MAKING MONOPOLISTIC COMPETITION MORE USEFUL

Kiminori Matsuyama

Working Papers in Economics E-92-18

The Hoover Institution
Stanford University
August 1992

This paper was written while the author was a National Fellow of the Hoover Institution at Stanford, whose hospitality he gratefully acknowledges.

Hoover Institution
Working Paper No. E-92-18
August 1992

MAKING MONOPOLISTIC COMPETITION MORE USEFUL

Abstract

The monopolistic competition literature exclusively focuses on the group equilibrium where firms in the group supply products that are close substitutes. While faithful to Chamberlin, this restricts the ability of the monopolistic competition paradigm to yield the predictions that are significantly different from those of perfect competition. This paper demonstrates that the case where firms supply complementary products naturally arises in the standard models of monopolistic competition and explores its implications. It is shown that the presence of positive feedback effects leads to agglomeration phenomena, which are discussed in the context of urban and interregional economics, as well as economic development and technology choices.

Keywords: The Chamberlinian Monopolistic Competition, Complements and Substitutes, the Dixit-Stiglitz (AER, 1977) model, Multiple Equilibria, Positive Feedback

JEL Classification Numbers: D40, L10

Kiminori Matsuyama
Department of Economics
Northwestern University

1. Introduction

The publication of Edward Hastings Chamberlin's book, The Theory of Monopolistic Competition (1933) was widely heralded in its time as the beginning of a revolution that would replace the classical paradigm of perfect competition (Samuelson 1967). Six decades later, however, the impact of the theory remains small outside of few areas, such as international trade (Helpman and Krugman 1985) and the growth theory (Romer 1991; Grossman and Helpman 1992). The position of Chamberlinian monopolistic competition in general economic theory, while secured and respectful, has not quite matched the expectations generated earlier. It was a revolution that failed.

One reason for this failure is that his theory is conceptually tied to the Marshallian partial equilibrium framework. The notion of an industry, or a "group" to use the Chamberlinian terminology, plays the central role in his theory. It is the process of entry to and exit from the group that distinguishes his theory from the classical theory of monopoly. The grouping of firms also provided him with a justification to analyze the interdependence of firms within the group in isolation, without worrying about any repercussion effect from the rest of the economy. Many economists approach the model of Chamberlin in a partial equilibrium perspective; that is, they treat it as a model of an industry. Chamberlin himself motivated his new theory as an attempt to bridge the gap between the two extreme modes of the partial equilibrium analysis available at the time, the theory of monopoly and the theory of a competitive industry. The Chamberlinian monopolistic competition is typically considered, and in fact taught, as a part of industrial organization.

The problem is that the notion of Marshallian industries, or any classification of firms into groups, while unequivocal in the context of homogeneous goods, loses its definiteness once we start thinking about product

differentiation and the heterogeneity of goods sold by different firms. Chamberlin and his followers typically define the group to be a collection of firms supplying products that are sufficiently close substitutes for each other. Aside from being somewhat arbitrary, this grouping criterion substantially restricts the ability of the monopolistic competition paradigm to yield predictions that are difficult to come by in the classical paradigm of perfect competition. After all, the assumptions of the homogeneity of goods and perfect competition are merely theoretical abstractions, meant to capture the environment in which many firms compete with similar goods and hence face very elastic demand curves. This point was clearly made by Stigler (1949, p.24), who wrote "I personally think that the predictions of this standard model of monopolistic competition differ only in unimportant respects from those of the theory of competition because the underlying conditions will usually be accompanied by very high demand elasticities for the individual firms." Of course, many economists continue to find some features of the theory of Chamberlin attractive. They include the explicit treatment of the price setting behaviors of firms rather than assuming the price taking behaviors and the fictitious auctioneer, and its ability to explain why firms operate at the downward sloping part of its average cost curve. Even these achievements, however, would fail to impress those who read Milton Friedman's critique of monopolistic competition in his Methodology of Positive Economics (1953) and believe that the descriptive accuracy of a theory has nothing to do with its analytical relevance. I think that it is fair to say that most practical economists regard the Chamberlinian theory of monopolistic competition as a mere mixture of the monopoly and perfect competition. They tend to believe that the theory of perfect competition supplemented by a discriminating use of the theory of monopoly would be

sufficient for practical purposes, and that the theory of Chamberlin is an unnecessary complication, which brings very little that is new. Those who study industrial organization find his theory rather boring because it lacks all exciting elements of strategic interactions of Cournot and Bertrand.[1] And it is quite ironic that mathematical economists have recently used monopolistic competition to test the robustness of perfect competition, the mode of analysis that the theory of monopolistic competition was invented to replace.[2]

With a couple of reformulations, however, the theory of monopolistic competition could be made more useful (that is, yielding predictions that are difficult to come by in the classical paradigm) and hence more exciting. First, the theory should be recast in the framework of general equilibrium. As argued by Triffin (1940), once the focus of analysis shifts from the equilibrium of supply and demand within an industry to the interdependence between firms producing different goods, there is no compelling reason why attention should be limited to a prespecified set of firms.[3] At least in principle, we need to deal with the interdependence between all firms in the economy. In other words, the theory can be reformulated to analyze the interaction between different groups of firms. Second, the grouping of firms (and hence products) could be based on factors other than substitutability, such as technical similarity and geographical proximity. What is an appropriate grouping criterion, of course,

---

[1]They certainly like the spatial model of competition better because of its explicit treatment of product choices and localized interactions.

[2]Hart (1979, 1985), Roberts (1980), Perloff and Salop (1985) and Jones (1987); see Beath and Katsoulacos (1991, Chapter 7) for a survey.

[3]Chamberlin's response to this criticism is equivocal, to say the least. He wrote, "The upshot of the matter seems to be that the concept is not very useful and is even seriously misleading in connection with monopolistic competition (1962, p.201)," despite that he kept the concept of a group central to the analysis.

depends on the nature of questions asked, but such a reformulation gives rise to the possibility of grouping complementary products together. This is important because the entry-exit process would have very different consequences when the group consists of complementary products.[4] Entry of a new firm into a group will increase the profits of other firms in the group, which offer complementary products. This in turn attracts more firms into the group. Analyzing the interdependence of different groups in the presence of such a positive feedback can throw new lights on the formation of the industrial structure of the economy.

And this is precisely what I attempt to do in this paper. Through a series of models, it is demonstrated that the theory of monopolistic competition, once reformulated this way, could be a useful apparatus for explaining agglomeration phenomena in a variety of contexts, such as urban economics (e.g., retail store clustering), interregional economics (e.g., polarization and industrial localization), development economics (e.g., underdevelopment traps), and technological choices (e.g, standardization). Indeed, these agglomeration phenomena have been the subjects of intensive research in recent years.[5] But the classical paradigm of perfect competition has difficulty of coming up with satisfactory explanations of why such an agglomeration may occur. The existing

---

[4]Chamberlin himself argued in later versions that "The "common sense" definitions of industries in terms of which practical problems are likely to be studied seem to be based much more upon technological criteria than upon the possibility of market substitution (1962, p.202n)." Nevertheless, he continued to assume in the formal analysis that products in the group are close substitutes.

