrier against the bringing together of price theory and business behaviour into a coherent and satisfactory relationship. What is clearly needed is the development of a comparative anatomy of market forms by which significant likenesses and significant differences may be distinguished.

In her *Economics of Imperfect Competition*, Mrs. Robinson told us that "no sooner had Mr. Sraffa released the analysis of monopoly from its uncomfortable pen in a chapter in the middle of the book than it immediately swallowed up the competitive analysis without the smallest effort." But, if the argument of this Part has been correct, it would appear that in the theory of imperfect (and monopolistic) competition the only thing "swallowed up" by traditional monopoly analysis is something very much like traditional monopoly.

For the rest, a new start — freed from the preconceptions of imperfect and monopolistic competition — is demand if we are to make significant headway with the problem of price determination within the "middle ground" which, after all, represents the general case.
PART II

THE MEASUREMENT OF AND CHANGES IN THE DEGREE OF MONOPOLY POWER
VI. MONOPOLY AND ITS MEASUREMENT

Ever since the capitulation of value theory to the general monopolistic assumptions sponsored by Mrs. Robinson and Professor Chamberlin, there have been intermittent attempts to devise an analytical measure of the degree of monopoly power. ¹ If one believes the degree of monopoly to be a significant economic variable, either in the short- or long-period, then the importance of possessing at least a conceptual measure of it can hardly be over-estimated. As Machlup puts it: "It is easier to justify the efforts devoted to the question of measurement than it would be to justify a failure to attempt an answer. Sheer intellectual curiosity compels us as economic theorists to work on this problem, for we could not with good conscience go on talking about "great" or "little" monopoly power, or about various degrees of monopoly, without trying to ascertain the meaning of these words. And this implies at least the possibility of 'conceivable' measurement,

¹. For a comprehensive survey of the various measures which have been proposed from time to time, see F. Machlup, The Political Economy of Monopoly, Baltimore, 1952, Ch. 12.
even if practical measurement were to remain impossible."²

Preferably, of course, the measure selected should be susceptible to empirical application. It may well be better to accept an inferior measure which is capable of empirical application than a more refined measure which is not. Thus, for example, even if the well-known Lerner index \( \frac{p - m}{p} \), where \( p \) is price and \( m \) is marginal cost, were in other respects adequate, it would suffer a great deal because of its restricted empirical possibilities, data on marginal cost being notoriously difficult to obtain.

All analytical measures of monopoly power are based, either directly or indirectly, upon some conception of a gap between price and "cost". Hence, the problem of measuring monopoly power splits into two subsidiary questions:

(1) Which is the relevant price and how is it determined?

(2) What is the appropriate "cost"?

The various analytical measures of monopoly power so far proposed differ essentially in that they are based on different answers to these questions, particularly the

² Machlup, op. cit., p. 469.
latter.

For example, Marx's measure of surplus value, which is a first cousin of Kalecki's percentage gross margin measure and must be looked upon as a prototype of modern measures of monopoly power, takes wages (variable capital) plus materials and depreciation (constant capital) as the only real or social cost, so that surplus value is measured by the sum of all net factor incomes other than wages, expressed as a percentage of the wage-bill. Among the moderns, Lerner assumes price to be determined by the equation of marginal revenue with marginal cost and takes the difference between price and marginal cost, expressed as a percentage of price, as his measure of monopoly power. For him, then, marginal cost is the relevant cost. Kalecki, starting with the Lerner index and upon the further assumption that marginal cost is normally equal to average prime cost (wages + raw materials), arrives at the percentage gross margin as his measure of the degree of monopoly. It is interesting to notice that this definition of what constitutes monopoly is even more sweeping than Marx's conception of surplus value because, whereas Marx would allow some salaries (as payments to special forms of labour) and depreciation as legitimate social costs,
Kalecki's measure describes these as monopoly elements.

A necessary prerequisite to the measurement of the degree of monopoly is, of course, a firm conception of what we are measuring - monopoly. The forms of monopoly are infinitely various, so much so that it has been held impossible to reduce them to a common denominator, a necessary condition for measurement. For example, Sweezy has stated: "No reasonably general laws of monopoly price have been discovered because none exist" and "We must not... expect to be able to reduce the theory of monopoly price to quantitative precision; anyone attempting to do so will only succeed in getting lost in a maze of special cases."³

Is there a common factor resident in the great variety of situations which we commonly call monopolistic? That is the question. Can we discern in cartels and cornered supplies, in toll-gates and trade unions, a common element? If we cannot, we must abandon the possibility of measurement.

It seems clear, when the problem is so stated, that there is one factor common to these situations: they

involve some barrier to the "free" flow of economic resources. Wherein does the power of a trade union lie? It rests firstly in the control — of greater or less degree — which it is able to exercise over the training of the necessary skills and the enforcement of entrance qualifications, and secondly in the sanctions which it can bring to bear against non-union labour offering for employment within the particular industry. The essence of what we call "pure competition" as a market situation is that the unit-firm is freely reproducible, i.e. new firms may enter the industry without incurring costs additional to those which have to be met by established units. Contrawise, the essence of "monopolistic competition" is that the "product" of a seller cannot — at least so far as the customers of that seller are concerned — be duplicated. As Chamberlin has put it:

"Mr. Kaldor has rightly pointed out that the assumption that 'entrance to the field in general and to every portion of it in particular was unimpeded' implies that 'every producer could, if he wanted to, produce commodities completely identical to those of any other producer ...' Logically, this is what 'free entry' in its fullest sense must mean, and it is quite incompatible with a differentiated product. With respect to the particular product produced by any individual firm under monopolistic competition, there can be no 'freedom of entry' whatever. No one else can produce a product identical with it, altho he may be able to produce others which are fairly good substitutes for it."4

Since, ultimately, the ability to produce substitutes is universal—in the last analysis as an alternative outlet for income—it follows that "freedom of entry" is a matter of degree which, if we can measure it, immediately suggests a measure of the degree of monopoly power in terms of our conception of monopoly as a barrier to the "free" flow of economic resources.

Let us take as our starting point the notion of perfect reproduceability, in the economic sense, of a good or service. This implies both (a) that the good or service can command the same price as those currently on the market, i.e. it can be made homogeneous with them; and (b) that it can be produced at the same cost. By contrast, it is the essence of imperfect ability to reproduce a commodity or service that either (a) additional costs are involved, such as those associated with advertising, sales promotion and quality competition; or (b) a lower price must be accepted, insofar as it is necessary to induce substitution by price-competition. Hence, the departure of a given situation from one of "free entry" or "perfect reproduceability" may be expressed in terms either of a cost or a price disadvantage, or a combination of both, attaching to a newcomer or potential newcomer as compared with the
entrenched "monopolist". In this way we may reduce the infinite variety of the forms of monopolistic impediment to the free flow of economic resources to the common denominator of cost and/or price advantage.

A measure of the degree of monopoly power along these lines was proposed by Mr. Kaldor a few years ago. Kaldor's suggested measure proceeds from the premise that "in a purely competitive industry, where entry is completely free, the maximum price which 'insiders' can charge without attracting 'outsiders', and the minimum price which they require in order to continue to supply the same market, are approximately equal to each other." Hence, he suggests, the divergence from equality of these two prices provides a measure of the degree of monopoly power enjoyed by an industry. Thus, "if we denote the price which just fails to cover the costs of production (including, of course, a normal rate of profit under 'cost') of potential new entrants by \( P \) and the costs of production (inclusive of normal profit) of the representative firm by \( C \), then \( P - C \) is the amount by which the selling price of the representative firm can exceed its own costs, and \( \frac{P - C}{P} \) is the measure of the

It is this approach to the measurement of the degree of monopoly which we will use as a starting point for our investigation of the measurement of monopoly power which is undertaken in this Part. However, even at first sight, Kaldor's proposed formula is bristling with difficulties: the notion of "full cost" as applied to an established firm and the concept of the "representative firm", to mention only two. These difficulties will be the subject of detailed examination in the next two chapters, theoretically in Chapter VII and statistically in Chapter VIII. Nevertheless, adoption of Kaldor's formula as a point of departure does enable us to clear up certain preliminary problems at this stage.

As has been mentioned above, the first question to be determined is what is the relevant price to take in a price-minus-cost measure of monopoly power. Since we have decided that our measure is to be designed to reflect the degree to which impediments to the free flow of economic resources are present, it is clear that we must allow an appropriate time-interval in which the adjustments demanded by a particular situation may take place. It follows, then, that the relevant price will be the long-
period price. Perhaps we should further justify this decision, since the two best-known measures of the degree of monopoly have been related to the short-period. Lerner proposed that his measure should apply to "the very shortest period", while Kalecki's measure, derived from the Lerner formula, is also based upon the short-period variables.

There does appear to be a good case for making an index of monopoly power independent of temporary, short-period changes in demand and/or cost conditions, and certainly this procedure would be in the spirit of most economic thought to date. At first sight, it may be thought that the Lerner index would not reflect such short-period variations, because it depends on the ratio of price to marginal cost and, hence, would only change with changes in the elasticity of demand. However, it is almost certain that elasticity of demand is subject to short-period variation for all sorts of reasons. Moreover, it would, in particular, be hard to sustain the proposition that short-period elasticity of demand is independent of the level of short-period price. Further, it is worth noticing that the degree of monopoly calculated on the basis of short-period price and marginal cost will bear no clear relation to the degree
of monopoly based upon similar criteria, but calculated in respect of the long-period. This point has been made by Machlup who says: "It would not disturb me to find two different degrees of monopoly for a particular firm, one referring to short-run considerations, the other to the long run."  

With Kalecki the difficulty is more apparent, since his measure is based on the assumption of constant marginal costs or, in other words, perfectly elastic supply. However, he is forced to recognize that there may be some short-period situations in which output may be straining against capacity so that price rises to ration sales. The consequent increase in the percentage gross margin Kalecki attributes to "bottleneck factors". Presumably, if a temporary reduction in the percentage gross margin were forced, e.g. by short-run dumping by

6. F. Machlup, op. cit., p. 512. Cf. also: "Long-run marginal cost included new investment outlays while short-run marginal cost does not include any part of investment cost. But new investment outlays are made only when the cost of increased output with the given capacity would be too high, that is, when long-run marginal cost - the additional cost of producing additional output with increased productive capacity - is less than short-run marginal cost. It follows that a rule to use only short-run marginal cost for calculations of the degree of monopoly may sometimes - namely, when capacity is utilized almost to the limit - result in a lower monopoly index than if long-run marginal cost were used." (p. 512n.)
"outsiders", this would have to be treated as a sort of negative "bottleneck factor" deductible from the degree of monopoly. Theoretically at least, the degree of monopoly as defined by Kalecki would be susceptible to quite sharp short-period fluctuations, particularly so far as the individual firm or industry is concerned.

A measure of the degree of monopoly referred to a long-period basis is free of these objections and therefore has much to commend it. The long-period itself will be defined as the period necessary, in virtue of technical or physical conditions of production, for resources to flow into a given industry. The appropriate price then for our monopoly index will be the long-period price; and the precise manner in which this long-period supply price is determined will be the subject of detailed examination in Section 1 of the next chapter.

Meanwhile, it is necessary to give some preliminary attention to the cost element of Kaldor's measure of monopoly power which we have adopted as our starting point. The relevant cost, Kaldor suggests, is the "full cost" of the "representative firm". As already indicated, both these concepts are very slippery and will require detailed investigation. This will be the task of Section 2 of the following chapter. Here, rather, we are
concerned with the broader question: what elements should we allow to enter into the calculation of "cost"?

There are two problems here concerned with the treatment of (1) selling costs; and (2) rental elements.

(1) Selling Costs

On this subject, Kaldor has put the position succinctly:

"It would be wrong to suppose, however, that the difference \( p - c \) is normally retained by the firms in the form of profit. In an industry that is not a pure monopoly but consists of a number of separate concerns, each of which is striving to obtain a growing share of the market, a considerable part, if not the whole, of the difference will tend to get taken up by the expenses incurred in order to enlarge the size of the market, the expenses consequent on 'non-price-competition', commonly referred to by economists as 'selling costs'. The distinguishing mark of 'selling costs' is that they arise in consequence of the fact that the price is higher than the purely competitive price, and their magnitude will be proportionate to this difference, i.e. it will depend on the magnitude of the obstacles facing outsiders. Selling costs exist with all kinds of market organization (except with the perfect market which dispenses completely with goodwill). They arise, therefore, with 'wholesalers' domination' just as much as with 'manufacturers' domination'. But it is only in the latter case that they become quantitatively important as a proportion of final price. A further complication arises from the fact that the difference \( p - c \) cannot be taken as given, irrespective of whether selling costs are incurred or not, or of how they are incurred. In so far as the services provided in conjunction with these selling costs are such that they increase the manufacturer's goodwill and attach his customers more firmly to himself — which they mostly are — the level of \( p \) is
raised. Hence the incurring of selling costs may be regarded as partly 'competitive' (i.e. aiming to enlarge the firm's share in the market) and partly 'protective' (i.e. aiming to increase the firm's monopoly power), though these two kinds of effects may not always be clearly distinguishable from each other. 7

Thus, selling costs are both a cause and a consequence of monopoly power. Insofar as they are a consequence, no problem is involved, since the degree of monopoly will stem from some other source, e.g. economies of large-scale, which may be determined independently. However, 'protective' selling costs are creative of monopoly power and must enter into any measurement of it as a separate factor. As Kaldor has pointed out, it is difficult to identify any particular selling expenditure as belonging to the one category or the other. However, one way out of the difficulty which suggests itself is to take as an index of the protective effect of selling outlay the degree to which an entrant firm would have to match the selling costs of insiders in order to obtain an economic foothold in the market. Then, entry supply price, \( P \), will include such necessary outlays; while \( G \), the full production cost of established firms, will, of course, be exclusive of

the selling costs of those firms. In this way, the selling outlay mandatory for a potential entrant will enter into the determination of the magnitude of $P - C$. The precise relationship between the selling outlays required of potential entrants and the degree of monopoly shelter enjoyed by established firms in virtue of them will be examined in the following chapter.

Meanwhile, there remains, on the subject of selling costs, only one other problem requiring preliminary notice, viz. the problem of distinguishing "selling costs" from "production costs". In the real world the distinction becomes blurred at the margin; and we require some theoretical criterion by which the distinction may be drawn. Kaldor treats the issues involved at length — and, in the writer's view, adequately; and it is not proposed to re-cover the same ground here. Kaldor concludes that selling costs are "the excess of the total expenditures actually incurred, at all stages of the chain of production and distribution, over the amount that would have been incurred, if all separate services performed in the course of the productive and distributive process had been priced separately."  

Although this definition is derived from the criterion

of perfect competition, it does not involve the description of all expenditures other than those which would occur in the very specialized form of market organization - a form which could not conceivably apply to the marketing of many goods and services of the real world, however little the monopoly power present - which is the subject of the perfectly competitive model as "selling costs". Even a socialist society would undoubtedly - and the Soviet Union does - undertake, for example, a minimum of advertising and packaging of a wide range of goods and services. Thus, a cinema, in any society, would be bound to advertise to some extent the programme on offer from time to time. Classified advertising which is largely informative in character represents nearly 40 per cent of total press advertising even in our economy. ⁹

Again, cigarettes would presumably need protection from damp and heat, such as is provided by cellophane, just as much under socialism as capitalism. It may even be thought worthwhile making the packaging attractive. In short, a socialist society would incur expenditures which would not be found in the theoretical "perfect market". It would be difficult to justify similar expenditures being described as monopolistic merely

---

because they are made under capitalism. It is the advantage of the Kaldor criterion that these types of service, which could command a price if offered separately, do not fall into the category of "selling costs" and hence do not rank as monopoly elements.

(2) Rental Elements

Our problem in relation to rents is somewhat analogous with that of selling costs; and arises out of the fact that any monopolistic advantage may be capitalized and appear as a money cost of production (as rent) either to the established firm or potential entrant. An established firm may be making barely normal money profits, yet there may be a considerable degree of monopoly resident in the situation. A good example of this is provided by Australian hotels, where the capital cost of purchase of a licence or its rental is so high that the purchaser or lessee may make very modest profits on the capital he must invest. (Incidentally, this is a principal reason why so little of hotel profits is ploughed back into improved facilities and services in the case of individually owned or leased premises. The brewery-tied hotels do much better in this respect.) Other examples of the capitalization of monopoly privilege come readily to mind.
Thus, a firm owning a patent may make large monopoly profits, whereas a firm which is otherwise similarly situated but which has to licence (rent) a patent may make very small profits. Again, the purchase of goodwill - in large part the capitalization of past selling expenditures - may be an alternative to current selling outlay. It is clear that rents paid in respect of such goodwill must be accorded the same treatment as current selling costs, from the point of view of calculation of the monopoly surplus in a given situation. The rent of urban land sites is another example. A new firm entering the retail trade may well have a direct choice between paying a high rent for a particularly favoured site or spending heavily on advertising to attract customers to a less favoured location. Clearly, the one form of expenditure must be placed in the same category as the other.

Thus, we find that rents arise directly out of many situations which have almost always been regarded as monopolistic. The problem then arises of where we are to draw the line between rents which may be looked upon as a monopoly surplus and those which may not. One way out would, of course, be to reckon all rents as monopoly elements in the manner of Kalecki. However, this solution must be rejected for the following reasons:

(1) Although there is every precedent for the treatment
of rent as an "unearned surplus", to describe this surplus as "monopolistic" would be to stretch the meaning of the word "monopoly" far from its traditional and usual sense. Substantial monopoly surpluses would then become compatible with market situations approaching the conditions of perfect competition.

(2) Rent is an expenditure alternative to capital outlay. For example, one may pay a high rent for land with a natural water supply or, alternatively, pay a lower rent and expend capital on providing such a supply. Clearly, if we count all rents as monopoly elements, we would also be bound to include quasi-rents accruing to capital. The result would be a portmanteau conception of monopoly which would swallow up many distinctions which it is valuable to make. For example, rents arising out of the productive advantages of natural resources are quite compatible with the best allocation of resources from the point of view of production—though not of course necessarily from the point of view of distribution. By contrast, the presence of monopoly implies some distortion from the optimum allocation of resources.

Therefore, we are left with the problem of distinguishing rents arising from monopoly situations from those which do not. A first approach to an answer
to this problem may be made by invoking our distinction drawn in the previous section between production and selling costs. Using the same criterion, we may then distinguish between rents arising out of a selling advantage and those arising out of a production advantage. Then, all rents in the former category may be isolated as monopoly elements. This class would include, for example, rents payable in respect of capitalized goodwill, restrictive selling licences (as with hotels), and urban selling sites (the excess of the rent over the opportunity rent earnable in production).

However, rents deriving from monopolistic selling advantages do not exhaust those rents which may be legitimately described as monopolistic. Two further types, falling in the category of production rents, may be identified: (1) those deriving from patents and the like; and (2) those deriving from the exclusive or near-exclusive possession of the supply of some natural raw material, e.g. in Australia B.H.P.'s control over the rich iron ore deposits of Iron Knob and Iron Monarch; or in Africa the diamond cartel. Both involve a barrier to the entry of new producers, the one due to law, the other to nature. The treatment of rents earned by patents is simple enough; they are measured by the excess revenue earnable by the patented process over that earnable if unpatented, freely
available, techniques were used. In the case of exclusive access to natural raw materials, we are, however, faced with the subsidiary problem of distinguishing the monopoly rents implied in such a situation from rents ordinarily earnable from any other "free gift of nature," such as a fertile tract of land. This distinction is best drawn in terms of the transfer-earnings of the asset. In a competitive regime, the rent payable for "land" must be sufficient to keep it employed in the particular line of production. It will therefore be just greater than the rent it could earn in an alternative line of production. On the other hand, monopoly elements are present when "land" is not transferable from one line of production to another. New "land" cannot come into the particular line of production; there is a large gap between the earning capacity of this highly specialized type of "land" and its earning capacity in an alternative line of production. For example, Iron Knob and Iron Monarch and its environs would certainly be extra-marginal if it were not for the iron ore deposits there. Consequently, we define the excess of rents of specialized natural resources over their transfer-earnings as a monopoly surplus. In effect, then, we class all rents other than those which would occur in a regime of perfect competition as
Having settled these preliminary questions, we are now ready to begin the task of developing a measure of the degree of monopoly power, upon the basis of the general principles outlined above.
CHAPTER VII

THE DEGREE OF MONOPOLY POWER

1. Entry Supply Price

In this chapter it is proposed to investigate in detail the measure of monopoly power suggested by Kaldor, taking into account the general considerations discussed in Chapter VI. The more straightforward of the elements of Kaldor's measure is \( P \), the price required to attract new entrants into an industry. To give precision to the concept it is necessary to consider the conditions which will determine whether entry is attractive or not and, if it is, upon what scale. Let us first prepare a schedule relating full production cost of the potential entrant to the various volumes of output which he is likely to consider. Full production cost will include "normal profit", the price necessary to attract free capital into the industry, which will depend on the ruling long-term rate of interest and the degree of risk attaching to the particular line of production. It will also include all

---

1. The basic argument of this chapter was published under the title: "The Degree of Monopoly Power", Economic Record, May 1952, pp. 1-12. However, some important modifications are introduced into the version given here.
"production rentals" which the newcomer has to meet. (The outlay involved in such rentals is, of course, merely an alternative to the expenditure of capital or labour to achieve a given productive result, e.g. in the simplest case, a factory may be rented or built.) Full production cost will, however, exclude selling costs and "selling rentals" in virtue of the argument of the previous chapter. The full production cost schedule so prepared will, then, define a cost curve which we reproduce as curve CC' of Figure 6.