[5]For retail clustering, see Brown (1989) for a survey. For other topics, a very partial list of literature includes Arthur (1990b), Faini (1984), Krugman (1987), Lucas (1988), Matsuyama (1992a) and Young (1991) for polarization and industrial localization; Azariadis and Drazen (1990) and Matsuyama (1991) for underdevelopment traps; Arthur (1989), Farrell and Saloner (1985, 1986), Katz and Shapiro (1985,1986) for technology choices. See Arthur (1988, 1990a) for a general overview of the economics of agglomeration.

studies, which maintain the perfect competition assumption, simply assume some sorts of technological externalities (e.g., the benefit of adopting an industrial standard is an increasing function of the extent to which the standard is adopted) and analyze the implications. What exactly constitute such externalities is not a question asked frequently, nor modelled carefully. The assumed externalities are often defended by appealing to the observed patterns of agglomeration, the phenomena that the model is supposed to explain. The analysis presented in this paper has the advantage of explaining when market mechanisms generate agglomeration phenomena and when they do not, without assuming the presence of technological externalities. Arguably, this demonstration proves that the theory of monopolistic competition can meet the challenge presented by its opponents who believe "[t]he sole test of the usefulness of an economic theory is the concordance between its predictions and the observable course of events (Stigler 1949, p.23)."

There already exists the voluminous literature on general equilibrium models of monopolistic competition, including Dixit and Stiglitz (1977) upon which the following analysis are built. Dixit and Stiglitz were aware of the possibility that entry of a firm may increase the profit of an incumbent firm in their model, but dismissed it as "implausible." The restriction they imposed effectively implies that the intragroup substitution dominates the intergroup one, thereby downplaying the interdependence across groups. While faithful to Chamberlin, this restriction turns their model effectively into the partial equilibrium one. This practice has been followed by those who later applied their model, often without any discussion.[6] What I will argue below is that,

---

[6]The paper by Dixit and Norman (1980) is one of the few articles, which discusses this assumption, by arguing that differentiated goods need to be "good enough substitutes to warrant the label "product group"(p.282)."

under certain situations, it is plausible to assume that the intergroup substitution be larger than the intragroup one. I will then show that this alternative assumption implies that products in the same group are complementary for each other, leading to a variety of agglomeration phenomena.

The rest of the paper is organized into four parts. In the next two sections, I demonstrate how the case in which the firms in the group supply complementary products arise naturally. In the model of section 2, all products are differentiated, but they are classified into two groups. Section 3 considers the original Dixit and Stiglitz model, in which there exist one group of differentiated goods and the single homogeneous good. This section also provides a detailed tax analysis, which may be of independent interest. In section 4, I extend the model of section 2, so that a product could be a substitute to some products and complementary to others at the same time. Finally, I present some concluding remarks in section 5.

2. The Basic Model

As the starting point, let us consider the following variant of the Dixit-Stiglitz (1977) model. The goods produced in this economy are divided into two groups, $i = 1$ and 2, and within each group there exists a continuum of differentiated products. Each product is supplied solely by a specialist firm. Let n_i be the number of firms (and products) in group i. Following the standard practice in the monopolistic competition literature, I use the dichotomy between the shorn-run, in which the numbers of firms are given, and the long run, in which they are determined endogenously through the entry-exit process.

2.A. The Short Run Analysis.

Let $x_i(z)$ denote consumption of product z of group i. The "representative agent" has the identical preferences, given by the following two-tier utility function, $V(X_1,X_2)$, where

$$X_i = \left[\int_0^{n_i} [x_i(z)]^{1-\frac{1}{\sigma}} dz\right]^{\frac{\sigma}{\sigma-1}}, \qquad \sigma > 1 . \qquad (i = 1, 2) \tag{1}$$

It is assumed that the upper-tier utility function V is homothetic. The sub-utility function, X_i (i = 1, 2), represents a composite of differentiated products of group i. This specification of the composite is standard. It assumes that, within each group, all products enter symmetrically and the direct partial elasticity of substitution between every pair is equal to σ. The restriction $\sigma > 1$ implies that $u(x) = x^{1-1/\sigma}$ satisfies $u(0) = 0$ and $u(x) > 0$ for any $x > 0$; no product is thus essential and the agents have preferences for variety. Let $p_i(z)$ be the price of product z of group i. (Throughout the paper, the labor is taken as the numeraire.) The demand function for each product satisfies

$$x_i(z) = \left[\frac{p_i(z)}{P_i}\right]^{-\sigma} X_i, \qquad (i = 1, 2) \tag{2}$$

and

$$\frac{X_1}{X_2} = \Phi\left(\frac{P_1}{P_2}\right), \tag{3}$$

where

$$P_i = \left[\int_0^{n_i} [p_i(z)]^{1-\sigma} dz\right]^{\frac{1}{1-\sigma}}, \tag{4}$$

can be interpreted as the price index of composite i. Function Φ is a decreasing

function, defined implicitly by $V_1(X_1,X_2)/V_2(X_1,X_2) = P_1/P_2$, where the homotheticity of V implies that the left hand side solely depends on the ratio X_1/X_2. The elasticity of substitution between X_1 and X_2 is defined as

$$\epsilon\left(\frac{P_1}{P_2}\right) \equiv - \frac{d\log\Phi(P_1/P_2)}{d\log(P_1/P_2)} > 0 .$$

The assumption of homothetic preferences is a convenient one, as the relative demand for each product is independent of the aggregate income, $I = P_1X_1 + P_2X_2$, as well as of its distribution across agents.

Each product is supplied by a single, atomistic firm. Producing one unit of output in group i requires a_i units of labor. Taking the demand function given, the firm chooses its price $p_i(z)$ to maximize its profit. In doing so, it treats P_i and X_i (i = 1, 2) as fixed parameters; although this firm has some monopoly power over its own product market, it is negligible relative to the aggregate economy. Because every firm faces a downward sloping demand with the constant price elasticity σ,

$$p_i(z)\left(1 - \frac{1}{\sigma}\right) = a_i ,$$

for all i and z. Thus, all firms in the same group set the same price and produce the equal amount. Dropping the index z,

$$X_i = n_i^{\frac{\sigma}{\sigma-1}} x_i , \qquad (1')$$

and

$$P_i = n_i^{\frac{1}{1-\sigma}} p_i . \qquad (4')$$

Furthermore, $p_1/p_2 = a_1/a_2$ and hence the relative price index depends solely and inversely on the relative number of firms (and products);

$$\frac{P_1}{P_2} = \left[\frac{a_1}{a_2}\right]\left[\frac{n_1}{n_2}\right]^{\frac{1}{1-\sigma}} . \qquad (5)$$

This is to say that, because the agents have taste for diversity, an expanding range of products in one group makes that group as a whole more attractive than the other. Finally, let L be the labor endowment of this economy. Then, the labor market clearing condition is given by, from (1'),

$$L = n_1 a_1 x_1 + n_2 a_2 x_2 = n_1^{\frac{1}{1-\sigma}} a_1 X_1 + n_2^{\frac{1}{1-\sigma}} a_2 X_2 . \qquad (6)$$

As shown, the labor requirement for providing X_i declines with n_i, which is nothing but the mirror image of the taste for diversity embedded in the preferences.