Let us now consider the demand side. Where monopoly elements are present, we cannot simply draw a demand curve for the entrant firm, because this will not be independent of the selling outlay undertaken by the new firm nor of the commitments it is prepared to undertake in the form of "selling rentals", i.e. in the purchase of existing monopolistic selling advantages. Since our treatment of these latter rents will be the same as for current selling expenditure, we may proceed on the understanding that the portmanteau term "selling costs" is used to include these rents. Then, in the general case of imperfect markets, the entrant firm will have the choice, between limits, of attracting an economic volume of sales either by price-competition with existing firms or by resort to selling outlays or a combination of both. Theoretically, for each such output, \( q_x \), there will be a whole range of varying
combinations of price- and non-price-competition which a potential entrant might consider, although in practice the choice is likely to be much more restricted.

At all events, for each output there will be a best combination of these elements - best in the sense of yielding the greatest net revenue per unit of sale. Let us denote this maximum net revenue obtainable in respect of any output, $q_x$, by $R_x$. Then to each such $R_x$ there will be a corresponding best price $P_x$. If we plot these values against the outputs concerned, we obtain a net revenue curve $RR'$ and a price curve $PP'$ as in Figure 6.

![Figure 6](image)

2. The price which will be relevant is, of course, the final price the entrant may expect to be able to get in the market, after allowance has been made for the fact that existing sellers may greet his entry by lowering their prices, that is to say, it must take full account of oligopolistic reactions to his own decision to enter.
From Figure 6 it is easy to see that the price which "insiders" may charge without attracting "outsiders" is the price just below that given by the point of tangency between the curves RR' and CC'. For, if RR' lies wholly below CC', there will be no point of entry which is profitable and, consequently, "insiders" can charge a higher price. On the other hand, if RR' lies above CC' at any point, then at that point entry is attractive and the influx of newcomers may be expected to force down RR' to the point of tangency.

The difference $P_x - R_x$ in respect of any output $x$ at which the firm may enter we shall call, following Kaldor, the cost of entry in respect of that output.

The curves we have drawn in Figure 6 are those which may be expected in a substantially imperfect or monopolistic market situation in which the revenue curve RR' may be presumed to fall fairly rapidly as the new firm's target output is increased, this for three reasons: (1) the need to invoke greater price-competition to break existing attachments of a larger number of buyers; (2) the larger the projected incursion into the markets of established sellers, the more likely is retaliatory price-cutting by those sellers, thus requiring further reduction in the price which the entrant may expect to receive; and (3) sooner or later, diminishing returns to selling outlay are likely to set in. It is impossible to say a priori
whether the price or the revenue curve will fall the more rapidly. On the one hand, the possibility of retaliatory price-cutting will tend to make non-price competition more attractive; on the other, diminishing returns to selling outlays will tend to make it less attractive. Whether price- or non-price competition becomes the more or less important as the scale of the new firm's target output increases will clearly depend upon how these opposing tendencies balance out.

As a market situation approaches that of pure competition, the curves PP' and RR' will move towards coincidence and their slopes will approach the horizontal. This means, of course, that in such a situation a small price differential is all that is required to attract any volume of sales desired. The second condition of the pure competition model may also be noticed. For pure competition to obtain, individual firm outputs must only represent a small fraction of total market supply. In other words, economies of scale must be negligible compared with the size of the market which is merely another way of saying that operation at a point of minimum cost is consistent with a large number of firms. In this case, then, the curve CC' will be U-shaped over the relevant range. Entry supply price will be given by the point of tangency between the horizontal PP' (= RR') and the U-shaped CC' as in Figure 7. This is, of course, the normal long-
period equilibrium position of the special case of pure competition. In short, then, the analysis advanced here is able to embrace the limiting case of pure competition, in correspondence with our conception of monopoly discussed in Chapter VI.

There are two further matters on the subject of entry supply price which demand consideration. Firstly, the analysis given above, as it stands, is far too static. In the real world, a firm considering entry into an industry is not faced with a situation in which it may or may not secure an economic share of market output at any particular point in time. Rather it is a question of whether such a firm can hope to build up to such an economic output over a period of time. This period may
be considerable, even years; and during such time a new firm may have to be prepared to incur losses. Such losses represent, of course, a cost of entry—and often a formidable one. We must, therefore, take account of this factor in our analysis.

To do this we must take a look at a potential entrant's position, not as at a point in time, but over a period. This period must be that over which a firm is able to foresee, with a reasonable degree of probability, future trends—trends of costs and of demand principally. Let us divide this period into smaller intervals—say the accounting year, for example. Then in respect of each such "year" we will have a corresponding pair of curves $CC'_t$ and $RR'_t$. In general the position of these curves will change through time. If at no point throughout the period under consideration does the curve $RR'_t$ stand above $CC'_t$, then at no time will entry be profitable. However, it will often be the case that, while $RR'_t$ will lie below $CC'_t$ over the early years of establishment, it will rise above $CC'_t$ in respect of later years as the new firm's market is established and consolidated. Then, for any given planned or possible rate of growth of the firm's market, specific losses will be incurred in the earlier years. In respect of each such year this loss will be equal to $q_x(C_x - R_x)$ where $q_x$ is the output for the year.
Clearly, this loss represents a cost of establishment which must be a charge against subsequent years. Within the framework of our analysis this situation is best dealt with by successively deducting from the revenue of each year the cost of liquidating losses anticipated in respect of the immediately preceding year, such losses, of course, having been charged with the appropriate rate of interest. It will then remain a condition of entry that in some year the corrected RR' stands above CC' in respect of some output. The successive reduction of RR' in relation to PP' specifically allows anticipated establishment losses to enter into the final "cost of entry" as conceived above. The practical importance of this change from a purely static model is that it brings into view at once that, other things being equal, the cost of entry into an industry will be lower in an expanding industry than in one that is stationary or declining.

The second matter which requires consideration to complete this part of our analysis is the question of defining the "industry". So far we have taken this for granted and talked of the "industry" and of "insiders" and "outsiders" without considering the well-known difficulties of these concepts. In many cases, of course, the "industry" will be fairly easy to define in practice; in others it will not. However, in this context we need not worry too much about the exact bounding of an "industry"
as such. Here we are concerned with entry, and the main significance of the industry will be: (a) from the point of view of the potential entrant, the most favourable place to enter; and (b) from the point of view of the existing firm, the most vulnerable spot for encroachment by new firms. Perhaps this may best be made clear by an illustration. Suppose we have a situation of pure chain-relationship. Clearly, it will be most profitable for a new firm to enter at the "weakest" link in the chain. Similarly, the "insider" cannot be content merely to estimate the danger of entry in his sector. He must also look at the weakest link, because entry here will produce repercussions which are likely to be passed along the chain to other firms in it.

Therefore, an established firm will be concerned with entry from two main points of view: (1) identification of the most favourable point of entry; and (2) calculation of the repercussions on itself both directly and by transmission through the price-output policies of associated firms. Formally this problem could be set out in terms of impact coefficients something along the lines of cross-elasticities. (Cross-elasticities, however, are not themselves suitable for this because they are not adapted to measuring cross-relationship with "potential firms" whose outputs and price are not given.)
In practice, however, the industry is likely to be a business-political entity representing a grouping of firms with a common interest which may stem from other than purely economic factors. This grouping may be international, national or local, and it may change from time to time although the actual firms may remain individually the same. We have our international cartels, of course. Contrawise, we get local manufacturers producing for a local market who will consider similar producers in the next State as "outsiders". But probably the most important demarcation is that along national lines. Indeed, we have special associations of national producers and business jargon talks of national products and nationally-advertised products. Here we have groups of national producers with the domestic producers representing the "insiders" and foreign producers "outsiders". This type of grouping gains its special significance, of course, in virtue of the fact that tariffs run on national lines.

At all events, in identifying the most dangerous threat to their market, producers will have to take account of the existence of established "foreign" firms which may or may not represent a more immediate danger than "domestic" entrants into the industry. The concept of entry therefore needs to be modified to this extent. If entry supply price as defined above is lower than the "landed" price of "foreign" goods then price will tend to settle
around the "domestic" entry supply price. If, on the other hand, "foreign" producers can enter the market at a lower price than new "domestic" entrants then clearly price will tend to be held down to this figure. It follows that the degree of monopoly obtaining in any particular industry will be very strongly influenced by transport costs and by tariff policies.

2. The "Full Cost" of an Existing Firm

We now turn to consider C, the "full cost" of production of the representative firm. For the time being, we put aside the problem of defining the "representative firm" and concern ourselves with the full production cost of any existing firm.

No difficulty arises in the case of current production costs, which we shall designate by c, that is with those costs for which the firm may re-contract and which it has to meet so long as production continues. These costs will include prime costs (wages and raw materials), current overheads (salaries, printing and stationery, telephones and postages, etc.), normal profit on working capital (since this can be varied in response to both output and value changes) and prime depreciation. We assume, as is usually the case, that prime depreciation - wear and tear - is continuously made good by maintenance expenditures which
are determined by current values. Obsolescence, of course, is in an altogether different category.

It is when we come to consider those costs of production which are fixed, either because they are embodied in past contracts (e.g. leaseholds) or represent expenditures which are "sunk and gone" into fixed capital equipment, that our real difficulty begins. Clearly, the purchase price (the accounting value) is irrelevant here. Obviously something in the nature of reproduction value is required. But reproduction value is a loose concept when applied to investment in capital equipment, for there may be no present market for the actual physical units being used in production. Actual equipment will differ from currently purchasable equipment in that it will, in general, have suffered obsolescence (or wear and tear not made good) of greater or lesser degree. The value of the contribution of the existing equipment can, therefore, only be established in terms of its efficiency in producing the given product as compared with the efficiency of currently purchasable equipment, that is the most up-to-date technique available to a newcomer. This value we shall call product reproduction value.

Also, there is the question of excess capacity to be taken into account. In general, some excess capacity will exist. Often this general under-capacity working in manufacture is attributed to technological factors.
But this is not the only reason nor, necessarily, the most important reason. No entrepreneur building a new plant, or enlarging an old one, would dream of designing its capacity so as to "just fit" his anticipated immediate output. Some slack will always be provided. Apart from the diseconomies of subsequent ad hoc extensions and additions, every entrepreneur will want some excess capacity with which to meet sudden increases in demand, since it is much easier to form the buying habits of new customers than to break these habits once formed. At all events, irrespective of whether excess capacity is of the nature of a margin for expansion or of a technological character, the same considerations will operate for the existing firm and for the potential entrant.

But where excess capacity exists in excess of the "normal" the case is different, since this clearly represents an over-investment of capital in relation to current market conditions. The newcomer, on the other hand, can choose a scale of plant more appropriate to these conditions. Therefore, it is clear that capital sunk in excess equipment of this category cannot enter into the product reproduction value of the existing firm's capital.

In terms of accounting practice, our argument may be summed up as follows. In so far as the supply price of equipment has altered (in terms of money) this represents
a capital loss or gain which has actually been incurred. In so far as the equipment has suffered "normal" obsolescence or depreciation this will have been a charge against past production, the value of the equipment having been appropriately written down. Where unforeseen obsolescence (e.g. due to innovation) occurs, this also represents a capital loss which has in fact been incurred and cannot be a charge against future production. Exactly the same may be said in respect of idle capacity in excess of the "normal". Although accountants do not actually treat these phenomena in the way described, it is interesting to notice how, in practice, they do modify their formal approach to capital values in the face of large changes in the product reproduction value of the capital invested in fixed assets. Often, in the face of substantial changes in money values, assets will be revalued. Again, it is quite common accounting practice to provide large amortization charges in times of high income and a negligible charge in times of low earning capacity. Moreover, in the face of large and more or less permanent contractions in income-earning capacity, it is common practice for firms to "reconstruct" or "reorganize" their capital structures (in terms of book values) to bring these values more into keeping with the earning capacity of the capital concerned.
To give precision to the concept of product reproduction value, let us call the capital cost of the most up-to-date technique available to a newcomer at the output corresponding to entry supply price, $k_e$. (Capital cost is defined as the estimated amortization charge plus normal profit on capital invested.) For the moment, we abstract from economies of scale or, in other words, we assume an identical output for the existing firm and for the new entrant. Then the full production cost of the new producer is defined as $C_e = k_e + c_e$, where $c_e$ is the current cost as defined above of the new firm. Clearly, then, $C_e$ is the reproduction value of the product and $k_e$ is the contribution to this value of the capital employed. If $c'$ is the current cost per unit of product of an existing firm, then $c' + k' = C$ is the full cost of production of that firm. But since the value of the product is defined as $C_e$ then $C' = C_e$ and, hence, $k' = C_e - c'$; and $k'$, related to the expected life of the equipment, gives the product reproduction value of the existing firm's capital equipment. This is so because the lower efficiency of existing equipment as compared with that available to the newcomer can only manifest itself in a rising $c'$ per unit of product. Thus, we define the full production cost of an existing firm as being equal to the full cost of a new producer using the best available technique and producing an identical output. If outputs
are in fact identical — or, what amounts to the same thing, if economies of scale are absent — then clearly \( P - C \) will be identical with \( P - C_e \) which is equal to \( E \), the cost of entry as defined in Section 1. This, then, constitutes the first element of our definition of the degree of monopoly power.

We must now consider the more general case where the output of the existing firm differs from that of the newcomer. It is easy to see from the previous discussion that, if \( C_a = k_a + c_a \) is the full cost of production of an entrant firm at the actual scale of output enjoyed by the existing producer with whom we are concerned, then \( C' = C_a \) and \( k' = C_a - c' \). But, in general, \( C_a \) will not be equal to \( C_e \), for the existing firm may enjoy economies of scale not available to a newcomer. Thus, \( C_e - C_a \) measures the shelter enjoyed by the existing producer in virtue of his established scale of output. Economies of scale, which we designate by \( S \), are very important in practice and may often constitute the most formidable obstacle to competition from new producers. Thus, \( S = C_e - C_a \), measuring the economies of scale enjoyed by the established firm, is the second element to be embraced in our measure of the degree of monopoly power, and \( P - C = E + S \).

We must now turn our attention to a point which we neglected in the foregoing analysis. It will be recalled that we defined the capital cost of an existing producer
as \( k' = C_a - c' \), where \( C_a \) is the full cost of a new producer at the actual scale of output enjoyed by the existing firm and \( c' \) is the existing firm’s current cost. This is the general case; but it is important to notice two special cases. Firstly, we consider the case where \( c' > C_a \) and, consequently, \( k' \) is negative. This means that the current cost of the existing producer is greater than the full cost of a new entrant at the same output. This could easily happen in the face of rapid technological progress. Under competitive conditions this would mean that the existing firm would have to re-equip or cease production. The firm has become extra-marginal. But under monopoly conditions a firm may continue in production so long as its monopoly elements, \( E + S \), are sufficient to shelter it from the consequences of \( c' > C_a \). In other words, the firm does not become extra-marginal until \( c' - C_a (= -k') \) is greater than \( E + S \) or, what amounts to the same thing, until \( c' \) is greater than \( P \), that is when price fails to cover the current costs of production. This illustrates a facet of monopoly which is of fundamental importance, namely, that the barriers to the exit of marginal firms are exactly of the same order as the barriers to the entry of new firms. It will be clear, then, that we need make no amendment to the conclusions reached above for this case. A negative \( k' \) will, of course, make a big difference to accounting profits, but will not affect the
degree of monopoly shelter enjoyed by the firm.

A second special case must now be noticed, namely, the case where the current costs of production of an existing firm are lower than those of a newcomer at the same scale of output, i.e. where $c' < c_a$. This means that the productive efficiency of the existing producer cannot be reproduced by a newcomer. This could be the result of the possession by the existing firm of secret or patented technical processes not available to an entering firm or (and this factor is becoming increasingly important in the modern world) of the possession of what is nowadays called "know-how". This type of monopoly shelter is very important in certain sectors of a modern economy; and it must be embraced by our measure of the degree of monopoly power. This element we call $T$, the shelter enjoyed by virtue of secret or patented processes of production or secret "know-how", and clearly $T = c_a - c'$.

3. The General Measure of Monopoly Power

We are now in a position, tentatively, to define our general measure of the degree of monopoly power which for the time being we will denote by $M''$. Then, $M'' = \frac{P - C}{P}$, where $P - C$ is equal to $E + S + T$; and where:

$E = P - C_e$ is the "cost of entry";

$S = C_e - C_a$ measures economies of scale; and
\[ T = c_a - c' \] is the shelter provided by the possession of secret or patented technical processes or equipment or "know-how". Of course, in any particular situation, any of these elements may be equal to zero.

To show these relationships graphically, we reproduce the curves of Figure 6, leaving out the net revenue curve RR' for the sake of clarity. We add a "shadow" cost curve TT' for those cases in which shelter of the nature of T is enjoyed. Clearly, the ordinates of this curve will be less than those of CC' by an amount \( c_a - c' \). Then, from Figure 8 it is clear that \( M'' = \frac{P - C}{P} = \frac{PT}{PE} \) and \( E = PR, S = RS \) and \( T = ST \).
One final modification is required at this stage. When firms act unilaterally in the pursuit of profits, long-period price will tend to settle automatically around entry supply price as the result of the entry and efflux of firms. However, when there is cohesive action by the firms within an industry or when oligopolistic warfare breaks out price may deviate from entry supply price in one direction or the other. In particular, in a "tight" oligopoly situation price may be held above entry supply price even in the long-period. Price can be maintained at this higher level because entry may be held off by the threat of cohesive reduction of prices by insiders to a level which may well deter the most venturesome potential rival. In practice, strongly oligopolistic industries resort to all sorts of variants of this strategy. For example, a common occurrence is to threaten potential or actual newcomers with long and very expensive litigation over alleged patent infringements and associated legal technicalities. Again, threats of boycotts, various forms of physical intimidation and the use of political influence have been used from time to time to prevent the entry of new firms. Clearly, in so far as such tactics are successful, "insiders" may be able to hold their price above the purely economic entry supply price.

On the other hand, where oligopolistic cohesion breaks down and firms enter on protracted warfare with the end of
achieving hegemony for one or a few firms, price may be held well below entry supply price over a considerable period of time.

Therefore, to take account of these factors, we introduce a coefficient into our measure of monopoly power which may be described as a coefficient of oligopolistic cohesion. If we use \( k \) to denote this term, then our monopoly measure may be recast in the form:

\[
M' = k \cdot \frac{P - C}{P}
\]

Clearly, \( k \) may be rather greater than unity when firms are able to maintain price above entry supply price by resolute cohesive action; \( k \) will be less than unity when there is a breakdown of intra-group discipline which forces price below entry supply price.

4. Monopoly Forms and Monopoly Strategy

One of the principal advantages of the measure of monopoly power developed above is that it is directly referable to the strategy and activities of monopolies in practice. By contrast, the Lerner index of monopoly power is several times removed from the practices of monopolists as they are manifest in the real world. It is only in the most indirect way that one may link the everyday strategy of monopolists with the aim of reducing elasticity of demand and thereby altering the relationship
between marginal cost and price. It will therefore be useful to pause at this stage and briefly consider our measure of monopoly power in terms of the forms and strategy of monopoly as they appear in the real world.

**Economies of Scale and Numbers.**

Perhaps the most important single tendency towards monopoly is found in the concentration of production, either in virtue of horizontal or vertical integration. Horizontal integration has its principal raison d'etre in economies of scale. The reduction in the number of firms which results from such integration is not so much a cause of monopoly as a consequence of it. Fewness of producers is a symptom of the existence of monopoly power - resident in economies of scale or some other form of shelter from "outside" competition - not a cause in itself; for it is clear that a small number of producers could exert no monopoly power beyond the shortest period, if it were unsheltered from the competition of new producers. By and large, then, one should expect the existence of monopoly power of the sort we have denoted by S to be reflected in a reduced number of producers.

Usually, when economies of scale are discussed, attention is focussed on economies in prime production costs arising out of technological factors. However, there are other economies of large-scale operation which
should not be overlooked. These occur in the realms of marketing and finance - working capital in particular - and may also derive from the spreading of more or less indivisible overheads over a larger output. This kind of economy of scale is perhaps the most important inducement towards vertical integration. As Kaldor has pointed out, "The economies of 'vertical' expansion are largely pecuniary rather than technological in character." Two further points may be made in this connection. Firstly, because of the economies possible in the spreading of overheads, elimination of duplication of function, reduction in working capital requirements and the streamlining of marketing arrangements, we may still expect a growth in concentration, even after technological economies of scale have been exhausted, e.g. where multi-plant production exists. Secondly, it will be noticed that the element S in our measure of monopoly power embraces all economies of scale, whether they be in prime costs or overheads.

3. This is not to deny that there are sometimes technical economies of scale involved in vertical integration, e.g. in the steel industry, nor that horizontal integration often brings some economies in overheads.

The Degree of Oligopolistic Cohesion

We have argued above that monopoly power does not derive from fewness of numbers, per se, but rather that fewness is a consequence or symptom of the existence of monopoly power. This argument needs qualification in one important respect, namely that discussed in Section 3 where it was argued that fewness of numbers will tend to facilitate intra-group discipline and thereby enable producers to exploit their position to a point more closely approaching the theoretical maximum. Many of the conventions and organisational forms of business are directly explainable in these terms. Price leadership, basing point systems, and trade associations are typical arrangements for the stiffening of intra-group discipline. Likewise, pricing conventions and those aspects of the businessman's ethos which are expressed in terms of antipathy towards "cut-throat competition" or "unfair trading practice" serve a similar function.

Costs of Entry and Defensive Strategy

The two chief ways in which monopolists may raise the cost of entry to newcomers are (a) the development of goodwill and product-attachment by selling expenditure - in the broadest sense; and (b) vertical integration, or control over marketing outlets which stop short of
integration. Successful use of these strategies will be reflected in the element $E$, the costs which a new producer must meet to enter the market.

We have already discussed vertical integration insofar as it derives from economies of scale, usually of an overhead nature. However, a good deal of the drive towards vertical combination stems, not from this source, but from attempts to raise $E$, the extra costs associated with a newcomer's entry into the market. This strategy can best be illustrated by an example. In Australia, vertical integration between the breweries and hotels has advanced a long way. Under the "tied-house" system, a potential newcomer in the brewing field is faced with the very substantial obstacle of securing or establishing alternative retail outlets. This, of course, greatly enhances the degree of monopoly power enjoyed by the existing producers. A similar strategy, stopping short of actual integration, seems to be implicit in recent attempts by oil companies to introduce "one-brand" service stations.

But it would be a mistake to believe that this strategy is directed solely towards enhancing the degree of monopoly power and thereby raising price. In large part the aim is to insulate output, not so much against the newcomer, but against rival producers, particularly in respect of possible contractions in demand. This has
two aspects. Firstly, loss of output involves the firm in some loss of economies of scale and hence of monopoly profit. Secondly, the rendering of some capital equipment idle involves the firm in capital loss of the nature discussed in Section 2 above. Moreover, the importance of output is the greater the greater is the degree of monopoly. For under competitive conditions much of the loss in group output will be taken up by the disappearance of marginal firms. But we have seen above that the ease with which marginal producers may be forced out of an industry is of precisely the same order as the ease of entry of newcomers. If the degree of monopoly is substantial, then we may expect contractions in demand to be met, not by the disappearance of high cost producers, but by the spread of idle capacity throughout the industry. It is important to notice that this form of "restrictionism" is not a cause of monopoly but a consequence of it. It is a mistake, therefore, to look upon the mushroom growth of cartels and price-maintenance agreements in the slump as indicative necessarily of a rise in the degree of monopoly. Output may be far more important to a firm than price over any given practicable range of variation. Consequently, in a slump, there will be an almost irresistible temptation for the firm suffering severe setbacks in output to cut price in an attempt to "beat the gun". Price-maintenance agreements in the face of
the demand contraction of the slump should be seen, therefore, as non-aggression pacts rather than as monopoly alliances. And, as we know from contemporary history, non-aggression pacts are not indicative so much of a heightened state of peace as of a greater threat of war.

Again, in many cases oligopolists are aware of just these consequences. Therefore, during an expansion of demand they leave a fringe of demand to be satisfied by smaller firms. In the slump, the bulk of output loss is taken up by the closing down of these smaller producers. It is important to notice here, however, that tactics of this kind are only practicable when the price-leaders enjoy monopoly shelter of the nature of S and T and also where E is small. Where E is large, different tactics are necessary, since a large E would protect once-established marginal firms from the consequences of all but large price-cuts, which the price-leaders will not be anxious to make. It is in these cases that selling costs play an important role.

It is sometimes puzzling why oligopolists, who certainly combine in their dealings with labour and pursue a common price policy, should dissipate a substantial

---

proportion of their profits on competitive advertising. In many instances, it is obvious that, as far as the redistribution of consumer demand is concerned, much of such advertising cancels out in its effect. The answer is sometimes advanced that advertising outlays are not scaled down because it results in an increase in overall demand for the product of the group and, although it does little to alter the intra-industry status quo, it benefits each of the members of the group in virtue of the enhanced group demand. However, it seems certain that, in those cases where such a result would be significant (and it is hard to believe that, except perhaps in the case of new products, group demand is very responsive), the same result could be achieved at considerably less cost to the group. It would seem, therefore, that apparently competitive selling expenditures of this kind have their rationale in two more fundamental aims. Firstly, selling expenditures of existing firms will raise $E$ and thus give added protection against the encroachments of "outsiders". Secondly, they will tend to insulate the demands of individual firms and thus act as a deterrent against ruinous intra-group price competition. To revert to our earlier analogy, "wasteful" selling outlays are no more inconsistent with oligopolistic agreement than competitive armament building is inconsistent with a pact of non-aggression, and for the same reason.
Patents, Designs and Technical Improvements

Exactly the same considerations apply to the modern trend towards competitive (and often duplicative) research by oligopolists. These firms are not particularly interested in the acquisition of knowledge for its own sake. The essence of their efforts is the attempt to keep one jump ahead of competitors or potential competitors, either in the technique of production or the quality of product. Research outlays perform the same function in respect of T as selling expenditures perform in respect of E.
VIII. THE DEGREE OF MONOPOLY POWER

(Continued)

1. Statistical Application of the Measure of Monopoly Power

We have now developed a tentative analytical measure of the degree of monopoly power; and it remains in this chapter to submit it to empirical test. For this purpose, we shall use the data of the Australian Production Bulletin (Secondary Industry) for the year 1938-39. In choosing this year we have had to balance a number of considerations. The later the year we take the better, in general, is the information available from the point of view of comprehensiveness and dis-aggregation. On the other hand, the war-time years and the post-war period of inflation have had a distorting effect on the economy and, in particular, on the money values in which some of our important information is expressed, e.g. capital values. Also, price controls, of which some remnants still exist, have, in the post-war period, almost certainly caused significant departures from the course which business would have taken given a "free" market.
Again, short-period bottleneck factors have been important practically throughout the post-war period. For these reasons, then, we have decided upon the last pre-war year, viz. 1938-39.

The data available are, of course, not exactly what we require, an almost inevitable feature in economic matters. In particular, the values given for capital and depreciation are book values — sometimes imputed — which are almost certainly highly formalized and which certainly would not correspond with the definitions used in Chapter VII. Also, information by industry about selling costs and overhead expenditure other than salaries and depreciation is non-existent, as is data on economic rents. No details are available as to stocks and consequently no calculation of the period of production is possible, so that no estimate can be made of working capital requirements which probably constitute an important element in the cost structure of some industries.

However, even allowing for these limitations, one would expect, if the argument of Chapter VII is substantially correct, to obtain some significant correlations between the percentage net profit margin\(^1\) and i.e. net of salaries and depreciation.
on the one hand and capital intensity on the other. Also, there should be some correlation between the N.P. margin and such independent indices of monopoly power, e.g. the degree of concentration of production or the degree of shelter against imports, which we are able to assemble. In fact, these correlations do not occur.

Searching around for the reason, one is inevitably made aware of the strong inverse correlation between the percentage net margin (and the percentage gross margin) and the fraction of value of output (price) represented by raw materials. Indeed, this can be seen at once by inspection. One example is sufficient to illustrate this effect. Thus, Sugar Refining is known independently to be one of Australia's most highly monopolistic industries. Raw materials represent 87 per cent of value of output, while the percentage net profit margin is only 8.4 per cent. By contrast, Toys, a fairly easily entered industry and one highly competitive with imports, has a net profit margin of 24 per cent, but with a materials bill representing only 45 per cent of value of output.

In Table 1 we have listed 60 industries of the
139 which are shown independently in the Production Bulletin. This selection is not strictly random, since we have selected the industries taking into account the following considerations:

(1) Industries wholly or substantially run by Government or semi-Government bodies, i.e. those run on a non-profit basis, have been excluded.

(2) Service rather than manufacturing industries, e.g. Boot Repairing and Dry Cleaning, have been omitted.

(3) Preference has been given to the better-defined industries over those which are in hotchpot or portmanteau classifications.

(4) At least one industry from each of the 15 broad industrial groupings has been included.

The resulting 60 industries is therefore a large and representative, although not strictly random, sample of Australian manufacturing. It is found that there is a simple inverse correlation between the percentage net margin and the ratio of raw materials to value of

2. Heat, Light and Power industries have been ignored because these are normally run by Governments or controlled public utilities. This practice is followed throughout this chapter.

3. One reason for this is the high proportion of Working Proprietors in these industries. This tends to seriously underestimate labour cost.
output of -0.69. This very strong correlation swamps other factors which enter into the determination of the percentage net profit margin.

It is clearly desirable that the influence of this factor should be eliminated from our measure of monopoly power. Now it can be shown that:

\[ M' = \frac{P - C}{P} = \left(\frac{P - C}{V}\right)(1 - \frac{m}{P}) \]

where \( m \) is the material bill and \( V \) is value added or income. Then, if we hold the term \( m/P \) constant, the ratio \( \frac{P - C}{V} \) suggests itself immediately as an appropriate index of monopoly power. This ratio we denote in future by the undashed \( M \).

It is interesting to pause for a moment to consider the significance of the choice of income or value added in place of value of output as the denominator of our measure of monopoly power. Two views may be taken of the matter. Firstly, from a philosophical standpoint, one would expect that only factors of production would earn a surplus in the process of production. Materials, once having been given a market value, would tend only to pass that value, without increment, to the final value of output.
In other words, in a market economy, the monopoly elements involved in the production of raw materials will have already been incorporated into their value. Then, if materials merely transfer their value to the final product, i.e., if they earn no surplus, the weighting of the denominator by the inclusion of materials will certainly skew the percentage margin, as indeed comes out in the figures. It should be pointed out, however, that insofar as an established firm has a monopsonistic advantage in the buying of raw materials not accessible to a new firm, the lower price of raw materials will be reflected in \( P - C \) and will thus enter into the overall degree of monopoly power enjoyed by such established firm.

At a lower level of generality, it is fairly clear, once one thinks about it, that a potential entrant firm, the firm which we have argued sets the limits to the monopoly power exercisable by established firms, will not be guided by the return it can expect upon turnover as such, except in the usually minor respect that it must cover its working capital costs which in turn will depend upon the value of the materials outlay as well as income-payments. Rather it will look to
get a minimum return on its capital. The idea is commonly put forward that firms determine their prices by adding a fixed percentage mark-up on cost, including material cost. Although this may appear as the formal method of price determination adopted as a matter of convenience, it can only operate within very narrow ranges. If there is a substantial and permanent increase, say, in the value of materials in relation to other costs, a fixed percentage mark-up would increase profits in relation to capital (and income) more than would be necessary to cover the increase in working capital requirements which would be involved. These profits could not for long be protected from new competition. However, failing substantial and permanent movements in material cost in relation to other costs, it is likely that firms tend to accept the good with the bad, as it were, and iron out short-period fluctuations by the application of fixed pricing formulae. However, this would involve no increase in long-period profits which, we have argued, will determine the degree of potential new or actual/competition to which established firms are exposed.
We may now proceed to examine our amended monopoly power formula in relation to our data. We should now expect to find some correlation between the share of net profit in value added or income, which we shall hereafter refer to as the "profits share of value added", and such independent criteria of monopoly as we may be able to construct. Two such indices suggest themselves:

(a) the degree of shelter from imports which an industry enjoys, an important factor in an economy such as Australia's; and

(b) a measure of the concentration of production in the various industries.

Also, one would expect some correlation between the profit share in value added and the degree of capital intensity. Let us examine these matters in turn.

The Degree of Shelter from Imports

It is very difficult, of course, to measure precisely the degree of shelter from imports enjoyed by one industry as compared with another. Not only are tariff structures very complicated, but other

4. Again "net" in that it is net of salaries and depreciation, the only overheads for which we have information.
considerations such as costs of transportation would have to be taken into account. However, a rough measure of the exposure of domestic producers to foreign competition could be based upon the actual volume of imports which enter the country taken in relation to the value of home-production. Unfortunately, it is not normally possible to get imports in classifications comparable with those of domestic manufacturing production. However, in the 1927 Tariff Committee's Report, Appendix G, we have imports and domestic manufactures compiled by the Commonwealth Statistician in comparable categories in respect of the years 1923-4, 1924-5 and 1925-6.

In Table 2 we have set out the profits share in value added\(^5\) in respect of 78 industrial classifications,\(^6\) together with imports as a percentage of the value of home-production. It is found that the simple correlation between these variates is \(-0.43\) which is highly significant for such a large sample.

Thus, given that the domestic monopoly power of producers is substantially influenced by the degree of

---

5. But this includes depreciation, information on depreciation not being available in these earlier years.

6. Only a handful of industries, such as Government-run ones, have been left out. The Tariff Report gives 86 separate classifications.
shelter from imports enjoyed by the domestic producer, we find this facet of monopoly power reflected in our amended index.

The Degree of Concentration

A few years ago, C.P. Haddon-Cave\(^7\) suggested the following formula as a measure of the degree of concentration of production:

\[
\text{Number of factories employing half the workers of an industry \over \text{Total Number of factories in the industry}}
\]

If we are prepared to use a rough interpolation, especially in the class of factories employing 101 hands or more, this ratio can be calculated from Table 9 - "Classification in Each Industry by Number of Hands\(^*\) - of the Production Bulletin.

There are, of course, some very serious limitations to such an index as a measure of concentration, the two most important being:

(a) It is an index of the concentration of factories, not of firms; and

(b) It necessarily calculates numbers on an

---

Australia-wide basis, whereas for many industries the "competitive field" does not extend beyond the boundaries of the one State. (Clearly, 6 firms competing throughout the Commonwealth represent something different from 6 firms each with a pure monopoly, or virtually so, within their own States.)

These difficulties are inherent and there is little we can do about them. However, there is a more fundamental difficulty with the Haddon-Cave index which we must consider. As a matter of fact, this index is not so much a measure of concentration in the usually accepted sense as of skewness of distribution. Consider two industries, the first of which embraces 100 firms of which 50 produce one-half the output (employ one-half of the workers), the second of which is composed of 2 firms, one of which produces half the output. In either case, half the output will be produced by 50 per cent of the firms (or factories in our case); and consequently the Haddon-Cave index will reckon both as having the same degree of concentration. From this example it is clear that the index is not very
sensitive to numbers as such, when 100 firms may take the same index of concentration as 2 firms. What we really need is an index which combines both the number of firms engaged in an industry and the proportion to whom falls the lion's share of output.

If we could feel confident that the boundaries of the Statistician's classifications were congruent with those of the competitive fields of the industries concerned, little difficulty would be experienced in constructing such an index. However, no-one would be found to assert such confidence, least of all the Statistician.

As a compromise, then, an index on the following lines is suggested:

(a) We take the number of firms employing half the workers (producing half the output) as of basic importance.

(b) However, the total number of firms engaged is also important, but less so. Therefore, we take the square root of this figure.

(c) Multiplication of (a) and (b) gives a figure which is a compound both of total
numbers and the skewness of their
distribution.

(d) However, since the absolute number of
firms (factories) appearing in any
industrial classification is largely
arbitrary, we damp our index by this figure.

Thus, our final index of concentration is given
by:

\[ c = \frac{n}{\sqrt{T}} \]

\[ = \frac{n}{\sqrt{T}} \]

where \( n \) is the number of factories producing half the
output (employing half the workers) and \( T \) is the total
number of factories within the classification concerned.

This ratio is found to be fairly stable and seems
to reflect sensibly what one would feel about
comparative degrees of concentration from first
principles. Thus, applying it to our example above,
in which 50 of 100 factories produce half the output
and 1 of 2 factories do likewise, we would get values
of 5 and \( 1/\sqrt{2} \approx 0.7 \) respectively. Although it gives
a broad result along these lines, it is not however
capable of fine distinctions which would tend to be
required when the numbers between which comparison is made are small. However, the data themselves are not sufficiently well-based to make such distinctions anyway.

At all events, we shall use this index as a rough measure of the degree of concentration of production. We may to some extent console ourselves that its deficiencies will be offset by our use of capital-intensity data which are likely, at least in some degree, to reflect firms' monopoly power.

Capital Intensity

For many purposes, the best measure of capital intensity is the ratio between capital and income (value added). However, in this context we are interested in testing the determinants of the mark-up on cost which itself constitutes an important element of value added. Since our independent numerical criteria of the contribution to the mark-up made by elements of monopoly are so limited, it seems better here to take a measure of capital intensity which is independent of the size of the monopoly mark-up, viz. the concept of capital per man. This is probably best estimated by the ratio of capital to wages and salaries. Again, it is not worthwhile attempting too
fine distinctions because, as has been pointed out, the figures for capital values are unlikely to be particularly precise.

* * * * *

We are now prepared to return to a consideration of our 1938-39 data. As already mentioned, no appropriately classified figures for imports are available in respect of 1938-39. Therefore, to eliminate the influence of this factor from our calculations, we confine ourselves to those industries which are sheltered or virtually so. Also, in view of the nature of our measure of concentration, we have had to keep for present purposes to even stricter criteria of industrial homogeneity. (Some industrial classifications are so hotchpot, so widely-drawn, that an index of concentration would be meaningless.) Forty industries more or less meeting these requirements have been assembled in Table 3, together with the percentage net mark-up on value added which we have designated by $M$. $M = P - C/V$ where $V$ is value added and $C$ includes salaries and wages and depreciation. Also, one further minor refinement
has been introduced. In view of our elimination of many industries from consideration for reasons stated above, it has been necessary, to maintain a fairly large sample, to take in some industries which have a large number of Working Proprietors in relation to employed hands, e.g. Bakeries. Where there is a high proportion of Working Proprietors, labour-cost tends to be seriously underestimated, as may readily be seen from an example. In the case of "Boots and Shoes" which has a negligible number of Working Proprietors, the share of wages and salaries in value added for 1938-39 was 71 per cent; in the case of "Boot Repairing", by no means a monopolistic industry but where Working Proprietors represent slightly more than half the work-force, the share of wages and salaries was only 30 per cent.

To counteract this effect, at least approximately, the wage and salary bill for each industry has been "blown up" in proportion as Working Proprietors constitute a certain percentage of employed hands, the recipients of wages and salaries.