It is easy to see that, for any distribution of the firms across the two groups, eqs. (3), (5), and (6) can be solved for the unique short-run equilibrium. The profit of each firm in equilibrium can be calculated as

$$\pi_i = (p_i - a_i) x_i = \frac{p_i x_i}{\sigma} = \frac{P_i X_i}{\sigma n_i} . \qquad (7)$$

Note that, from (6) and (7), one can get the national income account, $I = P_1X_1 + P_2X_2 = L + n_1\pi_1 + n_2\pi_2$, which is nothing but the Walras's Law.

2.B. The Long Run Analysis.

The number of products in each group is determined through the entry-exit process. Imagine that there is a continuum of entrepreneurs with unit mass in this economy. Each firm needs to be set up and managed by an entrepreneur. Hence, n_i can also be interpreted as the number of entrepreneurs who supply products in group i, and $n_1 + n_2 = 1$. Entrepreneurs chooses the group that earns higher profits. From (3), (5), and (7), the relative profit depends on the

relative number of firms,

$$\frac{\pi_1}{\pi_2} = \left[\frac{n_2}{n_1}\right]\left[\frac{P_1}{P_2}\right]\left[\frac{X_1}{X_2}\right] = \left[\frac{a_1}{a_2}\right]^{1-\sigma}\left[\frac{P_1}{P_2}\right]^{\sigma}\Phi\left(\frac{P_1}{P_2}\right)$$
$$\equiv \psi\left(\frac{P_1}{P_2}\right) = \psi\left(\left[\frac{a_1}{a_2}\right]\left[\frac{n_1}{n_2}\right]^{\frac{1}{1-\sigma}}\right) \equiv \Psi\left(\frac{n_1}{n_2}\right). \qquad (8)$$

The long run equilibrium is given by a distribution of firms across the two groups, in which no entrepreneur gains from reallocating.

As it stands, the model constructed above may have multiple equilibria. To see this, note that (8) implies that $0 < n_1 = 1 - n_2 < 1$, is a long run equilibrium (that is, $\pi_1 = \pi_2$) if the associated relative price P_1/P_2 satisfies

$$\left[\frac{P_1}{P_2}\right]^{-\sigma} = \left[\frac{a_1}{a_2}\right]^{1-\sigma}\Phi\left(\frac{P_1}{P_2}\right).$$

Since the right hand side could be any positive-valued, decreasing function, an arbitrary number of solutions for this equation may exist.

The Case of Substitutes: $\sigma > \epsilon(\bullet)$.

One way of dealing with this problem is to impose further restrictions on the upper-tier function V or equivalently on Φ. If the two groups in the model are to be interpreted as the two industries or sectors, then it is plausible to assume $\sigma > \epsilon(\bullet)$. For example, suppose that group 1 represents restaurants and group 2 a variety of retail stores. Then, it makes sense to assume that a chinese restaurant is a closer substitute to italian or thai restaurants than to video shops and bookstores, and that compact discs are close substitutes to books and videos than to any ethic foods. Under this assumption, $\sigma > \epsilon(\bullet)$, the ratio of the profit per firm in the two industries,