In addition to \( M \) - the profit share of value added - as calculated, Table 3 shows:
(a) the degree of concentration for each industry according to our modified Haddon-Cave formula which we designate by \( c \); and

(b) the degree of capital intensity - the ratio of capital employed to the wage and salary bill - denoted by \( k \).

The simple correlation between \( M \) and \( c \) works out at \(-0.287\); that between \( M \) and \( k \) at \( 0.468 \). The simple correlation between capital intensity and concentration works out at \(-0.11\). The total correlation between \( M \) as an estimate of the degree of monopoly, on the one hand, and capital intensity and the degree of concentration as calculated, on the other, is found to be \( 0.525 \).

There other factors which would tend to contribute to the profits share of value added in addition to those we have considered, e.g. the amount of working capital required in the industry. Although no accurate estimate of working capital is possible, we could use the ratio of materials to value added as an index at least of abnormal working capital requirements arising from variations in the carrying costs of stocks. However, the concentration of such secondary factors is swamped by the error factors in our data both for capital intensity and the degree of
concentration of production. It is unlikely that we could improve our result by further refinements.

This comes out plainly enough once we introduce information of a qualitative kind which is not susceptible to systematic enumeration. Take by way of illustration those industries for which M has a value of 65 or greater. These are:

<table>
<thead>
<tr>
<th>Industry</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breweries</td>
<td>74</td>
</tr>
<tr>
<td>Soap</td>
<td>70</td>
</tr>
<tr>
<td>Ice Cream</td>
<td>69</td>
</tr>
<tr>
<td>Sugar Refining</td>
<td>65</td>
</tr>
<tr>
<td>Cereal Foods</td>
<td>65</td>
</tr>
<tr>
<td>Butterine &amp; Margarine</td>
<td>65</td>
</tr>
</tbody>
</table>

Now Breweries, Soap, and Sugar Refining are known independently to be among the most highly monopolistic of all Australian industries, although no arithmetic concept of concentration could be expected to bring this out fully. Ice Cream is generally monopolistic on a State-to-State basis, there being virtually no interstate competition for obvious reasons. Also, butter substitutes may only be manufactured under what amounts to monopoly licence, production being highly restricted by the State to prevent competition with butter. Cereal Foods, as well as Soap and Ice Cream, is a highly advertised commodity, with the bulk of the market in the hands of very few companies.
Likewise, at the other end of the scale, we have as the industries with $M$ of 30 or less:

- Saddlery 30
- Sugar Mills 30
- Joinery 29
- Cooperage 28
- Furniture 26

With the exception of Sugar Mills, these are essentially artisan manufactures conducted on a relatively small scale, with entry comparatively easy. The presence of Sugar Mills in the list may occasion some surprise at first sight, but this is almost certainly a result of the fact that the greater part of the industry is in the hands of growers' cooperatives.

It is clear that many of these factors would be difficult to embody in strict quantitative form, even if all the data were available, which it is not of course. The main point of introducing these observations here is to indicate that "profit share in value added" can be sensibly related to independently known facts of a qualitative kind about the degree of monopoly in the industries concerned. The percentage net margin (and the percentage gross margin) bear no such sensible relationship.
2. The Degree of Monopoly and the Percentage Gross Margin

It is interesting at this stage to compare briefly the measure of monopoly power elaborated in these two chapters with its nearest relative, viz. the percentage gross margin which was sponsored by Kalecki as a measure of monopoly power. Our measure differs from Kalecki's in three important respects:

(1) The Short-period versus the Long-period

The first difference between the two conceptual measures of monopoly power is that ours is explicitly related to the long-period, whereas Kalecki's, like Lerner's, is conceived to measure short-period phenomena.

In part, this difference is not a real one. For on the cost side Kalecki assumes constant marginal cost over the short-period with which he is concerned. Then, as far as changes in cost in relation to output is concerned, these will occur only as the result of long-period decisions. Moreover, on the demand side, short-period changes in
the volume of demand have no effect, since these merely imply movement along a horizontal marginal cost curve. It is, of course, the elasticity of demand which is the primary determinant of the firm's monopoly power in terms of Kalecki's thesis. But even this is unlikely to be determined, for the most part, on short-period considerations. Schumpeter was probably not far from the facts in his assumption that 'We will throughout act on the assumption that consumers' initiative in changing their tastes ... is negligible and that all changes in consumers' tastes is incident to, and brought about by, producers' action.' And most producers' action such as the changing of products and the addition of new products or the disappearance of old ones is essentially referable to the long-period. In short, Kalecki's measure of the degree of monopoly power is much more a creature of long-period factors than he allows.

(2) The Treatment of Overheads

The second difference of our measure from Kalecki's

is that it reckons as cost those overheads which must be met if production is to continue in the long-period, e.g. salaries, depreciation and ordinary overhead expenses such as postages, audit fees, lighting and so on. Kalecki's inclusion of such expenditure as elements of monopoly power has often been, correctly I believe, the subject of criticism. The problem of maintaining production in the long-period presents itself at once as soon as we consider the conditions for long-period equilibrium of Kalecki's model. Clearly, the attainment of such equilibrium as a result of the influx and efflux of firms must be conditional on firms being able to cover average total cost in some sense.

(3) The Influence of the Raw Material Factor

The third distinction between our position and that of Kalecki is, of course, that our measure is expressed in terms of income, whereas Kalecki's is referred to value of output which is equal to income plus raw material cost. This matter has already been discussed at some length. It is interesting, however, to notice how this influence comes out in Kalecki's own figures. In his Studies in Economic Dynamics,
he shows that in the year 1937, for example, the percentage gross margins in the United States were 24.8 per cent for manufacturing, 12.9 per cent for wholesaling and 28.7 per cent for the retail trades. Does this mean that manufacturers and retailers enjoy something like double the monopoly power of wholesalers? The answer almost certainly lies elsewhere. The material cost (purchase of goods turned over) is much higher in relation to value added for wholesalers than it is for manufacturers and retailers, for fairly obvious reasons. This higher cost of materials in relation to value added is reflected in the lower percentage gross margin. This confirms the argument advanced in this chapter; and indicates that this inverse relationship between the ratio of material cost to price and the percentage gross margin is a general one.

It may, of course, be objected that Kalecki never intended the percentage gross margin to measure the absolute value of monopoly power enjoyed or to measure comparative degrees of monopoly, but only to reflect changes in them. However, if the percentage gross margin does not measure such comparative degrees of

9. p. 16.
of monopoly power, then the taking of a weighted average for the economy or sector of the economy is quite meaningless.

Finally, the inverse correlation between the proportion of price represented by the material bill and the percentage gross margin also illuminates a fundamental aspect of Kalecki's overall thesis. Kalecki, it will be remembered, argued that the share of wages in income was a function of (a) the degree of monopoly as measured, i.e. by the percentage gross margin, and (b) the ratio of the raw material bill to the wage bill. This share had been fairly constant both throughout the trade cycle and in the long-period because of the scissorss movement of these two factors:

(1) In the long-period, there had been a steady increase in the degree of monopoly power which was more or less offset by a secular tendency for the prices of raw materials to fall in relation to the price of manufactured goods and to wages.

(2) Over the trade cycle, the share of labour tended to have remained constant because, in the slump, the degree of monopoly would rise, but this would be counterbalanced by the fact that
the price of raw materials fall more than industrial prices and money wages, and conversely in the boom.

Keynes, who welcomed Kalecki's work as something "which may prove to be an important piece of pioneer work," was nevertheless worried about the fact that Kalecki's two basic determinants of labour's share in national income worked in such a way as to always more or less exactly compensate for each other. Keynes says: "His own explanation is based on the assumptions that marginal real costs are constant, that the degree of the imperfection of the market changes in the opposite direction to output, but that this change is precisely offset by the fact that the prices of basic raw materials (purchased by the system from outside) relatively to money wages increase and decrease with output. Yet there is no obvious reason why these changes should so nearly offset one another ..."

However, if it is true, as has been argued here, that Kalecki's two determinants are inversely correlated, the difficulty of explaining why their movements tend to offset one another disappears at once.

11. loc. cit., p. 49.
The explanation of why labour's share in income tends to be fairly stable must be sought elsewhere.
<table>
<thead>
<tr>
<th>Industry</th>
<th>Percentage Net Margin*</th>
<th>Raw Materials to Value of Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cement &amp; Cement Goods</td>
<td>26</td>
<td>47</td>
</tr>
<tr>
<td>Glass</td>
<td>23</td>
<td>46</td>
</tr>
<tr>
<td>Glass Bottles</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td>White Lead, Paints &amp; Varnish</td>
<td>26</td>
<td>59</td>
</tr>
<tr>
<td>Oils, Vegetable</td>
<td>16</td>
<td>74</td>
</tr>
<tr>
<td>Oils, Mineral</td>
<td>27</td>
<td>66</td>
</tr>
<tr>
<td>Oils, Animal</td>
<td>14</td>
<td>64</td>
</tr>
<tr>
<td>Soap &amp; Candles</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>Chemical Fertilizers</td>
<td>14</td>
<td>68</td>
</tr>
<tr>
<td>Matches</td>
<td>22</td>
<td>52</td>
</tr>
<tr>
<td>Agricultural Implements</td>
<td>11</td>
<td>46</td>
</tr>
<tr>
<td>Brass &amp; Copper</td>
<td>18</td>
<td>55</td>
</tr>
<tr>
<td>Wireworking (inc. Nails)</td>
<td>15</td>
<td>64</td>
</tr>
<tr>
<td>Stoves &amp; Ovens</td>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td>Lead Mills</td>
<td>22</td>
<td>69</td>
</tr>
<tr>
<td>Wireless Apparatus</td>
<td>12</td>
<td>55</td>
</tr>
</tbody>
</table>

* Net of Salaries and Depreciation.
<table>
<thead>
<tr>
<th>Industry</th>
<th>P.N.M.</th>
<th>R.M./V.O.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watches &amp; Clocks</td>
<td>16</td>
<td>41</td>
</tr>
<tr>
<td>Cotton</td>
<td>16</td>
<td>61</td>
</tr>
<tr>
<td>Wool</td>
<td>13</td>
<td>62</td>
</tr>
<tr>
<td>Rope &amp; Cordage</td>
<td>20</td>
<td>56</td>
</tr>
<tr>
<td>Fellmongery</td>
<td>4.2</td>
<td>82</td>
</tr>
<tr>
<td>Tanning</td>
<td>12</td>
<td>67</td>
</tr>
<tr>
<td>Tailoring &amp; Slop Clothing</td>
<td>16</td>
<td>51</td>
</tr>
<tr>
<td>Shirts, etc.</td>
<td>13</td>
<td>60</td>
</tr>
<tr>
<td>Boots &amp; Shoes</td>
<td>13</td>
<td>54</td>
</tr>
<tr>
<td>Grain Milling</td>
<td>8.3</td>
<td>84</td>
</tr>
<tr>
<td>Cereal Foods</td>
<td>26</td>
<td>60</td>
</tr>
<tr>
<td>Cattle &amp; Poultry Food</td>
<td>10</td>
<td>79</td>
</tr>
<tr>
<td>Chaffcutting</td>
<td>9.7</td>
<td>76</td>
</tr>
<tr>
<td>Bakeries</td>
<td>21</td>
<td>61</td>
</tr>
<tr>
<td>Biscuits</td>
<td>20</td>
<td>57</td>
</tr>
<tr>
<td>Sugar Mills</td>
<td>6.2</td>
<td>79</td>
</tr>
<tr>
<td>Sugar Refining</td>
<td>8.4</td>
<td>87</td>
</tr>
<tr>
<td>Sugar Confectionery</td>
<td>23</td>
<td>57</td>
</tr>
<tr>
<td>Jam, Fruit &amp; Vegetable Canning</td>
<td>13</td>
<td>70</td>
</tr>
<tr>
<td>Pickles, Sauces &amp; Vinegar</td>
<td>24</td>
<td>58</td>
</tr>
</tbody>
</table>
Table 1 (Cont'd.)

<table>
<thead>
<tr>
<th>Industry</th>
<th>P.N.M. %</th>
<th>R.M./V.Q. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacon Curing</td>
<td>7.3</td>
<td>82</td>
</tr>
<tr>
<td>Butter</td>
<td>4.7</td>
<td>90</td>
</tr>
<tr>
<td>Butterine &amp; Margarine</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>Salt Refining</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Aerated Waters</td>
<td>32</td>
<td>48</td>
</tr>
<tr>
<td>Breweries</td>
<td>44</td>
<td>41</td>
</tr>
<tr>
<td>Distilleries</td>
<td>25</td>
<td>62</td>
</tr>
<tr>
<td>Winemaking</td>
<td>19</td>
<td>66</td>
</tr>
<tr>
<td>Malting</td>
<td>17</td>
<td>66</td>
</tr>
<tr>
<td>Bottling</td>
<td>60</td>
<td>25</td>
</tr>
<tr>
<td>Tobacco &amp; etc.</td>
<td>15</td>
<td>73</td>
</tr>
<tr>
<td>Ice Cream</td>
<td>40</td>
<td>43</td>
</tr>
<tr>
<td>Sausage Skins</td>
<td>11</td>
<td>62</td>
</tr>
<tr>
<td>Joinery</td>
<td>17</td>
<td>50</td>
</tr>
<tr>
<td>Cooperage</td>
<td>18</td>
<td>44</td>
</tr>
<tr>
<td>Boxes &amp; Cases</td>
<td>15</td>
<td>61</td>
</tr>
<tr>
<td>Bedding &amp; Mattresses</td>
<td>18</td>
<td>64</td>
</tr>
<tr>
<td>Manufactured Stationary</td>
<td>30</td>
<td>49</td>
</tr>
<tr>
<td>Cardboard Boxes, etc.</td>
<td>21</td>
<td>50</td>
</tr>
<tr>
<td>Papermaking</td>
<td>18</td>
<td>56</td>
</tr>
<tr>
<td>Industry</td>
<td>P.N.M.</td>
<td>R.M./V.O.</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>Tyres, Motor &amp; Cycle</td>
<td>16</td>
<td>62</td>
</tr>
<tr>
<td>Pianos, etc.</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Surgical, Optical &amp; Scientific Instruments</td>
<td>24</td>
<td>39</td>
</tr>
<tr>
<td>Toys, Games &amp; Sports Requisites</td>
<td>24</td>
<td>45</td>
</tr>
</tbody>
</table>

\[ n = 60 ; \quad r = -0.69. \]
# TABLE 2

**Profits Share of Value Added and Imports in Relation to Home Production**

(Year 1925-26)

<table>
<thead>
<tr>
<th></th>
<th>Percentage Profits share in Value Added</th>
<th>Imports as a Percentage of Home Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanneries</td>
<td>39%</td>
<td>5%</td>
</tr>
<tr>
<td>Sausage Skins</td>
<td>48%</td>
<td>79%</td>
</tr>
<tr>
<td>Oil &amp; Grease</td>
<td>73%</td>
<td>449%</td>
</tr>
<tr>
<td>Soap &amp; Candles</td>
<td>63%</td>
<td>3%</td>
</tr>
<tr>
<td>Bricks, Tiles, etc.</td>
<td>37%</td>
<td>15%</td>
</tr>
<tr>
<td>Glass (inc. Bottles)</td>
<td>51%</td>
<td>28%</td>
</tr>
<tr>
<td>Glass, Other</td>
<td>37%</td>
<td>100%</td>
</tr>
<tr>
<td>Lime, Plaster, Cement, etc.</td>
<td>54%</td>
<td>9%</td>
</tr>
<tr>
<td>Marble, Slate, etc.</td>
<td>36%</td>
<td>14%</td>
</tr>
<tr>
<td>Cooperage</td>
<td>30%</td>
<td>1%</td>
</tr>
<tr>
<td>Joinery, Boxes, Cases</td>
<td>31%</td>
<td>3%</td>
</tr>
<tr>
<td>Woodturning, Carving, etc.</td>
<td>44%</td>
<td>34%</td>
</tr>
<tr>
<td>Agricultural Implements</td>
<td>28%</td>
<td>35%</td>
</tr>
<tr>
<td>Brass &amp; Copper</td>
<td>52%</td>
<td>19%</td>
</tr>
<tr>
<td>Cutlery</td>
<td>47%</td>
<td>785%</td>
</tr>
<tr>
<td>Engineering, Ironworks, etc.</td>
<td>32%</td>
<td>52%</td>
</tr>
<tr>
<td>Product</td>
<td>Profit Margin</td>
<td>Imports</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Galvanized Iron, etc.</td>
<td>38</td>
<td>96</td>
</tr>
<tr>
<td>Nails</td>
<td>49</td>
<td>33</td>
</tr>
<tr>
<td>Stoves &amp; Ovens</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>Wireworking</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td>Electric Apparatus</td>
<td>38</td>
<td>336</td>
</tr>
<tr>
<td>Lamps &amp; Fittings</td>
<td>32</td>
<td>1033</td>
</tr>
<tr>
<td>Sewing Machines</td>
<td>67</td>
<td>867</td>
</tr>
<tr>
<td>Bacon Curing</td>
<td>57</td>
<td>1</td>
</tr>
<tr>
<td>Butter, Cheese, etc.</td>
<td>64</td>
<td>2</td>
</tr>
<tr>
<td>Butterine &amp; Margarine</td>
<td>68</td>
<td>1</td>
</tr>
<tr>
<td>Meat &amp; Fish Preserving</td>
<td>33</td>
<td>20</td>
</tr>
<tr>
<td>Biscuits</td>
<td>54</td>
<td>1</td>
</tr>
<tr>
<td>Confectionery</td>
<td>53</td>
<td>2</td>
</tr>
<tr>
<td>Cornflour, Oatmeal, etc.</td>
<td>61</td>
<td>9</td>
</tr>
<tr>
<td>Flour Mills</td>
<td>62</td>
<td>-</td>
</tr>
<tr>
<td>Jam, Fruit Canning, etc.</td>
<td>50</td>
<td>6</td>
</tr>
<tr>
<td>Sugar Refining</td>
<td>74</td>
<td>1</td>
</tr>
<tr>
<td>Aerated Waters, etc.</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td>Breweries</td>
<td>70</td>
<td>3</td>
</tr>
<tr>
<td>Condiments, Coffee, etc.</td>
<td>67</td>
<td>12</td>
</tr>
<tr>
<td>Distilleries</td>
<td>70</td>
<td>232</td>
</tr>
<tr>
<td>Tobacco, Cigars, etc.</td>
<td>65</td>
<td>7</td>
</tr>
<tr>
<td>Salt</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td>Item</td>
<td>Profit Margin</td>
<td>Imports</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Cider</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>Animal &amp; Poultry Foods</td>
<td>66</td>
<td>-</td>
</tr>
<tr>
<td>Woollen &amp; Tweed Mills</td>
<td>43</td>
<td>82</td>
</tr>
<tr>
<td>Cotton Mills</td>
<td>-43</td>
<td>2250</td>
</tr>
<tr>
<td>Boots &amp; Shoes</td>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td>Clothing</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>Dressmaking &amp; Millinery</td>
<td>39</td>
<td>14</td>
</tr>
<tr>
<td>Furriers</td>
<td>52</td>
<td>40</td>
</tr>
<tr>
<td>Hats &amp; Caps</td>
<td>42</td>
<td>30</td>
</tr>
<tr>
<td>Waterproof &amp; Oilskins</td>
<td>48</td>
<td>74</td>
</tr>
<tr>
<td>Shirts, Ties, etc.</td>
<td>39</td>
<td>4</td>
</tr>
<tr>
<td>Hosiery &amp; Knitted Goods</td>
<td>54</td>
<td>43</td>
</tr>
<tr>
<td>Rope, Cordage &amp; Bags</td>
<td>50</td>
<td>207</td>
</tr>
<tr>
<td>Tents &amp; Sailmaking</td>
<td>56</td>
<td>1</td>
</tr>
<tr>
<td>Papermaking, Paper Products</td>
<td>47</td>
<td>119</td>
</tr>
<tr>
<td>Musical Instruments</td>
<td>37</td>
<td>111</td>
</tr>
<tr>
<td>Perambulaters</td>
<td>37</td>
<td>18</td>
</tr>
<tr>
<td>Saddlery, Harness, etc.</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>Spokes, etc.</td>
<td>46</td>
<td>5</td>
</tr>
<tr>
<td>Billiard Tables, Furniture, etc.</td>
<td>34</td>
<td>3</td>
</tr>
<tr>
<td>Picture Frames</td>
<td>47</td>
<td>12</td>
</tr>
<tr>
<td>Window Blinds</td>
<td>49</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Profit Margin</td>
<td>Imports</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Seagrass &amp; Wickerware</td>
<td>31</td>
<td>7</td>
</tr>
<tr>
<td>Baskets, etc.</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td>Brooms &amp; Brushware</td>
<td>37</td>
<td>50</td>
</tr>
<tr>
<td>Chemicals, Drugs, etc.</td>
<td>67</td>
<td>71</td>
</tr>
<tr>
<td>Fertilizers</td>
<td>60</td>
<td>26</td>
</tr>
<tr>
<td>Paint, Varnishes, etc.</td>
<td>64</td>
<td>34</td>
</tr>
<tr>
<td>Essential Oils</td>
<td>63</td>
<td>241</td>
</tr>
<tr>
<td>Inks, Polishes, etc.</td>
<td>71</td>
<td>12</td>
</tr>
<tr>
<td>Surgical, Optical &amp;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scientific Instruments</td>
<td>43</td>
<td>140</td>
</tr>
<tr>
<td>Electroplating</td>
<td>36</td>
<td>64</td>
</tr>
<tr>
<td>Manufacturing Jewellery</td>
<td>40</td>
<td>102</td>
</tr>
<tr>
<td>Matches</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>Carbide</td>
<td>-4</td>
<td>88</td>
</tr>
<tr>
<td>Leather Goods</td>
<td>42</td>
<td>14</td>
</tr>
<tr>
<td>Rubber Goods, Tyres</td>
<td>42</td>
<td>63</td>
</tr>
<tr>
<td>Toys</td>
<td>55</td>
<td>408</td>
</tr>
<tr>
<td>Umbrellas</td>
<td>53</td>
<td>9</td>
</tr>
</tbody>
</table>

\[ r = -0.43 \]
### TABLE 3

**NET PROFIT SHARE IN VALUE ADDED, CAPITAL INTENSITY, AND THE DEGREE OF CONCENTRATION**

*(Year 1938-39)*

<table>
<thead>
<tr>
<th>Industry</th>
<th>Profit Share in V.A.</th>
<th>Capital Intensity</th>
<th>Index of Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lime, Plaster &amp; Asphalt</td>
<td>43</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Cement &amp; Cement Goods</td>
<td>48</td>
<td>4.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Bricks &amp; Tiles, etc.</td>
<td>31</td>
<td>2.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Glass Bottles</td>
<td>45</td>
<td>1.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Boiling Down Works, etc.</td>
<td>42</td>
<td>3.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Soap &amp; Candles</td>
<td>70</td>
<td>2.4</td>
<td>0.5</td>
</tr>
<tr>
<td>Chemical Fertilizers</td>
<td>44</td>
<td>6.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Matches</td>
<td>46</td>
<td>3.3</td>
<td>0.9</td>
</tr>
<tr>
<td>Brass &amp; Copper</td>
<td>38</td>
<td>1.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Wireworking (inc. Nails)</td>
<td>39</td>
<td>1.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Stoves &amp; Ovens</td>
<td>33</td>
<td>1.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Tanning</td>
<td>35</td>
<td>1.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Saddlery, Harness &amp; Whips</td>
<td>30</td>
<td>1.6</td>
<td>1.4</td>
</tr>
<tr>
<td>Grain Milling</td>
<td>50</td>
<td>4.2</td>
<td>3.0</td>
</tr>
<tr>
<td>Cereal Foods &amp; Starch</td>
<td>65</td>
<td>4.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Cattle &amp; Poultry Foods</td>
<td>48</td>
<td>4.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Chaffcutting</td>
<td>32</td>
<td>1.8</td>
<td>4.7</td>
</tr>
<tr>
<td>Industry</td>
<td>M</td>
<td>k</td>
<td>o</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td>Bakeries</td>
<td>44</td>
<td>2.6</td>
<td>8.6</td>
</tr>
<tr>
<td>Biscuits</td>
<td>45</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Sugar Mills</td>
<td>30</td>
<td>7.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Sugar Refining</td>
<td>65</td>
<td>9.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Sugar Confectionery</td>
<td>53</td>
<td>2.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Jam, Fruit &amp; Vegetable Canning</td>
<td>43</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Bacon Curing</td>
<td>39</td>
<td>2.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Butter</td>
<td>45</td>
<td>3.8</td>
<td>3.2</td>
</tr>
<tr>
<td>Condensed Milk</td>
<td>45</td>
<td>4.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Butterine &amp; Margarine</td>
<td>65</td>
<td>1.9</td>
<td>0.7</td>
</tr>
<tr>
<td>Salt Refining</td>
<td>50</td>
<td>7.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Aerated Waters</td>
<td>56</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>Breweries</td>
<td>74</td>
<td>4.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Winemaking</td>
<td>54</td>
<td>4.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Malting</td>
<td>51</td>
<td>4.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Tobacco, etc.</td>
<td>56</td>
<td>1.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Ice Cream</td>
<td>69</td>
<td>5.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Joinery</td>
<td>29</td>
<td>1.1</td>
<td>3.2</td>
</tr>
<tr>
<td>Cooperage</td>
<td>28</td>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Table 3 (Cont'd.)

<table>
<thead>
<tr>
<th>Category</th>
<th>M</th>
<th>k</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boxes &amp; Cases</td>
<td>34</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Billiard Tables, Furniture, etc.</td>
<td>26</td>
<td>1.0</td>
<td>3.2</td>
</tr>
<tr>
<td>Bedding &amp; Mattresses</td>
<td>46</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Cardboard Boxes, Cartons, etc.</td>
<td>42</td>
<td>2.1</td>
<td>1.3</td>
</tr>
</tbody>
</table>

\[ n = 40. \]

\[ r_{MK} = 0.468; \quad r_{Mc} = -0.287; \quad r_{kc} = -0.11 \]

\[ R = 0.525 \quad (38 \text{ d. of f.}) \]
IX. CHANGES IN THE DEGREE OF MONOPOLY

In this Chapter, it is proposed to consider the chief factors underlying changes in the degree of monopoly, both over the course of the trade cycle and in the long-period. However, before doing so, it is necessary that we briefly examine the doctrine of "countervailing power" which has achieved wide acceptance in recent years. Clearly, the weight given to this thesis will have a large bearing on what we may say about changes in the degree of monopoly.

1. Monopoly Power: Additive or Countervailing?

Galbraith sets out his doctrine of countervailing power in the following terms:

"With the widespread disappearance of competition in its classical form and its replacement by the small group of firms if not in overt, at least in conventional or tacit collusion, it was easy to suppose that since competition had disappeared, all effective restraint on private power had disappeared.

In fact, new restraints on private power did appear to replace competition. They were nurtured by the same process of concentration which impaired or destroyed competition. But they appeared not on the same side of the market but on the opposite side, not with competitors but with customers or suppliers. It will be convenient to have a name
for this counterpart of competition and I shall call it countervailing power.

... private economic power is held in check by the countervailing power of those who are subject to it. The first begets the second. The long trend toward concentration of industrial enterprise in the hands of a relatively few firms has brought into existence not only strong sellers, as economists have supposed, but also strong buyers as they have failed to see. The two develop together, not in precise step but in such manner that there can be no doubt that the one is in response to the other. *(1)*

Thus, it is argued, countervailing power may be exercised against a monopolist or group of monopolists from both sides, either from the supply side or the demand side. Suppliers of materials or factors of production (especially labour) are likely to combine to increase their bargaining strength vis-a-vis the monopoly and, by so doing, force the concession of some of the monopoly gains inherent in the situation to themselves. On the other hand, purchasers of the products of a monopoly or monopoly group will tend to band together to reduce the power of exploitation enjoyed by the monopolist. The fundamental question which we have to ask is whether this development of derivative monopoly does in fact act as an off-setting or countervailing factor or

---

merely adds more monopoly areas to the system. The problem is analogous to the old taxation question: to what extent may a monopoly be forced through pressure of countervailing power to absorb demands made upon it, to what extent may these demands be merely passed on to other sections of the community.

In tackling this question, we shall confine ourselves simply to market power. There is every reason to believe that in the modern age large corporations exert a considerable power over and above that which they enjoy in the market; but, although for other purposes this is a very important problem, it will not be our concern here.

In considering the effect of countervailing power in market relationships, it is, perhaps, useful to distinguish between countervailing power originating outside the system from that involving two units within the system. For this purpose, wage-earners and consumers may be considered to be outside the system. By contrast, one firm purchasing the products of another may be looked upon as something internal to the system. The point of the distinction resides

2. cf. President Roosevelt, for example: "The liberty of democracy is not safe if the people tolerate the growth of private power to a point where it becomes stronger than their democratic State itself." - Broadcast, 29 April, 1938.
in the fact that an increase in the monopoly power of one unit in the system may involve a decrease in that of another such unit; whereas it would mean straining the use of language and the blurring of useful social distinctions to represent the reduction of market power exerted by a monopolist as a result of countervailing power by consumers, for example, as constituting an increase in the monopoly power of such consumers.

It is probably best, then, to approach our problem, in the first instance, in terms of the simpler case, viz. where the countervailing power is exerted from outside the system. Let us begin by considering the effect of the combination of labour on the monopoly power of employers. At its simplest, our problem is whether an increase in wages forced from a monopolist by trade union action is likely to be absorbed by the monopolist, thereby reducing his effective monopoly power, or whether it will be merely passed on to subsequent units of the productive and distributive structure.

Let us consider the problem in terms of our model of Chapter VII. How will an increase in wages forced by trade union action affect the monopoly power of a firm, as measured by that model? The first thing
to notice is that, in general, an increase in wages established over a given area by the countervailing power of trade union organization will affect the position of existing firms and potential newcomers in that area alike. Thus, as a first approximation, it would seem likely that the wage-increase would be passed on into the price charged by the monopolist, since this would not render the monopolist any more vulnerable to competition from outsiders than before. However, three factors which work in the opposite direction will tend to modify this result, according to the particular circumstances.

(1) Where the wage-increase is not co-extensive with all producers or potential producers of the commodity in question.

In fact, a wage-rise enforced by trade union action will very rarely, if ever, affect all producers and potential producers alike. Even when a wage-increase is forced over a national area, it will not normally be operative in respect of foreign producers. Clearly, where a monopolist or group of monopolists has, sooner or later, to meet "foreign" competition as price rises, there is a definite limit to the extent to which an increase in wages may be
This fact has become of considerable importance in business strategies in the era of monopoly capitalism. Firstly, it gives rise to those familiar situations where industries are highly monopolistic at home, in fierce competition abroad. The inter-war period of competition between the higher-wage countries and Japanese manufacturers is very instructive in this respect. Certain subsidiary facets of international economic strategies are implicit in such a situation. Governments feel bound to support their monopolists in their struggle with foreign competitors; and this in turn leads to the whole range of restrictive trade practices, tariff walls and even subsidized dumping in the international sphere. These policies tend to be reinforced by the political and economic power of the monopolists vis-à-vis their Governments. In these terms we may explain a number of outstanding features of latterday capitalism:

(1) We get strong pressures to make the area of monopolistic agreement wider and wider until it is nothing less than world-wide. The twentieth century has seen the growth of many international cartels, perhaps the most
highly organized and most powerful being the chemicals triumvirate - I.C.I., I.G. Farben and Du Pont.

(2) Where such agreement is not attainable, we tend to get a situation of discriminating monopoly with the home market and the foreign market sealed into watertight compartments. Nazi Germany probably achieved the zenith of this form of discrimination. At all events, its effectiveness and endurance depends upon Governmental support, and this in turn hinges, to a large extent, on the extra-market power of monopolies.

(3) We tend to get the shifting of industry from high wage to low wage countries, as when British capital sponsors cotton-goods manufacture in Bombay and Shanghai.

These phenomena do not, of course, arise solely out of attempts to forestall reduction in monopoly power by the exercise of countervailing power. But they do provide escape mechanisms in this direction, and hence do set definite limits to the degree in which trade union action, for example, may effectively lower the monopoly power of modern big business.
Although these phenomena come out most sharply in the international sphere, they apply equally to areas within a nation. Just as lower-wage countries tend to put a brake upon the ability of firms to pass on wage-increases, so do lower-wage areas within a nation. Clearly, the power of a monopoly or group of monopolists on the one hand and the countervailing power of trade unions on the other will depend on the area over which each party exercises effective control.

(2) Where the elasticity of demand (consumers' sensitivity) increases as relative price rises.

It will often be the case, especially where expenditure on the product in question is a significant fraction of total income, that buyers will become more sensitive as the price of the one commodity rises in relation to other prices. Buyers will search more diligently for possible alternatives, and this must make it easier for new producers to establish themselves in the market. Thus, upward pressure on wages by trade unions may, for this reason, result in a squeezing of the monopoly power enjoyed in the particular industry, the full impact of the wage-rise not being passed on into price for fear of the easier conditions of entry which would result.
(3) Where substitution between factors is possible.

In many cases substitution between factors is strictly limited by technical considerations. However, there will be cases where substitution is feasible. At first sight, it might appear that the possibility of factor substitution would merely set a limit to the amount by which trade union action may force up wages, since if these were forced to the point where substitution on a considerable scale took place, unemployment would result and this would exercise a downward pressure on wages. In other words, it would be likely to reduce the bargaining power of the union. However, at any particular time a new producer (not being committed to a given plant and equipment) is in a much better position to undertake such substitution between factors. As a result, a new producer may be able to enter an industry on better terms if wages have risen. This fact may discourage existing firms from fully passing on a wage increase, thereby reducing the monopoly power which they effectively enjoy.

We may conclude then that it is impossible to say a priori just how far trade union action may be effective in reducing the degree of monopoly in any particular industry. In some cases, it is no doubt
fairly high. In others, it is probably negligible. It will all depend upon the particular market circumstances of the industry concerned. To some extent, also, it will depend upon the extra-market power which firms can exert to the end of getting Government or other official protection in the form of tariffs and the like.

What we can say for certain is that the extent by which countervailing power exercised by trade unions can offset market monopoly power is limited. In particular, the extent by which monopoly margins may be squeezed by trade union pressure depends on the degree in which one firm or group of firms may be forced into a more disadvantageous position vis-a-vis other existing firms or potential firms. It follows that the more general and ubiquitous the countervailing power exerted by trade unions the less likely is it, other things being equal, to be successful in forcing a reduction in monopoly margins. Thus, for example, a general increase in wage-rates reduces the possibilities of factor substitution because capital good prices will be affected equally with labour costs. Indeed, it is not the primary aim of individual trade unions to squeeze the margins of
the employers of their members in particular. It is only of minor importance to a particular union whether its successful wage demands are met by a reduction in the margins of their employers or by the additional wage-costs being passed on to the community in general. Certainly there seems to be a good deal of evidence that a large part of the gains of highly organized sections of the workers is passed through the system, coming to rest upon less-protected sections, e.g. white-collar workers, small businesses and fixed income groups.

This brings us to the question of the exercise of countervailing power by one business against the monopoly power of another. Clearly, such countervailing power may be exerted either from the supply side or the demand side. On the supply side, an increase in the price of raw materials forced by countervailing action by suppliers will represent, to the business purchasing them, an analogous situation to that in which a wage rise is forced by trade union action. Again, it is impossible to say a priori how far such a rise in materials prices will be absorbed or passed on. However, there is one important difference in this case from the case of a wage-rise. The overall degree of monopoly will not be decreased in any event. Insofar
as the price rise is passed on to final consumers, there will be a net increase in the total degree of monopoly in the system. Insofar as the price increase is absorbed by the firms buying the materials, the reduction in the monopoly power of these firms will be offset by the increased monopoly power of the suppliers of the materials. There is, in effect, a redistribution of monopoly power from one firm or group of firms to another.

On the demand side, when one business unit is buying from another, the effects are similar. Any increase in the monopsonistic power of the buyer may either be offset by a reduction in the monopoly power of the supplier or merely additive to it. In either event there will be no net reduction in the overall degree of monopoly. The overall monopoly power of business as a whole can only be reduced from the demand side by the exercise of countervailing power by final consumers, that is at the exact level where the organization of such power is most difficult to maintain, especially over a long period.

Stigler has made this point strongly. He says: "Nor is there any explanation, in Galbraith’s book or elsewhere, why bilateral oligopoly should in general
eliminate, and not merely redistribute, monopoly gains." He adds further: "It is natural to ask why the operation of the economy is improved when a monopsonist or a set of oligopsonists arises and shares the gains of a previously unhampered monopolist or set of oligopolists." Indeed, the effect may well be adverse to the balance of the economy, especially where, for example, monopsonistic buyers or monopolistic sellers use their power to squeeze firms in fairly competitive industries. Over long effective periods, the monopoly power of such firms may be forced below zero as a result. Thus, we may get a situation, such as in the United States during the 1920's, in which monopoly profits are very high indeed, while a large proportion of businesses is subject to protracted losses.

Another point worth noticing is that, if two firms on opposite sides of the market are faced with the choice between struggling for the redistribution of monopoly gains between them and making their monopoly power additive in effect, there will be a powerful incentive for them to combine in the common interest, either by full vertical integration or some

2b. loc. cit., p. 9.
agreement stopping short of actual integration.

Finally, Galbraith himself suggests that countervailing power will not operate in a period of excess demand or in conditions of a seller's market, because the power of buyer resistance will be small. This overlooks the fact that the very power of a monopolist lies in his ability to create for himself the conditions of a seller's market. Moreover, one would have thought that the increased power of sellers in conditions of excess demand would be matched by the increased power of the ultimate sellers, namely factors of production, especially labour.

Galbraith is undoubtedly correct that the power of exploitation increases in conditions of excess demand. But this arises out of the fact there is an overall bottleneck in the supply of means of production, which provides existing monopolists with more than normal protection against the encroachments of "outsiders".

2. Changes in the Degree of Monopoly Over the Long-Period

There has been a general inclination among economists to accept as a fact that the degree of monopoly has been steadily increasing over the last 50 years or so, at least in the developed capitalist countries such as the United States, United Kingdom, Germany and Japan. This belief is based on two
outstanding and generally admitted features of recent capitalist development:

(1) On the one hand, we have seen the carrying through from time to time of great merger movements resulting in business units under unitary control of an unprecedented size and scale. Alongside of this, there has developed a complex system of business techniques, falling short of full amalgamation, designed to tie business units more closely - the trust, holding company and cartel, subsidiary shareholding and pyramiding, the interlocking of directorates, and so forth. Other devices, such as the trade association, have developed for the purpose of regulating competition between independent firms in those industries, usually, with larger numbers of firms.

(2) Parallel with these developments, there has been a remarkable increase in certain forms of market imperfection - advertising and a whole new range of selling devices, product differentiation, trademarking, and the like.

However, against these trends, certain countering factors are generally conceded. In particular, the enormous advances in transport and communication have resulted in lower transportation costs relatively, on the one hand, and better knowledge
among both buyers and sellers on the other. Also, the development of mass production techniques have probably brought about a certain degree of standardization in secondary goods, even though this process has, as yet, nowhere near attained the possible in this respect. Nevertheless, the feeling has remained that, on balance, there has been a definite increase in the degree of monopoly since somewhere around the eighties of last century.

In part, this is an optical illusion, so far as straight-out market power is concerned. The illusion derives from the size-dimension of modern business. However, in many cases, the growth of large-scale industry has represented merely the substitution of, say, a national monopoly for a few local monopolies. Also, although business has grown in size, so has the market; and size, of itself, is only significant in relation to the size of the market. In Australia, for example, we do not get the same feeling of sheer size as in the United States, although in many cases the degree of monopoly is higher in Australia than the United States because of the smallness of the market compared with economic levels of output required by modern production techniques.
In passing, however, it should be stressed again that we are here dealing with market power. It is probably true that, so far as extra-market power is concerned, size, of itself, is very important.

Finally, we must take account of the development of countervailing power. We have seen that we must allow some effect to this factor. Not least of these effects will be the way in which countervailing power may re-distribute monopoly power. Indeed, one would expect that the growth of great monopoly power in some sectors of the economy would involve, among other things, the reduction in the market power of the more competitive sectors. Thus, we may find that the weighted average degree of monopoly may be constant, while the monopoly power of some sections is increasing, that of others is decreasing.