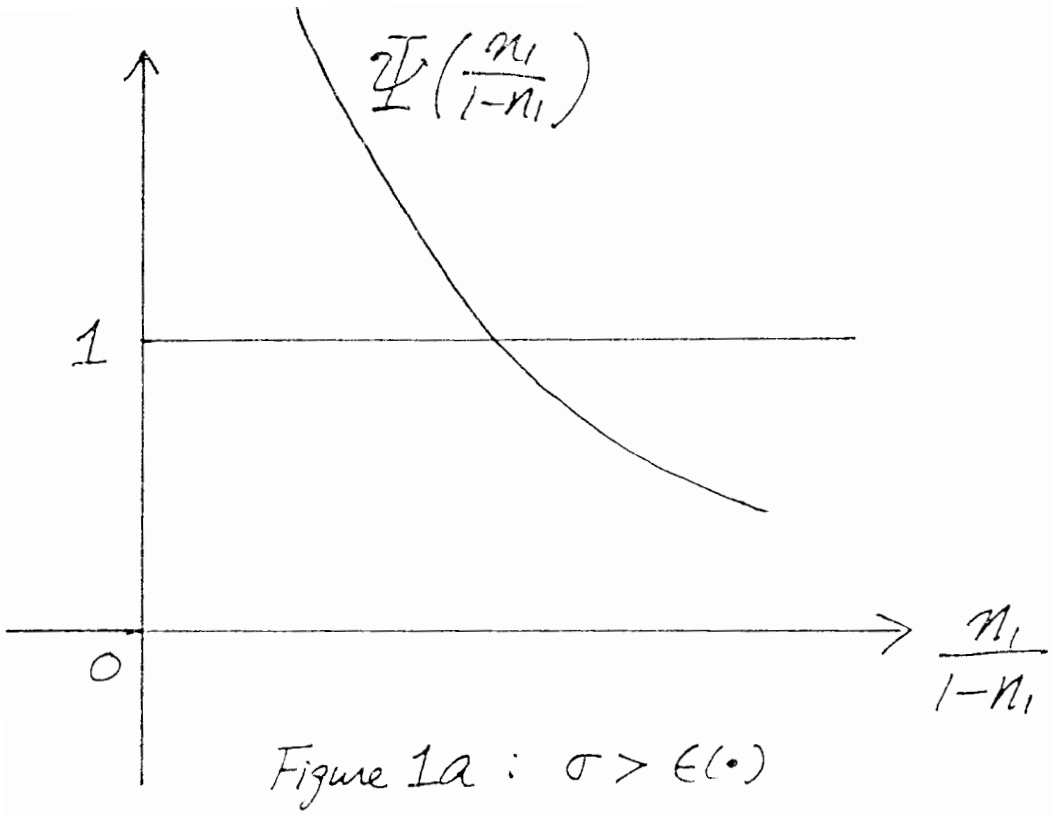

Figure 1a : $\sigma > \epsilon(\cdot)$

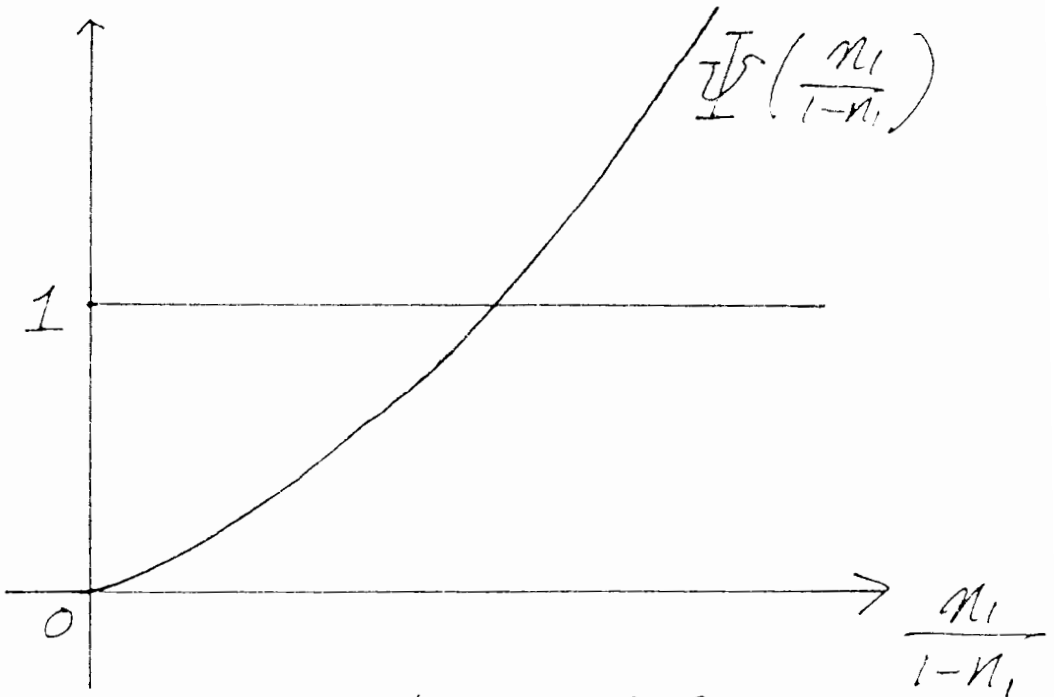

Figure 1b : $\sigma < \epsilon(\cdot)$

$$\Psi\left(\frac{P_1}{P_2}\right) = \left[\frac{a_1}{a_2}\right]^{1-\sigma}\left[\frac{P_1}{P_2}\right]^{\sigma}\Phi\left(\frac{P_1}{P_2}\right) = \left[\frac{a_1}{a_2}\right]^{1-\sigma}\Phi(1)\exp\left[\int_1^{P_1/P_2}\frac{\sigma-\epsilon(v)}{v}dv\right], \tag{9}$$

is an increasing function of the ratio of the price indices. Hence, from (5), $\Psi(\bullet)$ is decreasing, and $\Psi(+0) = \infty$ and $\Psi(\infty) = 0$. This is to say that the incentive to enter an industry declines as the number of firms in that industry increases. Any additional firm in an industry reduces the demand for any competing product, given the expenditure share of the industry. Although an expanding variety of the products may increase the expenditure share of the industry, the former effect dominates, as long as the products in the same industry are closer substitutes than those which belong to different industries. As a result, there is a unique equilibrium in the interior, as depicted in Figure 1a.

The equilibrium numbers of firms (and products) may be calculated, once the upper-tier function is specified. For example, if

$$V(X_1, X_2) = \left[\beta_1 X_1^{1-\frac{1}{\epsilon}} + \beta_2 X_2^{1-\frac{1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}},$$

then

$$\frac{n_1}{1-n_1} = \left\{\left[\frac{\beta_1}{\beta_2}\right]^{\epsilon}\left[\frac{a_1}{a_2}\right]^{1-\epsilon}\right\}^{\frac{\sigma-1}{\sigma-\epsilon}}. \tag{10}$$

The equilibrium number depends on the parameters in an intuitive way. More firms enter in industry 1, as the benefit of the goods produced in industry 1 is relatively large and the other industry produces the essential goods at the relatively low production cost (or the other industry produces the non-essential goods at the relatively high production cost).

That assumption that the intergroup elasticity of substitution is smaller

than the intragroup one is also faithful to the original formulation of the Chamberlinian monopolistic competition. In his classic work, Chamberlin was concerned with the industry in which the goods produced are fairly close substitutes. The group he contemplated was "one which has ordinarily been regarded as composing one imperfectly competitive market: a number of automobile manufacturers, of producers of pots and pans, of magazine publishers, or of retail shoe dealers." (1962, p.81; Emphasis in the original). In his analysis of "Group Equilibrium," he introduced the dd curve, the demand curve for each product with the prices of all other products held constant, and the DD curve, the demand curve for each product when all prices in the group move together. He then assumed that the dd curve is more elastic than the DD curve. As is clear from (2) and (3), this assumption corresponds to $\sigma > \epsilon(\bullet)$ in the present model.

It is worth noting that the products in the same group are substitutes in two related, but different senses. One is the standard Hicks-Allen notion of product substitution, thus concerned with the property of demand functions (Hicks and Allen 1934). The demand for each product declines if the prices of other products in the same group are reduced; that is, the Allen partial elasticity of substitution is positive. This is precisely why the dd curve is more elastic than the DD curve. The other is concerned with the property of the relative profit function. Entry of a firm in an industry reduces the incentive for other firms to enter the same industry. Viewed as an entry-exit game played by a continuum of players, the reallocation decisions by entrepreneurs are strategic substitutes, in the sense of Bulow, Geanakoplos, and Klemperer (1985). Because of this negative feedback effect, the assumption $\sigma > \epsilon(\bullet)$ guarantees the unique equilibrium in the interior.

The Case of Complements: $\sigma < \epsilon(\bullet)$.

The alternative assumption, $\sigma < \epsilon(\bullet)$, while unconventional, seems more plausible under some settings, as the following three examples suggest.

1) Shopping districts. Suppose that there are two shopping districts, say East and West, and the entrepreneurs choose the locations of their restaurants and retail stores. Shoppers do not have a clear idea on what they want to eat and to buy until they actually visit a shopping district and observe the choices available to them. Hence, they generally prefer to go to the shopping district with more restaurants and shops, although different shoppers may have different locational preferences.[7] Then, it makes sense to assume that the East and West are much closer substitutes to each other as than a restaurant and a bookstore in the same shopping district are to each other.

2) Regional economies. Suppose that there are two regions in the economy, North and South. The entrepreneurs supply differentiated intermediate inputs or "producer services" to the competitive final goods sector in each region. They provide a wide array of complementary services, ranging from equipment repair and maintenance to delivery and warehouse services, and engineering and legal supports to accounting, advertising, and financial services. These services are nontradeable, and hence the entrepreneurs need to choose where to locate. There are two final goods that are close substitutes to each other. Both regions

[7]Sattinger (1984) and Perloff and Salop (1985) have recently constructed a framework, in which the utility each individual gets from consuming a particular product is a random variable, drawn independently from the same distribution and there is no aggregate uncertainty. They show that this framework could justify the Dixit-Stiglitz type representative agent model of monopolistic competition. Anderson, De Palma, and Thisse (1989) explored the similar approach to investigate the relation between the location model and the representative agent model of product differentiation.

specialize in the final goods production. North produces X_1 by employing differentiated inputs, using the CES technology given in (1). Similarly, South produces X_2. Labor, as well as the final goods, are mobile at zero cost. Let the agents share the identical homothetic preferences over the final goods, given by $V(X_1, X_2)$. Then, this economy is isomorphic to the model presented above, and it is natural to assume $\sigma < \epsilon(\bullet)$. In this context, the value of diversity embedded in (1) can be considered as higher productivity due to increasing specialization. (This interpretation of a CES formulation is due to Ethier (1982) and Romer (1987).)

3) Technology Choices. Finally, imagine that there are two competing technologies or industrial standards. A variety of differentiated products can be produced based on either technology. The two technologies are not compatible with each other, hence any product can be used only in combination of other products in the network, that is, those based on the same technology. The services provided by the two networks are, however, similar except that consumers prefer the network that comes with a larger set of options.

The assumption, $\sigma < \epsilon(\bullet)$, implies that two products in the same group are now complements in the sense of Hicks-Allen. As can be seen from (2) and (3), the demand for a product increases if prices of other products in the same group are reduced; that is, the Allen partial elasticity of substitution is negative.[8] To take the shopping district example, bargain sales in apparel stores bring more customers and hence benefits restaurants in the same shopping district.

---

[8]Note that the two products in the same group, even though their direct partial elasticity of substitution, σ, is positive, becomes complementary in market demand in the presence of the third good that is their close substitutes. See Samuelson (1974) and Ogaki (1990) for the recent treatment on this issue.

Furthermore, the reallocation decisions of entrepreneurs are strategic complements in the sense of Bulow-Geanakoplos-Klemperer. As seen from (8) and (9), the relative profit is now decreasing in the relative price index, and hence, the profit function Ψ, as shown in Figure 1b, is increasing in the relative number of firms and satisfies $\Psi(+0) = 0$ and $\Psi(\infty) = \infty$. Reallocation of a firm from group 2 to 1, by expanding the available variety of group 1 and reducing that of group 2, causes a large shift in demand. This enhances an incentive for other firms to follow. Because of this positive feedback effect, the model now has three equilibria: the two on the boundary, ($n_1 = 0$ and $n_1 = 1$), and the third in the interior, given by the condition, $\Psi(n_1/n_2) = 1$.

The interior equilibrium ought to be ruled out for a couple of reasons. First, the existence of the interior equilibrium is guaranteed due to the assumption that there is a continuum of entrepreneurs. This assumption, while technically convenient, is not a realistic feature of the model. If the integer constraint on the number of entrepreneurs is taken seriously, the solution of $\Psi(n_1/n_2) = 1$ can be achieved only approximately. The positive slope of the profit function, however, implies that, at an approximate solution, all entrepreneurs have an incentive to move into the same group. Hence, equilibria of an economy with a large, but finite number of entrepreneurs do not exist anywhere near the interior equilibrium. In other words, the interior equilibrium of the limit economy could not be a limit point of a sequence of equilibria of large, but finite economies.[9] Second, the interior equilibrium has perverse comparative static properties. As clear from (10), the number of firms in group

---

[9]In the case of substitutes, the integer constraint does not cause a problem. Because of the negative slope of the profit function, no entrepreneur has an incentive to deviate from the allocation of entrepreneurs closest to the solution of $\Psi(n_1/n_2) = 1$. As the number of entrepreneurs goes to infinity, this allocation converges to the interior equilibrium.

1 must decline in order to restore the equilibrium when parameters change so as to increase the demand for group 1 products and hence their profits. This also suggests that the interior equilibrium is unstable under occasional perturbations in the process of transition from the short-run to the long run situations. For example, suppose that entrepreneurs naively gravitate towards the group with the higher profit per firm, in the spirit of the evolutionary game literature. Then the economy always converges to one of the two end points. Although a variety of sophisticated adjustment processes have been considered to address the problem of equilibrium selection (Arthur (1989), Kandori, Mailath, and Rob (1991), Matsui and Matsuyama (1991), and many others), it is difficult to come up with a sensible dynamic process, in which the interior equilibrium is stable.[10]

The model thus predicts that all entrepreneurs belong to the same group in the long run; either $n_1 = 1 - n_2 = 0$, or $n_1 = 1 - n_2 = 1$. This result suggests the tendency for restaurants and retail stores to cluster together, the tendency for the core-periphery regional patterns to emerge within the national economy, and the tendency for certain technology or industrial standard to dominate other alternatives.

2.C. Welfare

To see the welfare implications, consider the efficient allocations of this economy. From the symmetry of the CES, for any number of the products, it is always optimal to produce the equal amount of each product within the same group. Hence, the optimization problem for the short run can be reduced to the following

---

[10]It is possible to make the interior equilibrium stable in the perfect foresight adjustment process used in Matsuyama (1991), but the stability requires that the agent to have a negative time preference, and hence should be regarded as a theoretical curiosum.