There are two further difficulties of assessing just how the overall degree of monopoly has changed or is likely to change in the long period. Firstly, there is a distinction between a state of given monopoly power and the growth of monopoly power. Clearly, the market power exertable by individual firms may be reduced, as the result of a struggle for position and survival between rivals, while at the same
the outcome is certain to result in a higher degree of monopoly. Just as an economy can guarantee itself higher consumption in the future by foregoing it now, so monopolists may not fully exploit the potentialities of their present position for the sake of a better position later on.

Secondly, insofar as the degree of monopoly power - and of countervailing power - which may be exerted is a function of the level of output and of utilization, the overall degree of monopoly power will vary with long-period changes in these factors. Just to take one example, it is clear that long-period conditions of surplus labour, as in Britain throughout the period between the two world wars, will affect the countervailing power of trade unions. What makes the problem particularly intractable is the fact that, for the economy as a whole, the conditions of long-period levels of output and utilization, are almost certainly not independent of the degree of monopoly.

3. Cyclical Changes in the Degree of Monopoly

On the question of how the degree of monopoly varies with cyclical changes in the level of output, there have been greater differences of opinion among
economists. The earliest discussion of this matter was by Pigou who argued that the degree of monopoly could be expected to rise in the slump relative to the boom.³ He gave four reasons why producers would be reluctant to reduce price in relation to cost to the full extent during a slump:

(i) The mutual fear among producers of "spoiling the market" which "means selling a thing in bad times at such a price, and, therefore, in such quantities, that in subsequent good times the market is already stocked and producers cannot benefit by the then good demand."⁴

(ii) The existence of "psychological friction" - the reluctance of business-men to consider bygones as bygones particularly in respect of prices paid for stocks of materials.

(iii) The influence on demand of expectations. Reduction in price which engenders the belief that price will fall further may actually reduce current demand, and, hence, producers will be hesitant in cutting price to the extent required by the situation.

³ A.G. PIGOU: Industrial Fluctuations, London 1929, Ch. XIX.
⁴ op. cit., p. 186.
(iv) The convenience of a stable selling price in some types of business will lead to the maintenance of a higher price relative to cost in the slump, a lower price relative to cost in the boom.

Harrod, by contrast, has held that the degree of monopoly will fall during a slump because consumers, with lower real incomes, will become more sensitive to price and selective of quality. The reverse will hold true in the relative affluence of the boom. Thus, firms will be faced with a more elastic demand curve in the slump, a less elastic demand curve in the boom. This effect Harrod elevates to the status of a "law", namely "The Law of Diminishing Elasticity of Demand" which he tells us "means that, as output as a whole increases and individuals become more affluent, their sensitiveness to price differences declines."\(^5\)

More recently, Kalecki has argued that, on balance, the degree of monopoly is likely to rise during a slump.\(^6\) In the first place, he argues, we may expect that transport costs will not fall as much as the general level of prices in a slump and

will not rise as much in the boom. Hence, the degree of market imperfection will tend to move inversely with the level of output. Against this, however, must be placed the Harrod Effect which moves in the opposite direction.

However, he says: "Probably more important than the effect of cyclical changes in market imperfection upon \( \mu \) is the influence of 'tacit agreements' in a deep slump which may be classified as changes in the degree of oligopoly." The basis of these "tacit agreements" is the financial deterioration of firms in a slump. Industry demand is much less elastic than individual-firm demand. Price-cutting by all firms will lead to substantial reductions in margins without any great expansion of demand. Hence, producers will be reluctant to cut price and this reluctance, if it is general, amounts to a tacit agreement. Of course, in many cases firms will make firm agreements between themselves to avoid a common debacle.

It will be noticed that Kalecki's position is broad enough to embrace the specific factors set out by Pigou. Moreover, as was pointed out in Chapter IV,

7. op. cit., p. 18.
the Harrod Effect and the thesis that the degree of oligopoly rises in the slump are not necessarily inconsistent. In fact, the one may contribute to the other. The greater interdependence which derives from the increase in substitutability due to increased consumer sensitivity may serve to throw producers into each other's arms, as it were. As mentioned before, non-aggression pacts may well indicate a heightened threat of war, as when a firm's rivals move closer to it.

However, it is not sufficient to consider merely what firms would want to do in a slump; it is equally important to consider what they are able to do, in the way of protecting their margins. We have argued in Chapter VII that the ultimate sanction upon a firm's ability to exploit a monopoly position comes from the threat of "outside" competition. This sanction applies both to an individual firm acting unilaterally and to a group of firms acting in "tacit agreement", or in open agreement for that matter. Therefore, it is instructive to re-consider cyclical changes in the degree of monopoly in terms of our model of Chapter VII.

Changes in the facility with which outsiders may enter an industry will occur as a result of either
changes in demand conditions which make the marketing (selling) task of the new producer harder or easier, or changes in cost conditions which put the new producer in a more or less advantageous position vis-a-vis existing producers, or changes in the degree of imperfection in the capital market which affect the ease with which the new producer may obtain capital, or, finally, changes in the degree of oligopolistic cohesion within the industry.

On the demand side, the Harrod Effect, insofar as it does operate, will serve to make entry easier for the outsider and will thereby tend to reduce the degree of monopoly. Whereas the effect of increased consumer sensitivity may be cancelled out, as between firms, by open or tacit oligopolistic action in the slump, the Harrod Effect will, other things being equal, tend to lower the ceiling at which it will be profitable for firms to maintain price by agreement. This is the ultimate sanction of increased consumer sensitivity.

Against this effect on the demand side, however, must be placed the fact that, during a boom or upswing, demand will be expanding, whereas in a slump it will be stationary or contracting. With an expanding
demand - with new buyers of the particular commodity coming into the market - the effect of consumer attachments will be weaker. In general, it is easier to form consumer attachments than to change them. For this reason, the demand curve facing a potential entrant may be, on balance, more elastic in the boom than in the slump.

On the cost side, there are a number of changes as between boom and slump which may alter the comparative ease of entry in the one direction or the other. Firstly, insofar as transportation costs are "sticky" in the slump, this will make it more difficult for "foreign" producers to deliver their products to the market. Also, where tariff or other protective devices are based on a flat money rate rather than on an ad valorem system, the degree of protection will rise relatively as price falls. On the other hand, where protection takes the form of fixed quotas, the relative degree of protection will fall as price and output within the industry contracts. In addition, if the prime costs of "foreign" producers fall more than those of "domestic" producers, this will make the danger from outside competition greater. In the extreme case of dumping, there may be no limit to the reduction in local monopoly power which might result.
In the case of direct entry into the industry by new producers there is also a number of factors which may affect the comparative ease of entry one way or the other. Firstly, severe setbacks in the output of existing producers may virtually abolish protection deriving from economies of large-scale. Real prime costs of operating a large plant at low capacity may rise sharply compared with the cost of operating a new, smaller-scale plant. A new producer will, of course, be reluctant to commit himself to a small plant permanently and this must count against the ease of entry. Nevertheless, the scope for a mushrooming of small-scale temporary "backyard production" is likely to be greater in the slump. This tendency will be reinforced by the availability of persons with the necessary "know-how" who may have been displaced by the general contraction of activity.

If the prices of producers goods fall more than the general level of prices, this will put the new producer at an advantage, insofar as there is scope for substitution between labour and this lower-priced capital. The old producer, on the other hand, is not able to make this substitution, so long as he is committed to old capital investment. The buying up of bankrupted plants "for a song" may also operate to
make possible entry on more favourable terms.

Against this, however, must be set the fact that, in the liquidity context of the slump and in the prevailing climate of expectation, it will be more difficult for a potential producer to obtain the necessary finance to enter upon production. This difficulty will be greater than that merely involved in the almost certain rise in interest rates. Moreover, the difficulty of raising finance will be considerably enhanced where existing producers have links, e.g. common directorates, with banking and other financial institutions, such that pressure can be brought to prevent new producers from obtaining capital.

The effect of a slump on the degree of "oligopolistic cohesion" will also depend on a variety of factors. A large amount of excess capacity and, perhaps, large stockpiles are a standing invitation to the individual firm to slash price unilaterally in an attempt to "beat the gun." This temptation will be greatly increased if the firm is in liquidity difficulties. On the other hand, individual survival is probably best assured by firms entering into some agreement to stand or fall together. Adversity will, in many situations,
serve to close the ranks, as it were. Moreover, the discipline which can be maintained in an industry will depend in part on the immediacy of the threat from "outsiders". Firms in more vulnerable positions in relation to outside competition will be unwilling to go all the way with those better protected. On the other hand, any easing of the threat from outside the group is likely to make agreement easier to maintain.

Over and above all these considerations, however, is the influence of expectations. On the one hand, the existence of large amounts of excess capacity throughout an industry must tend to dampen the enthusiasm of a potential entrant. Firstly, it will make oligopolistic reaction more probable. Secondly, it raises the threat of internecine warfare breaking out at some future time, if the slump is protracted. Generally, the whole matrix of expectation will work in this direction. In the boom, there will be a natural tendency for expectations to be framed on the basis of current boom conditions, and vice versa in the slump. In the depths of a depression, it will take a good deal of courage to predict how long it will take before "good times" return, and to act upon the prediction.
After all, the new producer must commit himself to long-term capital investment.

Finally, we must consider how countervailing power is likely to work as between boom and slump. It is fairly clear that the countervailing power exertable by trade unions will fall in the slump, when unemployment is substantial, as compared with the boom. As far as consumer organization is concerned, this is probably more effective in the slump, although this is hard to decide a priori. However, it would seem that consumer attitudes which form the basis of the Harrod Effect will tip the scales in favour of more critical purchasing in the slump. How countervailing power as between business units is concerned will change in a slump is also a moot question. Probably, the effect of depressed business conditions will be to increase the power of those who have it, to decrease that of the weaker. To whom that has shall be given.

How then would we expect the average degree of monopoly to vary as between boom and slump. Clearly, this will depend on how the variety of factors we have considered balance out. Of course, for individual industries this balancing will be less
likely. Some industries will be subject to one set of influences, some to another. Hence, we should expect that in some cases the degree of monopoly will be substantially increased, in others substantially reduced. Just how this will end in the average is a difficult question to answer in view of the variety of forces at work, although one is left with an impression that, on balance, one would expect the average to rise. Moreover, insofar as firms with a lower degree of monopoly are more likely to be forced out of production than those with a high degree of monopoly — in Chapter VII we showed that the barriers to the exit of firms is of the same order as the barriers to entry — one would expect the average to rise. To what extent this happens is likely to depend on the depth and duration of the slump. In a short, sharp slump, with few firms forced out of production, we may get a reduction in the average degree of monopoly. However, if the slump is at all protracted, the dropping out of firms with a low degree of monopoly in preference to those with a high degree of protection must tend to raise the average degree of monopoly, other things being equal.
However the average may move, the distribution of monopoly power is hardly less significant. We have already noticed that there would seem to be a reasonable presumption that, other things being equal, monopoly power is likely to increase during the slump in those sectors most protected, to decrease in the less protected areas. Hence, one would expect that the dispersion of degrees of monopoly as between industries would tend to increase in the slump as compared with the boom. However, insofar as it is likely to be the weakest firms of the less protected industries which will go to the wall in greatest number, this will tend to raise the average degree of monopoly in those less protected industries. As a result, the dispersion of monopoly power between industries may fall on this account.

4. Some Figures

In Tables 9.1 and 9.2 are set out the shares of wages in value of production for Australian manufacturing industry between 1910 and 1952-53 and similar figures for the United States between 1899 and 1939. Taking the share of wages in the value of production as a (rough) reciprocal index of the degree of monopoly, let us reconsider our argument in the light of these figures.
The data of Tables 9.1 and 9.2 are graphed in Figures 9 and 10. The Australian figures are probably less instructive than the American for two reasons. Firstly, Australia is a dependent economy and hence is more susceptible to the influence of "arbitrary" factors impinging upon it from outside. Tariff policies will also be of greater significance for a country which imports a large proportion of its manufactured goods requirements. Secondly, the Australian wage system is in some degree arbitrary and will only reflect market conditions within broad limits.

Nevertheless, there are a number of interesting similarities between the Australian and American data, as well as some important differences. In the first place, both show a distinct rise in the share of wages (fall in the degree of monopoly) in the sharp slump of 1920-21. Also, both show a strong downward movement in labour's share in the deep slump of the early 'thirties. This is indeed the tendency one would expect in terms of the theoretical discussion of the previous section. Moreover, both series show a tendency to rise in the weak boom conditions around 1937. However, this tendency should be looked at as a reversal of the deeply depressed conditions of
the immediately preceding period, rather than as an
effect of an independent boom. Clearly, the way in
which the degree of monopoly is likely to behave in
a boom will depend on how it has fared in the preceding
slump.

Perhaps the most striking dissimilarity between
the two series is the fact that, though a rise in the
share of labour occurred in both countries in the sharp
slump of 1920-21, in the following period the
Australian figures show a tendency for the 1920-21
level to be maintained or increased at least until
1926-27, whereas the American series falls steadily
from the peak of 1921 to the depths of the depression
in the early 'thirties. The explanation of this
divergence of behaviour is almost certainly to be
found in the different degree of maturity in the
respective economies:
(1) The 1920's were years of rapid expansion of
Australian manufacturing industry, involving the
establishment of many kinds of manufactures, new to
Australia but established in the field of international
competition. In such periods of establishment of new
industries, the share of labour is likely to be high
because firms will be earning small profits or even
incurring losses in the formative stages, e.g. the
share of labour in value added for the cotton industry was over 100 per cent throughout the formative years of the 'twenties. Probably lags in the adjustment of tariff policies also helped to raise the share of labour in this period.

(2) By contrast, the United States was a mature economy by the 1920's. There is also a good deal of independent evidence that the degree of monopoly rose during this period, both as a result of mergers, the establishment of new industries on a monopoly basis, e.g. aluminium, and the flowering of product differentiation, advertising, and other facets of market imperfection.

Another interesting feature of the Australian figures is the behaviour of the series in the period after World War II. During the war the share of labour rose to record levels under the influence of price and profit controls. In the post-war period, the wartime level of labour's share in value added has been maintained. This could be the outcome of a number of influences. In the first place, the period was one of rapid expansion of manufacturing into new fields and, in this respect, resembles the period following World War I which we have already discussed. Also, it was a period of brimful employment
with very high levels of output and satisfactory profits for industry generally. Maintenance of the high wartime level of labour's share was no doubt assisted by the workings of the Arbitration system and, in particular, by automatic wage adjustments, in a period of satisfactory profits or better when employer resistance is likely to be at its lowest. In the background, we may also add as a factor the countervailing power of trade unions in a period of brimful employment.

In Figures 11 and 12 are photostat copies of graphs prepared by Steindl showing the movement of the share of wages in value added of 14 American industries. These can illuminate our ideas derived from the average; they show a number of interesting characteristics in this respect. Firstly, in the short slump of 1920-21, there was a synchronous movement of all 14 industries in the same direction; the share of labour tended to rise in each case. Moreover, the downward trend of the 1920's represented a fairly general pattern. On the other hand, following the onset of the Great Depression in late 1929, we find individual industries moving in either direction, a trend which continued throughout the 'thirties. While firms tended to come together during the 1920's,
Figure 11 in particular shows a distinct tendency for industries to move apart in the 'thirties.

A serious limitation upon the use of empirical data relating to the movement of the degree of monopoly is that the figures are necessarily ex post. Insofar as the degree of monopoly may be both a cause and a consequence of changes in the level of output, we have an inherent difficulty of interpretation. There is no way of telling how the degree of monopoly would have moved had the degree of utilization and the level of output remained constant. A similar difficulty presents itself in the problem of determining how certain factors influence the distribution of monopoly power between business units. Thus, in a contraction of the level of output, certain firms will be forced out of the distribution and the sample becomes, at least in part, a selected one. Firms do not drop out at random. What we would really like to know is how changes in certain variables would have affected the distribution of monopoly power between firms, other things remaining equal. It is an unfortunate fact about economic argument that it is very seldom that other things do remain equal.
TABLE 9.1

Share of Wages and Salaries in Value Added:

Australian Manufacturing*

<table>
<thead>
<tr>
<th>Year</th>
<th>%</th>
<th>Year</th>
<th>%</th>
<th>Year</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td>49.6</td>
<td>1924-25</td>
<td>55.3</td>
<td>1939-40</td>
<td>51.8</td>
</tr>
<tr>
<td>1911</td>
<td>51.0</td>
<td>1925-26</td>
<td>55.6</td>
<td>1940-41</td>
<td>53.5</td>
</tr>
<tr>
<td>1912</td>
<td>51.7</td>
<td>1926-27</td>
<td>55.8</td>
<td>1941-42</td>
<td>56.9</td>
</tr>
<tr>
<td>1913</td>
<td>51.6</td>
<td>1927-28</td>
<td>54.6</td>
<td>1942-43</td>
<td>59.3</td>
</tr>
<tr>
<td>1914</td>
<td>51.2</td>
<td>1928-29</td>
<td>54.3</td>
<td>1943-44</td>
<td>59.2</td>
</tr>
<tr>
<td>1915</td>
<td>50.1</td>
<td>1929-30</td>
<td>54.2</td>
<td>1944-45</td>
<td>57.3</td>
</tr>
<tr>
<td>1916</td>
<td>50.2</td>
<td>1930-31</td>
<td>52.8</td>
<td>1945-46</td>
<td>58.4</td>
</tr>
<tr>
<td>1917</td>
<td>49.4</td>
<td>1931-32</td>
<td>50.4</td>
<td>1946-47</td>
<td>58.3</td>
</tr>
<tr>
<td>1918</td>
<td>48.2</td>
<td>1932-33</td>
<td>49.8</td>
<td>1947-48</td>
<td>58.4</td>
</tr>
<tr>
<td>1918-19</td>
<td>49.1</td>
<td>1933-34</td>
<td>49.9</td>
<td>1948-49</td>
<td>59.7</td>
</tr>
<tr>
<td>1919-20</td>
<td>49.7</td>
<td>1934-35</td>
<td>50.7</td>
<td>1949-50</td>
<td>58.3</td>
</tr>
<tr>
<td>1920-21</td>
<td>53.3</td>
<td>1935-36</td>
<td>50.5</td>
<td>1950-51</td>
<td>58.3</td>
</tr>
<tr>
<td>1921-22</td>
<td>52.4</td>
<td>1936-37</td>
<td>50.7</td>
<td>1951-52</td>
<td>59.8</td>
</tr>
<tr>
<td>1922-23</td>
<td>54.0</td>
<td>1937-38</td>
<td>52.0</td>
<td>1952-53</td>
<td>58.7</td>
</tr>
<tr>
<td>1923-24</td>
<td>54.7</td>
<td>1938-39</td>
<td>52.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Source: Commonwealth Statistician's Production Bulletins.
<table>
<thead>
<tr>
<th>Year</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1899</td>
<td>44.3</td>
</tr>
<tr>
<td>1904</td>
<td>43.7</td>
</tr>
<tr>
<td>1909</td>
<td>42.2</td>
</tr>
<tr>
<td>1914</td>
<td>43.2</td>
</tr>
<tr>
<td>1919</td>
<td>42.2</td>
</tr>
<tr>
<td>1921</td>
<td>45.0</td>
</tr>
<tr>
<td>1923</td>
<td>43.0</td>
</tr>
<tr>
<td>1925</td>
<td>40.4</td>
</tr>
<tr>
<td>1927</td>
<td>39.9</td>
</tr>
<tr>
<td>1929</td>
<td>36.9</td>
</tr>
<tr>
<td>1931</td>
<td>36.9</td>
</tr>
<tr>
<td>1933</td>
<td>36.5</td>
</tr>
<tr>
<td>1935</td>
<td>39.9</td>
</tr>
<tr>
<td>1937</td>
<td>41.1</td>
</tr>
<tr>
<td>1939</td>
<td>37.8</td>
</tr>
</tbody>
</table>

*Source: J. Steindl, Maturity and Stagnation in American Capitalism, Table 9.*
Share of Wages and Salaries in Value Added:

Australian Manufacturing

% Share of Wages and Salaries in Value Added

1910
1912
1914
1916
1918
1920
1922
1924
1926
1928
1930
1932
1934
1936
1938
1940
1942
1944
1946
1948
1950
1952
1954
1956
1958
1960
1962
1964
1966
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
(305)
FIGURE 10
Share of Wages in Value Added:
U.S. Manufacturing

(306)
A study of profit margins in U.S. manufacturing industry.

Fig. 4a and 4b. The development of the share of wages in value added in various groups of industries, 1899 to 1939. Data from Table 9.
PART III

MONOPOLY AND THE RATE OF INVESTMENT
X. MONOPOLY AND INVESTMENT: THE PROBLEM

1. Brief Historical Background

The problem of monopoly has been the concern of economists ever since the beginnings of economics as a separate and self-contained discipline. Indeed, the very birth of economics was contemporary with the challenge by the new bourgeoisie to the economic philosophy of Mercantilism; and high among the deadly sins of Mercantilism was its encouragement and protection of monopolistic practices. The classical economists were virtually unanimous in their condemnation of monopoly, not only as a vehicle of exploitation, but also as a brake on economic progress. Adam Smith, for example, warned:

"By a perpetual monopoly all the other subjects of the State are taxed very absurdly in two different ways; first, by the high price of goods, which, in the case of free trade, they could buy much cheaper; and, secondly, by their total exclusion from a branch of business which it might be both convenient and profitable for many of them to carry on."(1)

Around the middle of the nineteenth century, John Stuart Mill was explaining that "to be protected

against competition is to be protected against idleness, in mental dullness."²

Although Marx did not believe competition to be particularly beneficent, he was, on this subject, in closer concord than usual with the main stream of contemporary economic opinion. However, Marx went a good deal further than the classical economists in his assessment of the dangers of monopoly as a brake on capitalist progress. On the one hand, he predicted, as one of the most significant features of capitalist development, that centralization of capital and concentration of production would develop as a fundamental trend which would ultimately destroy the liberal-competitive capitalism of the nineteenth century. In the long-run, this trend would play an important part in the eventual supercession of capitalism by socialism. For Marx, the growth of monopoly was in the nature of the hardening of capitalism's arteries.

By the turn of the century, the development predicted by Marx was beginning to become evident, especially in Germany and the United States. Although J.B. Clark in the United States could be

found reiterating that monopoly was "hostile to improvement" and represented a "weakening of dynamic energy and a reduction of progress", the ever-cautious and astute Marshall was already tempering the traditional view:

"For when the production is all in the hands of one person or company, the total expenses involved are generally less than would have to be incurred if the same aggregate production were distributed among a multitude of comparatively small rival producers. They would have to struggle with one another for the attention of consumers, and would necessarily spend in the aggregate a great deal more on advertising in all its various forms than a single firm would; and they would be less able to avail themselves of the many various economies which result from production on a large scale. In particular they could not afford to spend as much on improving methods of production and the machinery used in it, as a single large firm which knew that it was certain to reap the whole benefit of any advance it made."(4)

The last sentence, in particular, puts the "novelty" of the Schumpeterian position of some 50 years later in rather a new light. By and large, however, this approach remained exceptional in theory, if not in practice; and the traditional pro-competition and anti-monopoly attitudes continued to be dominant.

In the English tradition, attitudes were again thrown into the melting pot in the 1920's, following

---

World War I. The British economy was then making heavy weather of "recovery" from the war-time disruptions. This gave rise to two opposing views. On the one hand, envious eyes were cast at the cartelized Germany and trustified U.S.A., Britain's chief competitors. Britain was exhorted to emulate their example under the banner of "rationalization". On the other hand, there were many who believed Britain's limping recovery to be due to monopolistic rigidities which had developed in the economic system and which frustrated the readjustments required to put the economy on its feet. Alongside this, there was a growing concern among academics, not confined to Britain, about the relevance of traditional value theory in a world of product-differentiation, advertising and high-pressure salesmanship.

At least within the academic halls, the victory went to traditional attitudes towards monopoly. The elaboration of the doctrine of imperfect or monopolistic competition not only underlined traditional criticisms of monopoly - restriction of output and higher prices - but also was at pains to stress the generality of the phenomena. Not a little of the blame for the economic shambles of the 'thirties...
was allotted to monopoly.

Then came the "Keynesian Revolution" ...

Prior to Keynes, investment was more or less assumed to take care of itself. The rate of interest was believed to balance the demand for investment funds (investment) with the supply of them (savings). Keynes repudiated this comfortable assumption and selected investment as the primary parameter of economic activity. Also, the distribution of income gained a new significance over and above that of mere economic justice. For the distribution of income was the main determinant of the consumption function which, together with the rate of investment, determined the level of Effective Demand and, hence, of economic activity.

It was natural, then, for economists to become more interested in the effects of the growth of monopolistic forms upon the inducement to invest and the distribution of income. The latter question appeared fairly straightforward and it was generally taken for granted that, other things being equal, monopoly skews the distribution of income in favour of the monopolists. However, the problem of the inducement to invest is a much more complex and slippery business; and a lively controversy has
grown up around the question of to what extent monopoly inhibits or facilitates investment.

A typical post-Keynesian view of the matter was advanced by Hansen:

"We have noted that the approaching cessation of population growth and the disappearance of new territory for settlement and exploitation may cut off a half or more of the investment outlets which we were wont to make in the past. We are, thus, compelled to fall back upon that measure of capital formation which is associated with the advance of technique and the rise of per capita output. But current institutional developments are restricting even this outlet. The growing power of trade-unions and trade associations, the development of monopolistic competition, of rivalry for the market through expensive persuasion and advertising, instead of through price competition, are factors which have rightly of late commanded much attention among economists. There is moreover, the tendency to block the advance of technical progress by the shelving of patents.

Under vigorous price competition, new cost-reducing techniques were compulsorily introduced even though the scrapping of obsolete but undepreciated machinery entailed a capital loss. But under the monopoly principle of obsolescence new machines will not be introduced until the undepreciated value of the old machine will at least be covered by the economies of the new technique. Thus, progress is slowed down and the outlets for new capital formation, available under a more ruthless competitive society, are cut off. Capital losses which could not be avoided under rigorous price competition can be and are avoided under an economic system more closely integrated by intercorporate association and imperfect competition. If we are to save the one remaining outlet for private capital formation, deliberate action of a far bolder character than hitherto envisaged must be undertaken in order to make the price system and free enterprise sufficiently responsive to permit at least that
measure of capital formation to which the rate of technological progress had accustomed us in the past."(5)

The counter-attack against this view was led by Schumpeter. Schumpeter held that the view that monopoly inhibited investment in new technical processes was a coloured one and maintained that, if anything, possession of monopoly control over a market would encourage innovation. In particular, he argued:

(1) that technical progress depends on research and, in the modern age, only very large businesses are able to provide research facilities on the scale required to sustain such progress;

(2) that long-period investment of the huge funds nowadays required will only be undertaken by a firm which has a sufficiently secure place in the market as will assure it of reaping at least a substantial part of the benefits which it plans for;

(3) that a monopoly will scrap existing capital equipment as soon as production with a new technique is less costly than with existing

equipment and, in this respect, the monopolist is in no different position from a socialist planning authority; and

(4) that the argument that monopoly inhibits technical innovation cuts across the argument, often advanced by the same people, that a large part of the unemployment of the 'thirties was the result of the displacement of workers by machines — so-called "technological unemployment".

A typical rejoinder to this Schumpeterian position is that of Maurice Dobb:

Schumpeter's is "an argument which seems to ignore the extent to which monopolies spend time and energy in entrenching an established position against the encroachment of rival innovations and in resisting the intrusion of enterprising newcomers on to the field — to ignore the fact that consideration of the unfavourable effect of new methods on the value of capital sunk in older methods will (during the length of life of the old plant) exercise an influence, and a retarding influence, under monopoly, which it could not do under conditions of atomistic competition."

And so the debate has gone on.

It will be the argument of the following section that, in large part, the reason for the inconclusiveness of the debate is that a number of distinctly different problems have become entwined in the argument and that

a necessary pre-condition for a solution is to sort them out. A different answer will be obtained, according as we ask the question from one point of view or the other.

2. Elements of the Problem

Prior to Keynes, economists were concerned with the effects of monopoly from three main points of view:

(1) The effect of monopoly upon the internal efficiency of business under three principal heads:

(a) slackness and conservatism in management;

(b) adaptability to technical change; and

(c) adaptability to changes in demand conditions or to structural changes in the economy.

(2) The general efficiency of the economy in allocating resources which has been held to be distorted by monopoly. (This is the special problem of the "welfare" economists.)

(3) The exploitation of the consumer, looked at as a matter of economic justice.

All these aspects of the monopoly problem are fundamentally concerned with the best use of given resources, a matter with which pre-Keynesian economics
was primarily concerned.

However, with the acceptance of the Keynesian position, these problems become secondary to that of determination of aggregate output. We are immediately faced with the Keynesian paradox: inefficient use of resources may increase aggregate output. As Keynes himself says: "To dig holes in the ground, paid for out of savings, will increase, not only employment, but the real national dividend of useful goods and services." Once this position is accepted, the effects of monopoly will be looked at from two different angles:

(4) the effect on investment considered as an income-generating phenomenon; and

(5) the effect on the distribution of income looked at, not so much from a point of economic justice, but as a primary determinant of the consumption function, and also as a determinant of the funds available to non-monopoly businesses for the purpose of investment.

These latter two sets of problems are quite distinct from the first three, and demand a different approach.

During the course of the controversy about the

effects of monopoly on investment, these two kinds of problem have not been clearly delineated. Much post-Keynesian discussion about the relation of monopoly to the investment process has mixed up the basic Keynesian conception of investment as an income-generating element with questions of efficiency and the allocation of given resources, the major preoccupation of pre-Keynesian economists.

Two factors seem to account for this confusion:

(1) Investment in the Keynesian system is taken as an independent variable, something given from outside the system. Although later attempts to close the system have brought in induced investment as a dependent variable, autonomous investment — and in particular the rate of technical innovation — has remained outside the system. Because economists have tended to regard investment as data and have not systematically attempted to study the determinants of investment decisions, there seems to have been a natural tendency to approach the question along traditional, i.e. pre-Keynesian, lines.

(2) This tendency has been reinforced in virtue of the fact that both sets of problems, viz. the
efficient use of resources and the sufficient generation of income by investment, have certain factors in common. In particular, technical innovation is a concern of both efficiency and income-generation. Likewise, it is easy to slip from consideration of the effect of monopoly on the distribution of income as a basic determinant of the consumption function into the question of high prices and consumer exploitation.

The basic purpose of this Part is, then, to try and sort out the factors which are relevant from the point of income-generation. Questions of efficient use of resources are important but separate, and must be kept separate if we are to put our questions in the right form.

Looked at from another angle, we could say that problems of efficiency and allocation of resources is the special problem of a full employment economy. The problem of income-generation is the problem of the economy with a chronic tendency towards underemployment. As it happens, the former problem tends to loom larger at the present day. Indeed, there are economists who believe that the problem of income-generation has been solved once
and for all. Governments, with the aid of economists no doubt, can fix all that. But however much one may be sceptical about this — especially if the priming effects of armament production were to taper off — the fact still remains that a general theory should be able to provide answers appropriate to both types of situation. Economists of all people should not easily forget the era of the "new capitalism" of the 1920's with its promise of endless vistas of progress and prosperity. At all events, in the following pages we shall be concerned with the effects of monopoly on investment looked at as an income-generating phenomenon. If the future should prove this to have been an unnecessary pursuit, the world will have much to be thankful for.

Finally, there is one further distinction which it is essential to make. There is a fundamental difference between monopoly taken as a growing process and monopoly taken as a given structural situation. According as we are dealing with the one situation or the other, our answer about the effect of monopoly on investment may be quite different. It may well be that the growth of monopoly will involve substantial investment outlays. For example, the growth of monopoly in an industry brought about
by amalgamation or merger of a number of existing firms may involve considerable technical reorganization of production, with considerable investment expenditures associated with it. In fact, historically, great periods of amalgamation and merger in the United States have been associated with high levels of investment, e.g. in the period 1898-1903 and in the nineteen-twenties. It is, of course, difficult to separate cause and effect here; but it seems likely that the process of amalgamation and merger contributed to the investment outlays in those periods.

On the other hand, the process of growth of monopoly power may serve to damp down investment expenditures. Insofar as an increased degree of monopoly brings with it a restriction of output or, at least, a check to the rate of expansion of output in the particular industry, this may tend to reduce investment outlays below what they would have been in the absence of the increase in monopoly power.

At all events, these processes tend to be of a once-for-all character and, usually, to be irreversible. They will not be our concern here. Rather, we are interested in the effects of a given
state of monopoly power upon the rate of investment. This, by contrast, represents a continuing influence.

3. The Distribution of Income

One consequence of monopoly is, of course, to skew the distribution of income in favour of those sections enjoying the monopoly power. The increased proportion of income accruing to these sections will usually be at the expense of wage and salary earners and the incomes of non-monopolistic businesses. In general, then, we should expect this to result in greater savings and a corresponding lowering of the consumption function, so that a greater rate of investment will be required to sustain a given rate of economic growth. This is a very important facet of monopoly, but one with which we are not primarily concerned here.

However, it would seem fairly certain that, at least to some extent, redistribution of income will affect the inducement to invest. For example, reduction in the incomes of more competitive businesses in favour of monopolies may leave the former with an acute and continuing shortage of investible funds. To the extent to which distributive relationships do affect investment outlays, we are bound to take this into account when considering the effect of monopoly on investment.
XI. THE EFFECT OF MONOPOLY ON THE RATE OF INVESTMENT

1. Investment in Cost-reducing Innovations

In the controversy which has grown up around the question of how monopoly affects the rate of investment, perhaps the greatest emphasis has been placed upon the effects in relation to cost-reducing innovations. This is no doubt due to the fact, mentioned in the previous chapter, that it is here that we find a direct association with questions of efficiency and productivity. For us, however, these are secondary matters. We are here concerned with the problem from the point of view of the income-generating consequences of technical progress.

As outlined in the previous chapter, there has been a distinct cleavage of opinion about the effect of monopoly on technical innovation. On the one hand, it is held that monopoly will inhibit such investment. A monopoly is in a position to protect existing capital values. Therefore, it will not introduce an innovation unless such investment reduces cost to an extent that the savings made are at least equal to the capital cost (interest + amortization)
of the new technique. Hence, the introduction of cost-reducing innovations will be delayed by monopolists until a sufficient degree of amortization of existing capital values has been achieved. Innovations are then financed primarily from amortization funds released from depreciation of existing capital equipment. And, of course, this is not net investment.

Hansen has been one of the chief protagonists for this viewpoint. As we noticed in Chapter 10, he says: "Under vigorous price competition, new cost-reducing techniques were compulsorily introduced even though the scrapping of obsolete but depreciated machinery entailed a capital loss. But under the monopoly principle of obsolescence new machines will not be introduced until the undepreciated value of the old machine will at least be covered by the economies of the new technique. Thus, progress is slowed down and outlets for new capital formation available under a more ruthless competitive society are cut off. Capital losses which could not be avoided under rigorous price competition can be and are avoided under an economic system more closely integrated
by intercorporate association and imperfect competition.\textsuperscript{1}

Now it is easy to refute this argument on a purely formal level, to show that the monopolist is faced with the same calculations as a competitive firm or socialist management; and a number of economists have taken the trouble to do so. Thus, Austin Robinson puts the matter as follows:

"We must now enquire whether a monopoly may be expected to adopt improvements of technique as rapidly as will competing firms. ... Let us start by considering in what conditions a firm is ordinarily prepared to substitute new equipment for old. It pays it to do so only when total cost of the required output with the new equipment, including a sufficient return on the capital invested, is less than prime cost with the old equipment. The old equipment is there in any case, whether it earns any return or none. If prime cost with the old equipment is less than total cost with the new, the additional profits earned through putting in the new equipment will be less than sufficient to pay the interest on the extra capital invested in it. ... So far as they are concerned with these considerations, the actions of plants owned by monopolies and by competing firms may be expected to be identical.\textsuperscript{2}"

Schumpeter presents a similar argument. Thus, he says:

"Another doctrine has crystallized into a slogan, viz., that in the era of big business the maintenance of the value of existing investment -

\begin{itemize}
  \item 1. Fiscal Policy and Business Cycles, p.
\end{itemize}
conservation of capital - becomes the chief aim of entrepreneurial activity and bids fair to put a stop to all cost-reducing improvement. Hence the capitalist order becomes incompatible with progress."(3)

But, he goes on:

"Conserving capital values is the same thing as conserving profits. Modern theory tends in fact to use the concept Present Net Value of Assets (= capital values) in place of the concept of profits. Both asset values and profits are of course not being simply conserved but maximized. ... All that private management tries to do is to maximize the present net value of total assets which is equal to the discounted value of expected net returns. This amounts to saying that it will always adopt a new method of production which it believes will yield a larger stream of future income per unit of the corresponding stream of future outlay, both discounted to the present, than does the method actually in use. The value of past investment, whether or not paralleled by a bonded debt that has to be amortized, does not enter at all except in the sense and to the extent that it would also have to enter into the calculation underlying the decisions of a socialist management. So far as the use of the old machines saves future costs as compared with the immediate introduction of the new methods, the remainder of their service value is of course an element of the decision for both the capitalist and the socialist manager; otherwise bygones are bygones for both of them and any attempt to conserve the value of past investment would conflict as much with the rules following from the profit motive as it would conflict with the rules set for the behaviour of the socialist manager."(4)

4. op. cit., pp. 96-97.
The argument put forward by Robinson and Schumpeter is, of course, formally correct. However, it is only correct subject to some very weighty ceteris paribus conditions; and to accept it as it stands would be highly misleading.

One important exception, which Robinson acknowledges but Schumpeter does not, suggests itself at once. It is clear that many cost-reducing innovations depend for their efficacy upon large-scale output. For example, the substitution of rotary presses for flatbeds by a printer only becomes economic if they can be employed in the production of large outputs. It is easy to think of a host of such examples. Now it is clear that, if an innovation requires an expansion of output for its profitable employment, the attitude of a monopolist or oligopolist to its introduction will be distinctly different from the attitude of a competitive firm. A competitive firm may expand output at the expense of competitors: a monopolist cannot, at least without a fierce struggle with its attendant losses. Clearly, a monopolist or small group of monopolists will be less enthusiastic about using the price weapon to expand output to the degree necessary for the economic
employment of the innovation than will a competitive firm. In short, marginal revenue, which must be balanced against the cost of the innovation, will be lower for the monopolist than for the competitive firm.

But let us return to the general case where an innovation reduces cost independently of any contingent increase in output. In this case, looked at from a static standpoint, Robinson and Schumpeter are logically correct. However, it is surprising that Schumpeter of all people should choose to rest his case on these static considerations. For the one thing he has been at pains to stress throughout his economic writings is that it is new firms and new products (or new variants of the old) which are the vehicle of innovation and technical progress.

Whereas under competition there is more or less freedom for new firms or new products to capture the market required for economic operation, under monopoly there is not; otherwise the situation would not remain one of monopoly for long. Once we admit the possibility of new firms entering the industry, the argument of Robinson and Schumpeter falls to the ground. For a new firm will be interested not only in innovations which reduce total cost below existing
firms' prime costs, but is able to consider entry into the field so long as total cost falls below the going price, which is quite a different thing. By offering price competition, such new firms can force existing firms to re-equip or go out of business. This ability to force innovation under a competitive regime is greatly enhanced when the innovation, as is often the case, also involves an improvement in the product, as well as reducing the cost of producing it. New firms are not concerned at all with existing capital values, since they have no capital invested in the field. Thus, they can inflict capital losses on existing firms which those firms would not incur voluntarily, if they were able to control innovation as can the monopolistic firm.

This distinction between the behaviour of monopolistic and competitive firms is even more clearly demonstrated when we consider innovations which are purely capital-saving. Clearly, it will never pay an existing firm to invest in purely capital-saving innovations which merely replace equipment already in use. By contrast, a new firm will be indifferent as to whether an invention be capital-saving or labour-saving. Both will reduce total cost below the
going price.