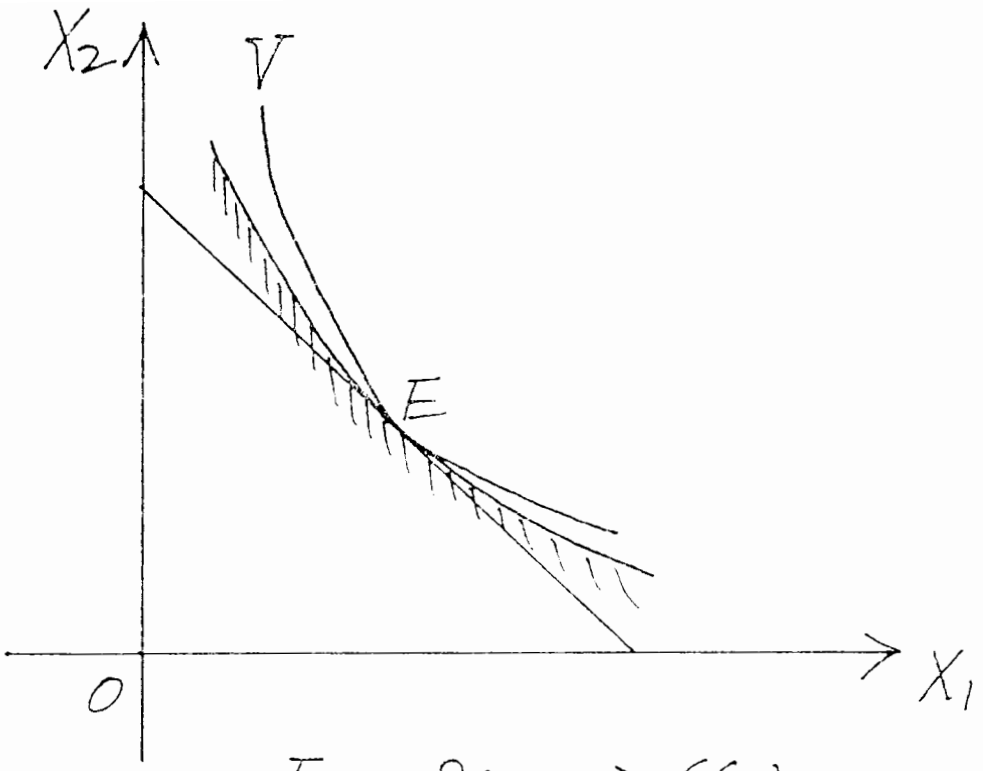

Figure 2a: $\sigma > \epsilon(\cdot)$

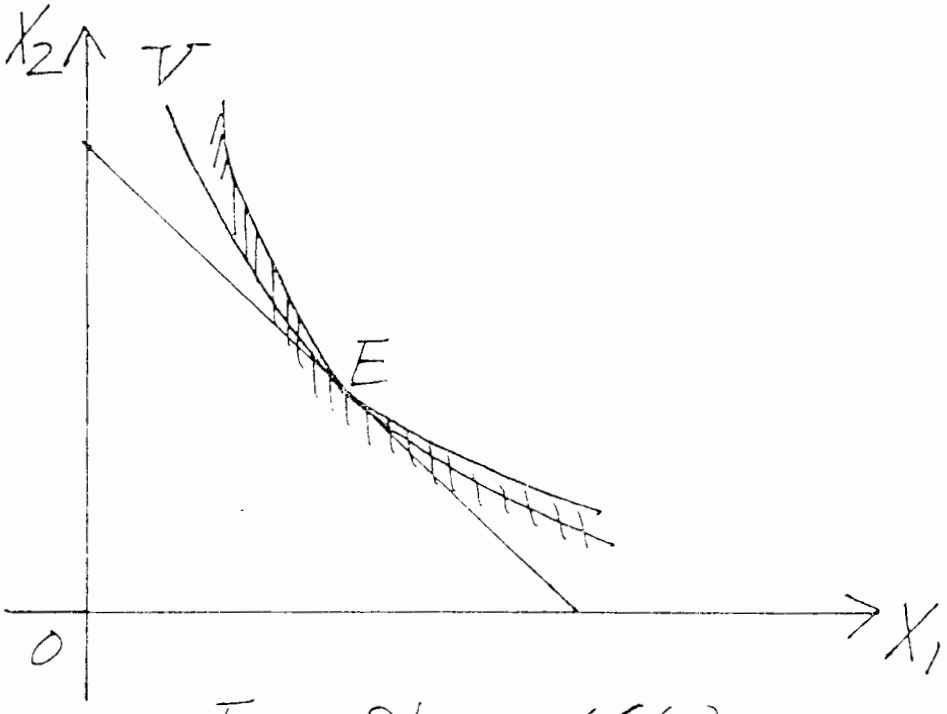

Figure 2b: $\sigma < \epsilon(\cdot)$

form: for any given n_1 and n_2, choose X_1 and X_2 to maximize

$$V(X_1, X_2) \qquad s.t., \qquad L \geq n_1^{\frac{1}{1-\sigma}} a_1 X_1 + n_2^{\frac{1}{1-\sigma}} a_2 X_2 \ ,$$

and the long run optimization problem becomes: choose n_1, n_2, X_1, and X_2 to maximize

$$V(X_1, X_2) \qquad s.t., \qquad \begin{cases} 1 \geq n_1 + n_2 \\ L \geq n_1^{\frac{1}{1-\sigma}} a_1 X_1 + n_2^{\frac{1}{1-\sigma}} a_2 X_2 \end{cases} .$$

It is easy to see that the (unique) short-run equilibrium solves the short-run optimization problem. On the other hand, the first order condition of the long run problem is satisfied by all long run equilibria, but the "unstable" one does not satisfy the second-order condition, as illustrated in Figure 2a and 2b. In both figures, the straight line represents the resource constraint in the short run, when the numbers of firms are equal to those at the long run interior equilibrium. Point E, where the indifference curve is tangent to the line, thus depicts the interior long run equilibrium. The long run resource constraint is the upper envelope of the lines generated by reallocating firms across groups. Algebraically,

$$L \geq \min_{n_1+n_2 \leq 1} \left\{ n_1^{\frac{1}{1-\sigma}} a_1 X_1 + n_2^{\frac{1}{1-\sigma}} a_2 X_2 \right\} = \left\{ [a_1 X_1]^{1-\frac{1}{\sigma}} + [a_2 X_2]^{1-\frac{1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} .$$

Point E is optimal when the indifferent curve is more curved than the long run resource constraint, that is, $\sigma > \epsilon(\bullet)$. This shows that the long run equilibrium is optimal in the case of substitutes (Figure 2a). In the case of complements, $\sigma < \epsilon(\bullet)$, the indifferent curve is flatter than the long run resource constraint; the optimality requires all firms to be concentrated in one group. The interior equilibrium is thus suboptimal (Figure 2b). One of the two "stable" equilibria is optimal, and the other is not in general optimal, depending on the relative

magnitude of $V(L/a_1, 0)$ and $V(0, L/a_2)$. The possibility thus arises where the market mechanism may fail to choose the right group and concentrate all firms into the other. This is to say that restaurants and retail stores may cluster together into an inconvenient location, that the core of the national economy may be located in the area where the severe climate makes the production costly, and the technically inferior industrial standard may be universally adopted.

Three remarks should be made here. First, the possibility of market failure demonstrated above is entirely due to a multiplicity of equilibria, hence cannot be easily resolved by the standard Pivouvian corrective taxes and subsidies. In fact, taxes and subsidies may simply introduce distortions, if the market happens to choose the right equilibrium. Second, one may argue that this kind of inefficiency could be resolved by coordination. However, if the economy initially is located at the inefficient equilibrium, and some frictions or inertia prevent all entrepreneurs from switching simultaneously, then coordinating expectations alone may not be able to help the economy escape from the inefficient equilibrium. This point has been made in an explicit dynamic setting by Matsuyama (1991, 1992b) and Matsui and Matsuyama (1992) in different contexts. Third, in the complement case, the interior equilibrium is not only unstable, but also less efficient than the two stable equilibria. Therefore, the very fact that we observe phenomena such as clustering of restaurants and retail stores, the core-periphery patterns in the regional development, and the universal adoption of a single industrial standard, does not immediately justify policy interventions, at least on the efficiency ground. In fact, if the economy happens to sit on the interior equilibrium, the government should push the economy out of it so that agglomeration emerges. Such an "symmetry breaking" policy is welfare-enhancing.

3. The Dixit-Stiglitz Model.

In the previous model, the optimal allocation is always an equilibrium. This result depends on the assumption that all products are priced at the same mark-up rate. In this section I consider the Dixit-Stiglitz (1977, Section 1) model, in which the presence of a competitively supplied good introduces distortions. Their model also has multiple equilibria under the assumption that the intergroup elasticity of substitution is larger than the intragroup one. Furthermore, a stable equilibrium could be less efficient than a unstable one in the absence of taxes and subsidies, the possibility that did not arise in the previous model. This result will be interpreted as capturing the possibility of a low equilibrium trap in economic development.