It is interesting to notice how Schumpeter shifted his argument to avoid these implications. He does not, it will be recalled, say in so many words that the inducement to invest in cost-reducing innovations will be the same for a monopoly as for a competitive firm. Rather, he argues that it will be the same for a monopolist as for a socialist management. Schumpeter is correct in this claim, save for the exception noticed earlier where introduction of a new technique is contingent upon an increase in output. However, the fact is that a socialist economy will not reproduce the conditions of a competitive economy in this respect. Under socialism it will also be the case that innovation will only become worthwhile when the total cost with the new technique is lower than prime cost with existing equipment. It will not pay a socialist management to scrap working equipment otherwise. Certainly a socialist authority would not set up new plants so as to force scrapping of equipment used by existing plants, unless this condition is fulfilled. Nor will it pay to scrap existing equipment in the face of a purely capital-saving innovation. In these respects the position of socialist management is
the same as for a monopolist. Moreover, it is easy to see the reason for this. In both cases property values are vested in one or a few hands which are able to protect them. But under competition the property used in the industry is diffused, lots of firms owning a small part of it. So long as a competitive firm, or potentially competitive firm, finds it profitable to invest in a particular innovation, it is unconcerned with what happens to the capital values vested in other people. It is quite prepared to inflict capital losses on others. Neither a monopolist nor a socialist management will take this individualist attitude. They must clearly concern themselves with the total effects of their action.

From the point of view of conserving resources, monopoly is in this respect superior to competition, which may destroy existing capital before it ceases to be socially useful. But this very superiority is inimical to the development of the economy once we look at the matter from the point of view of investment as an income-generating phenomenon. Misdirection of resources and the destruction of
still socially useful equipment does provide investment outlets under competition, although in a socially inefficient way. Monopoly on the other hand tends to conserve resources more, but in the very act of doing so it reduces investment outlets and hence tends to be a depressive influence on the rate of growth. This sort of paradox will crop up time and time again in these pages: it is a contradiction, not of logic, but of capitalism. Perhaps this is best illustrated from Schumpeter's argument. On the one hand, he alleges, critics of monopoly accuse it of sabotaging economic progress; on the other hand, they accuse technical progress of creating "technological unemployment." He says: "Incidentally, it should be noticed that the kind of restrictive practice under discussion, granted that it exists to a significant extent, would not be without compensatory effects on social welfare. In fact, the same critics who talk about sabotage of progress at the same time emphasize the social losses incident to the pace of capitalist progress, particularly the unemployment which that pace entails and which slower advance might mitigate to some extent. Well, is technological progress too quick or too slow for them? They had better make up
The answer is, of course, that technical progress may be both too quick and too slow. It may be too quick in the sense that it wastes economic resources which have a social usefulness. At the same time, it may be too slow in that it does not provide sufficient investment opportunities to fully employ available productive resources. There is no logical contradiction here!

Finally, there is one further point about the effect of monopoly upon investment in innovations. Insofar as the growth of monopoly is accompanied by a tendency for the entrepreneurial function to pass from the hands of industrialists into the hands of financiers, i.e. insofar as an era of monopoly also becomes an era of Finanzkapital, the incentive to invest in risky innovations is likely to be inhibited. The fear of loss is likely to bear more heavily with financiers and bankers than the prospect of super-profits. The safe and present monopoly is likely to be preferred to the risky potential one. And once the new venture is forsaken, a good deal of energy and resources are likely to go into the prevention of interlopers entering the field.

5. Capitalism, Socialism and Democracy, p. 97n.
2. The Widening of Capital

Over the last twenty years a great deal of attention has been directed to the effects of oligopoly upon price formation. The significance of the problem arises out of the fact that, generally speaking, oligopolists will be deterred from pursuing a policy of price-competition because they are acutely aware that price moves by themselves are likely to be matched by rivals, thus bringing to nullity the benefits which it was hoped to gain from the reduction in price. It is surprising therefore that so little attention has been given the problem of analogous attitudes of monopolists and oligopolists towards investment decisions.

The situation in respect of investment decisions is in many ways similar to that in respect of price-output decisions, but is probably of much greater significance for the economy as a whole. Monopoly and oligopoly become significant in the investment process because of the two-edged nature of investment. On the one hand, investment immediately increases income. On the other, it "permanently" increases productive capacity. There are two aspects to the increase in income. First, it is temporary: unless the investment
is followed by further doses, income will drop back to its previous level. Secondly, the income effect is diffused throughout the economy. A given piece of investment means only a very small increment of income to the particular industry making the investment. But whereas the income effect of investment is temporary and diffused, the increase in capacity resulting from the investment is more or less permanent and is specific to the particular industry undertaking it. This is one of the large paradoxes of the capitalist process.

It is true, of course, that the income effect is diffused and temporary for competitive firms and industries, as it is for monopolies. But under competition the capacity effect is also diffused! The addition to total industry capacity which will result from the investment plans of any one firm or potential firm will not be significantly great. Therefore, if all firms act on this assumption and proceed with their investment plans on this basis, the resulting expansion of income may well be sufficient to justify the investment outlays. However, even if the worst comes to the worst and it turns out ex post that the expansion of income has not been sufficient to employ total capacity—both old and new—the expanding firm has the knowledge that a good deal of
the capital loss involved will fall upon competitors. Indeed, it will usually be marginal firms which will take the brunt of the capital loss; and these firms are unlikely to be the ones which are making additions to capacity. In short, under competitive conditions, if capital losses are to result from investment, they can, from the point of view of the individual firm contemplating the investment, be inflicted upon other firms.

But under conditions of monopoly or oligopoly the position is quite different. By contrast, additions to capacity will loom large compared with the resulting increment of income. Redundant equipment will be left on the firms' very own doorsteps; and it will be very difficult to pass the baby onto rivals. In Chapter VII, I demonstrated that the barriers to the exit of marginal firms are of exactly the same order as the barrier to the entry of new firms – including the economies of scale enjoyed by the established firms. Moreover, the greater the barrier to new firms, the greater the degree of monopoly. Hence, the greater the degree of monopoly, the less prospect there is for new firms to inflict capital losses on existing firms and for existing firms to inflict them on their rivals. In greater or lesser degree, these losses, if they
eventuate, must be borne by existing firms themselves.

Domar has given an excellent hypothetical illustration of how this is likely to work out in practice. "We may imagine," he says, "representatives of our big business, perhaps together with those from big labor, sitting around a table. All prospective investments are placed on a platter and passed around. The representative of General Motors might, for example, find all of them splendid, all with the exception of Kaiser-Frazer's new plan for expansion. He takes this one project off and passes the platter on. Pennsylvania Railroad has no objection to Kaiser-Frazer and to all other such projects, but it might decide to veto the St. Lawrence waterway. Its views are seconded by United Mine Workers whose behaviour is undistinguishable from that of their business associates at the table. The platter moves on. Real estate interests think that it was a great pity that Kaiser-Frazer and the St. Lawrence were taken off the platter; those were fine projects to achieve prosperity and raise the American standard of living. The construction of a plant for prefabricated houses, however, is a different matter. That will destroy existing real estate values. And so it goes. By the time the platter completes its trip around the table
it is perfectly empty; all projects have been vetoed by one or other participant. And there goes our prosperity."^6

In another context, Domar has given a real example of this process at work. He points out that:

"The recent publication by the American Iron and Steel Institute, Background Memoranda on Steel Capacity, is an excellent demonstration of this point. The purpose of this memorandum is to show that the country possesses sufficient steel capacity, both relative to the peak (peacetime) year of 1929, and to the peak demands of the most important users of steel taken individually. And the only growth admitted into these prognostications was the growth of population. If a sufficient number of our industries make their plans along these lines, we will end up with some fifteen or more million unemployed."^(7)

The reason for the attitude of United States Steel is made quite explicit in a statement before the Special Committee to Study Problems of American Small Business by Walter Reuther, President of the United Automobile Workers of America, to the effect that representatives of the steel industry believed


that "their chances of achieving full employment were too risky to justify the capital investment in steel-making which full employment will require."\(^8\)

This fear of productive capacity is one of the most striking features of latter-day capitalism. Although throughout the 1930's American business was unable to find any substantial avenues for profitable expansion, American productive capacity was roughly doubled in the 5 years of World War II, when the growth of income was assured from outside the system, as it were.

3. The "Anarchy" of Competitive Production

We have just seen how the diffusion of the capacity effect of investment under conditions of competition is more conducive to investment than the more concentrated capacity effect which we get under conditions of monopoly and oligopoly. But even if monopolies went ahead with their investment plans, ignoring the possibility of income growth falling short of that required to profitably employ the newly-created capacity, their rate of expansion of investment would

---

probably fall short of that of competitive firms in otherwise similar conditions. This is due to the fact that under a competitive regime investment decisions are also diffused. This diffusion of investment decisions make it almost certain that competitive firms will overshoot the investment target for any envisaged rate of expansion. This characteristic of a competitive economy has been called by Marx the "anarchy of capitalist production," a term which brings out an aspect of competitive production which is recognized, but not stressed, by academic economists.

Under conditions of competitive production, the coordination of production plans is achieved by the market mechanism *ex post*, as contrasted with the *ex ante* coordination which would be possible under a planned economy. Consider an atomistic industry faced with an expanding demand. Demand begins to press against capacity; and each firm has the incentive to expand its capacity. This is, indeed, the indication of the market. Moreover, new firms are likely to be attracted into the industry, as price tends to rise to ration sales.

But a considerable period must elapse between
the decision to expand capacity and the bringing of the final product to market. Under competitive conditions there is no mechanism by which the total planned expansion of capacity may be equilibrated with the long-period, say full employment, requirements of the industry. At least there is no mechanism ex ante. Thus, under competitive conditions the aggregate of all independently planned additions to capacity will often be greater than that required on any long-period calculation. Indeed, this over-shooting of the target is almost inevitable for competitive industries facing a rapidly expanding demand, as for example during a strong upswing of the trade cycle. The fact that over-provision of capacity is being made does not become apparent until the final goods begin to saturate the market. By this time, further investment is probably under way. It is at this point that the stock market begins to panic.

It should be stressed that we are concerned here, not with excessive provision of capacity which results from a failure of Effective Demand, but with provision of capacity over and above what would be required at, say, a full employment level of output. John Strachey cites an example of this from a New Deal investigation into the
boot and shoe industry in the United States. It was found that, as a result of the expansion of capacity of the 'twenties which continued into the early 'thirties, capacity in that industry was three times that required upon the basis of need, abstracting altogether from effective demand. This is probably an extreme case; but it is certain that competitive industries do tend to over-provide capacity in this way, in greater or lesser degree. Necessarily, a substantial proportion of capacity which is created from time to time will run to waste.

From our point of view, there are two aspects to this matter. From the point of view of the best use of resources, once again we find that monopoly is probably superior to competition. Monopolists and oligopolists will have a much better idea of total industry requirements and of the investment plans which are under way at any given time. They can "plan" ahead more effectively. Also, their industries are not so much at the mercy of "outsiders". As a result they can take a broad, "planning" view of investment. And monopolists and their spokesmen have not been slow to point this out. Indeed, a sort of

Guild Capitalism under the hegemony of monopolies has often been advanced as an alternative to State Socialism. Certainly, the notion was basic to the Corporate State ideology of Fascism.

On the other hand, it is clear that too much investment is preferable, from the point of view of the growth of the economy, to too little. Although some of this investment will run to waste, it will serve to lift the economy to a level to which it might not otherwise have attained. In this respect, over-investment is in the same category as pyramid-building, gold-mining and digging holes. Via the ratchet effect upon consumption, such over-investment may well result in a permanently higher level of activity than would have otherwise obtained.

4. Rationalization of Capacity Requirements

Another factor which tends to reduce the amount of capital investment in an industry as the degree of monopoly rises is the reduction in excess capacity which often occurs as an industry becomes more integrated. Here we are not comparing monopoly with an idealized perfect competition in which excess capacity is impossible, but rather with the more usual (outside of certain standardized primary products) form of
imperfect competition in which there may be a large number of small firms—e.g., job printing—but where the market is far from perfect in the technical sense.

In an imperfect market where consumer attachment is important, firms will carry a reserve of capacity with which to meet unexpected increases in demand or to expand output permanently if the opportunity presents itself. Moreover, most firms produce a number of products or perform a number of different but related services. Equipment must be provided to produce each product or service, even where the particular demand may be well below technical capacity. It is easy to think of examples. Every job printer must possess a linotype, every small engineer a certain range of lathes, milling machines, and so on; yet each of these machines may be required for only a fraction of total working time.

Clearly, then, as small independent producers are merged into larger units, the total amount of excess capacity required in the industry is substantially reduced. Demands upon individual firms will be higher and also more stable—in virtue of the law of large numbers. Consequently,
a closer approximation to full running may be achieved. An excellent example of this trend is given by Austin Robinson in his book, *Monopoly*. Judge Gary found that, when the United States Steel Corporation was formed, 50 per cent less capital was required than was carried by the merging firms when working as independents. This is probably an exceptional case, but it does illustrate an important tendency in the latter stages of the growth of monopoly. The significance of the reduction in capital requirements as a result of merger and amalgamation is also attested by the clamour by United Kingdom industrialists after World War I for the rationalization of industry. A nearer approach to full running and consequent reduction in capital costs are not the least important gains which would follow from "rationalization."

To give a complete picture, it should be stressed that this tendency is only likely to operate in the later stages of monopoly growth. In the formative stages the reverse may hold true; the earlier stages of monopolistic strategies may involve increases in excess capacity. In this respect, the position is analogous to that of oligopolistic price formation.

In the formative period, when firms are struggling for position and expansion, price wars and cut-throat competition may well result in lower prices. Later, when the successful firms become entrenched, higher prices are almost certain to be levied. Actually, the connection is rather more than merely analogous, for excess capacity is an essential element of price warfare.

The great periods of capitalist expansion have almost always been associated with investment in heavily capital-using development. Anything which tends to reduce the capital requirements of industry generally must tend to damp down capitalism's inherent ability to expand.

5. The Growth of the Distribution Sector

During the last 20 years or so, economists have come to recognize the connection between monopoly and the growth of selling costs, advertising, the duplication of distributive outlets, and so on. The reason for this correlation between monopoly and distributive costs is not far to seek. Under monopoly conditions, an expansion of output by a firm can, in general, only be achieved by a substantial reduction in price. Moreover, in the more general
case of oligopoly, even this avenue of expansion is limited, since a price-cut is almost certain to be followed by rivals, with whom any expansion of demand has got to be shared. Under these conditions, price-competition tends to be superseded by methods of non-price-competition. A huge superstructure of selling expenditure, advertising, one-brand distributive outlets, and the like is erected above the productive structure. Of course, the purpose is not only to provide an offensive weapon to replace price-competition. Equally it has a defensive significance in the protection of established positions from outside competition.

An idea of the importance of these trends in quantitative terms has been given by the Twentieth Century Fund on the basis of its report, Does Distribution Cost Too Much? The Fund has pointed out that:

"Distribution – not production – is now the great frontier of the American business system. Distribution takes 59 cents of the consumer’s dollar as compared with only 41 cents for production processes. Workers in distribution increased nine times between 1870 and 1930, while the population increased only three times." (11)

Of course, this trend cannot be wholly ascribed to monopolistic practices; but assuredly a large part of it can be. Certainly it represents an enormous change in the structure of the economy.

The important point of this for our discussion is that distribution and service industries usually require a much lower rate of investment in capital equipment – with the possible exception of transportation – than do industries engaged in production. Hence, the larger the sector of the economy which is represented by this type of business activity, the less the demand for investment funds. Certainly distributive industries are likely to provide much less scope for investment in technical innovation than productive industries.

This conclusion goes in the opposite direction to that of Paul Sweezy, who suggests that the growth of the distribution sector offers an additional outlet for investment. This may be so in the formative stages; but, since there is a limit to the size which this sector may reach in relation to production, it must sooner or later constitute a dampening factor on the economy as a whole. Consider an economy which requires a net investment of 10 per cent of income to maintain a steady rate of growth at full employment.

If there is a large sector of such economy, for example the distributive sector, which requires only 5 per cent net capital formation in relation to income to sustain an equilibrium growth consistent with the overall rate of growth, then this would mean that a rate higher than 10 per cent must be sustained in the productive sectors, if the overall investment requirement is to be forthcoming.

6. The Distribution of Income

Monopoly will tend to skew the distribution of income in favour of the monopolists. Insofar as this raises the propensity to save, this will increase the rate of investment required to maintain any pre-determined rate of growth. However, we are not concerned here with this aspect of the matter. Rather we are concerned how changes in the distribution of income resulting from monopoly directly affect the inducement to invest. There seem to be two ways in which such changes may affect the rate of investment.

Firstly, insofar as the gain of income by the monopoly sector is at the expense of consumers (particularly wage and salary earners), this will tend to depress the rate of growth of consumption and hence the rate of investment, via the acceleration principle.
Secondly, the skewing of income distribution in favour of the monopoly sector may be at the expense of other business units, in particular the smaller and more competitive businesses. This may react unfavourably upon investment in three ways:

(1) profits which are more or less permanently low or losses which are protracted will tend to reduce the enthusiasm for investment of those firms suffering them;

(2) reduction in the share of income accruing to the non-monopolistic business sector will reduce internal accumulation in that sector and hence reduce the funds available in that sector for investment purposes; and

(3) this reduction in internal accumulation will, in itself, make it more difficult for businesses to borrow from outside the sector for investment purposes.

Steindl has worked over these aspects of the inducement to invest at some length; and they play a basic part in his model of the investment process.13

While one feels that he makes them carry a weight which is far too heavy for them alone, these elements are bound to play a not unimportant part in a more general and comprehensive theory of investment and of growth.

7. Conclusion

In these pages I have attempted systematically to set out the main ways in which monopoly may affect the inducement to invest. A good deal of further theoretical and empirical investigation is clearly yet to be done, before really firm answers may be expected. In particular, we need a general theory of investment. Investment must be brought into the theoretical system, not introduced from outside as data.

Even at this stage, however, certain presumptions seem apparent, if the foregoing argument is valid. Firstly, monopoly is probably superior to competition so far as the best direction of a given use of resources is concerned. Secondly, it would appear on a number of important counts to be inferior to competition in providing, of its own accord, the necessary investment to maintain a high rate of growth.
In particular, it seems to be lacking in inherent ability to provide investment outlets sufficient to lift the economy from a low level of activity. This indeed seems to be the most striking feature of the 1930's. It required preparation for war, and finally war itself, to lift the world economy from the doldrums of the 'thirties.

We may feel that the problem of internal income generation is no longer a serious one, and applaud monopoly for its seeming superiority in the use of resources. We may feel that Governments are now able to guarantee income generation, from outside the system as it were. However, it should not be easily forgotten that, historically, the only periods of sustained full employment or near full employment under capitalism have been associated with high levels of armaments expenditure. These have the inestimable advantage from the point of view of business interests that (a) they cannot be over-produced, and (b) they are a form of investment which does not come into competition with existing capital. On both counts they can obtain the unqualified approval of the capitalist class and, in particular, the monopoly interests among it who are usually most strongly entrenched in the basic industries which
benefit most directly from armaments building.

Whether such approval would be forthcoming for alternative forms of Government expenditure on the required scale, e.g. in housing, schools, hospitals and direct effort to raise the living standards of the under-privileged peoples the world over, is yet to be demonstrated.
BIBLIOGRAPHY
(SELECT)

PART I

(a) Books


JOAN ROBINSON: Economics of Imperfect Competition, London, 1933.


* Because of the immense literature on the subject of Part I and the availability of published bibliographies (e.g. in the later editions of Chamberlin's Monopolistic Competition), references for Part I are confined to those actually cited in the text.
(b) Articles and Reviews


Reply to Kaldor's "Professor Chamberlin on Monopolistic and Imperfect Competition", Quarterly Journal of Economics, May, 1938.


N. KALDOR: "Mrs. Robinson's 'Economics of Imperfect Competition', Economica, N.S., August, 1934.


"Professor Chamberlin on Monopolistic and Imperfect Competition", Quarterly Journal of Economics, May, 1938.


PART II

(a) Books

J.K. GALBRAITH: American Capitalism, Boston, 1953.


(b) Articles, Papers and Reviews


(vi)


(a) Statistical Sources


The Australian Tariff: An Economic Enquiry, Melbourne, 1929.


(vii)

PART III

(a) Books


(b) Articles, Papers and Reviews


W.P. REUTHER: "One Hundred Million Tons of Steel", Statement to Hearings of the Steel Sub-committee of the Special Committee to Study the Problems of American Small Business, U.S. Senate, July 21, 1947.