In the Dixit-Stiglitz model, the utility V depends on X and H, V(X, H), where X is the differentiated goods and H the competitively supplied outside good. One unit of the outside good is produced with one unit of labor. The main differences from the previous model are: i) there is only one group of differentiated goods producers (and hence the subscript will be omitted); ii) anyone can start new firms and introduce new products (hence there is no upper bound on the potential number of firms and products) except that f units of labor is required as the entry cost. In describing the equilibrium conditions below, I also consider corrective taxes and subsidies; τ - 1 denotes the tax rate on the labor used in manufacturing differentiated goods, and s the fraction of the entry cost subsidized.

3.A. The Short Run Analysis

As in the previous section, each differentiated goods producer faces a downward sloping demand with the constant price elasticity σ. The output price

is thus $p = a\tau/(1-1/\sigma)$, identical across firms, and

$$P = n^{\frac{1}{1-\sigma}} p = n^{\frac{1}{1-\sigma}} \left(\frac{\sigma}{\sigma-1}\right) a\tau \quad . \tag{11}$$

The relative demand for X and H is given by $V_1/V_2 = P$, or

$$\frac{X}{H} = \Phi(P) = \Phi\left(\frac{n^{\frac{1}{1-\sigma}} a\tau}{1-\sigma^{-1}}\right) \quad . \tag{12}$$

The labor market clears when

$$L = n(f + ax) + H = nf + n^{\frac{1}{1-\sigma}} aX + H \, . \tag{13}$$

Equations (11)-(13) can be solved for the unique short-run equilibrium for any given number of firms.

Let I be the aggregate income and $\alpha(P)$ the share of differentiated goods,

$$\alpha(P) \equiv \frac{PX}{I} = \frac{PX}{PX+H} = \frac{P\Phi(P)}{P\Phi(P) + 1} \quad .$$

Then, the operating profit per firm can be expressed as

$$\pi = (p - a\tau)x = \frac{px}{\sigma} = \frac{PX}{\sigma n} = \frac{\alpha I}{\sigma n} \tag{14}$$

and, from the income identity,

$$\begin{aligned} I &= L + n[\pi - (1-s)f] - nsf + (\tau-1)nax \\ &= L + n(\pi - f) + \left[1 - \frac{1}{\tau}\right]\left[1 - \frac{1}{\sigma}\right]npx \quad , \end{aligned} \tag{15}$$

the aggregate income satisfies

$$I = \frac{L - nf}{1 - \left\{1 - \frac{1}{\tau}\left[1 - \frac{1}{\sigma}\right]\right\}\alpha(P)} \quad . \tag{16}$$

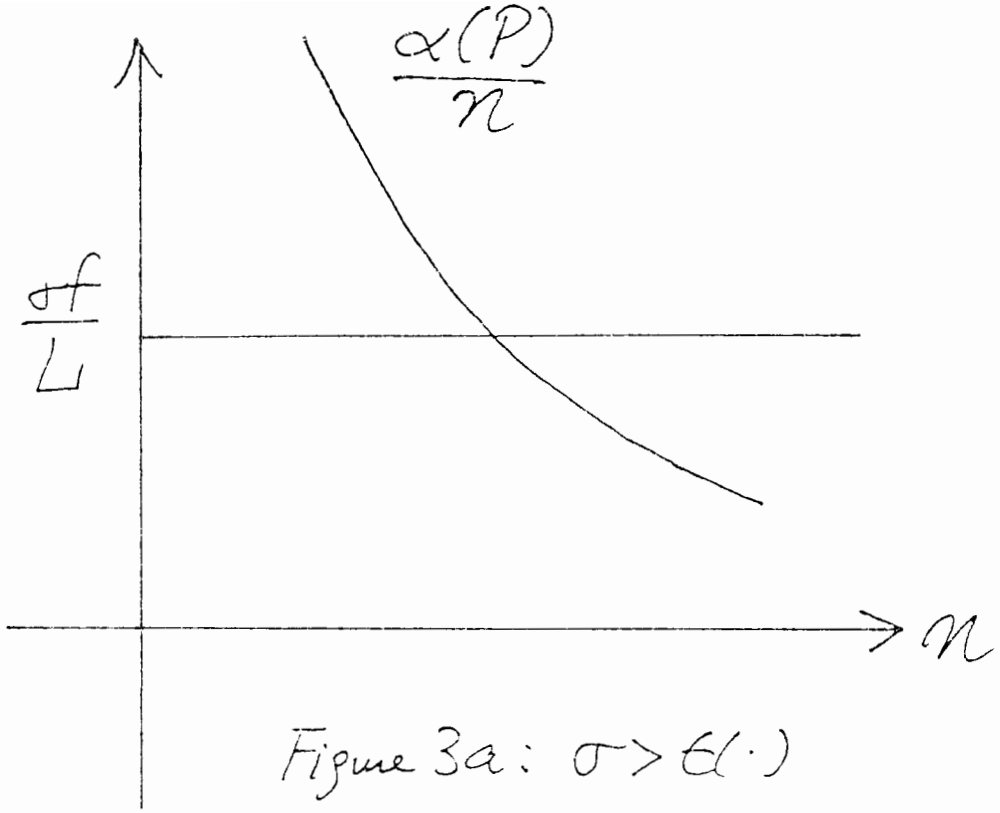

Figure 3a: $\sigma > \epsilon(\cdot)$

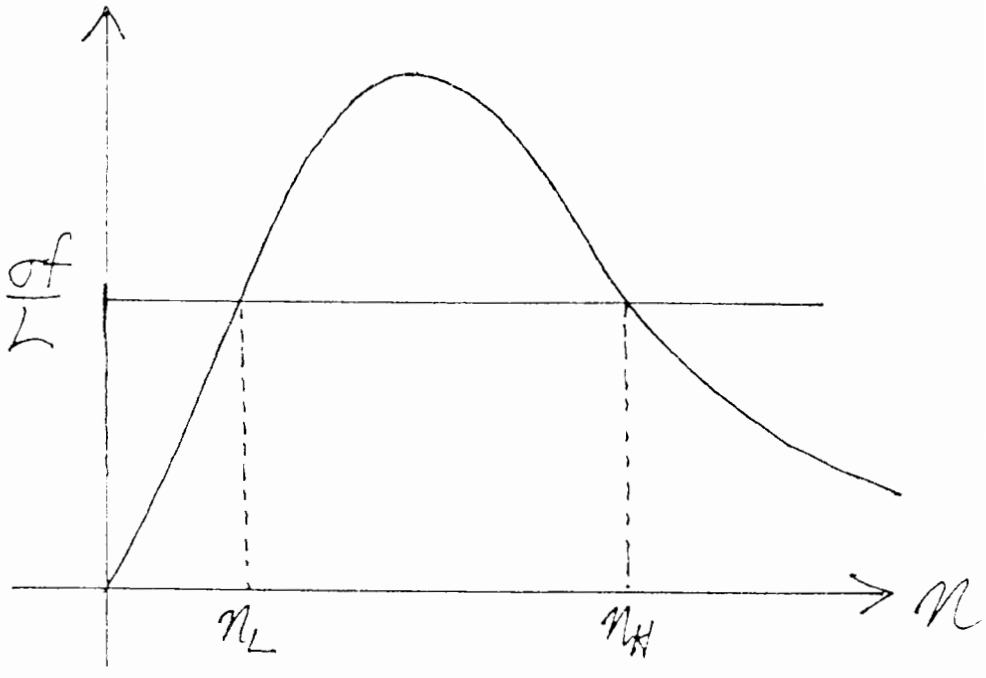

Figure 3b: $\sigma < \epsilon(\cdot)$

3.B. The Long Run Analysis

There is an incentive to set up a firm unless $\pi \leq (1-s)f$. This no-profitable-entry condition is equivalent to, from (14) and (16),

$$\frac{\sigma f}{L}\left[1 - s + \Theta\alpha(P)\right] \geq \frac{\alpha(P)}{n} ,$$

where P is given by (11) and

$$\Theta \equiv s\left\{1 - \frac{1}{\tau}\left[1 - \frac{1}{\sigma}\right]\right\} - \left[1 - \frac{1}{\tau}\right]\left[1 - \frac{1}{\sigma}\right] .$$

To see what is involved in this condition, consider the situation with no government intervention ($s = 0$ and $\tau = 1$). Then,

$$\frac{\sigma f}{L} \geq \frac{\alpha(P)}{n} = \frac{1}{n}\,\alpha\left(\frac{n^{\frac{1}{1-\sigma}}a}{1-\sigma^{-1}}\right) . \qquad (17')$$

As in the previous section, this model has an arbitrary number of long run equilibria. The uniqueness of equilibrium is ensured if the right hand side of this equation, the market share of each firm, declines with n, as shown in Figure 3a. With the declining market share, the entry of a firm has a negative effect on the profit of other firms. In other words, the entry decisions by the firms are strategic substitutes. Because of the negative feedback effect, the unique equilibrium is hence stable under any plausible entry dynamics. Dixit and Stiglitz [1977, eqs. (6) and (17)] showed that the condition for the declining market share can be written as

$$\left[1 - \alpha(P)\right]\left[\epsilon(P) - 1\right] < \sigma - 1 . \qquad (18)$$

They also demonstrated that this condition is equivalent to the condition for the dd curve to be more elastic than the DD curve. Thus (18) also ensures that the differentiated products are substitutes in the sense of Hicks and Allen. It

should be noted that the assumption that the intragroup elasticity of substitution is larger than the intergroup one, $\sigma > \epsilon(\bullet)$, is sufficient for the declining market share condition and the unique equilibrium. More generally, in the presence of government intervention, the long run equilibrium is unique and stable if

$$(1-s)[1-\alpha(P)][\epsilon(P)-1] < [1-s+\theta\alpha(P)](\sigma-1) \quad . \tag{18'}$$

Again, it is sufficient to assume $\sigma > \epsilon(\bullet)$ as long as $1 - s + \theta \geq 0$.

The uniqueness of equilibrium in the case of substitutes enables us to conduct comparative static. Differentiating (17) with the equality shows that there are more active firms in equilibrium with a high entry subsidy s, a large labor supply L, and a small entry cost, f. A change in τ has an ambiguous effect on n. In the special case of Cobb-Douglas preferences, $\epsilon(\bullet) \equiv 1$, one can show that subsidizing the differentiated goods manufacturing (a lower τ) leads to fewer firms and higher output per firm.

Dixit and Stiglitz were, of course, aware of the possibility that (18) could be violated if $\epsilon(\bullet)$ is sufficiently large. They noted that, in such a case, an increase in n lowers P, and "shifts demand towards the monopolistic sector to such an extent that the demand curve for each firm shifts to the right. However, this is rather implausible (1977, p. 300)."

Nevertheless, it seems worthwhile to look at the situation in which the intergroup elasticity dominates the intragroup one, $\sigma < \epsilon(\bullet)$, as the following examples suggest.

1) Urbanization. Suppose that X represents a composite of goods and services that, when used together, enhance the amenities of urban life. They include a

wide array of complementary offerings, ranging from delicatessens to health clubs to laundromats to theaters, each of which need to be highly specialized. On the other hand, H represents the basic good that could be produced in a rural setting. People prefer to lead a simple rural life, when the varieties offered in cities are limited. However, if more options are available, they would be willing to acquire urban lifestyles.

2) Specialization and economic progress. Suppose that V(X,H) represents a constant-returns to scale technology of the competitive consumer goods industry, when H is the labor input and X a variety of specialized capital goods and producer services. The range of specialized inputs available determines the stage of development. In a highly developed stage, a large number of specialist firms are active and cater to the needs of the consumer goods industry. The presence of the vast network of auxiliary industries make the consumer goods producers to adopt a more roundabout way of production and rely heavily on the intermediate inputs. On the other hand, in the stage of underdevelopment, the limited availability of specialized inputs forces the consumer goods producers to use a more labor intensive technology. Thus, productivity growth and economic progress in this economy are associated with a greater indirectness in the production process and an higher degree of specialization.

Under the assumption, $\sigma < \epsilon(\bullet)$,

$$\frac{P^{1-\sigma}}{\alpha(P)} = P^{1-\sigma} - \frac{1}{P^{\sigma}\Phi(P)} = P^{1-\sigma} + \frac{1}{\Phi(1)} \exp\left[\int_1^P \frac{\epsilon(v)-\sigma}{v}\, dv\right]$$

approaches infinity, as P goes to either zero or infinity. Hence, the market share of each firm, $\alpha(P)/n$, is a hump-shaped function of n, satisfying $\lim_{n\to 0}\{\alpha(P)/n\} = \lim_{n\to\infty}\{\alpha(P)/n\} = 0$. The entry of a firm increases the market

share of other firms, at least for a certain range. This is because the new firm contributes to an expanding variety, thereby generating demand for complementary products. In the context of urbanization, more coffee houses and health clubs make the urban lifestyle more acceptable, an increasing number of people attracted to the city light, thereby generating demands for other services. In the context of economic development, new firms add a variety to the set of support industries, which induces the consumer goods producers to adopt an even more roundabout way of production, thereby generating demands for other related industries.

For a sufficiently small $\sigma f/L$, the model has thus at least three equilibria. Figure 3b illustrates the equilibrium condition (17') for the case in which the market share function has a single peak. One sufficient condition for the single-peakness is that V is a CES; $\epsilon(\bullet) = \epsilon > \sigma$. Based on the same logic given in the previous section, the equilibrium in the middle, $n = n_L$, where the horizontal line cuts the upward sloping part of the market share function, could be ruled out. This equilibrium is not robust in that it cannot be considered as the limit of an equilibrium of large, but finite economies. Under any plausible dynamics, it is unstable. On the other hand, the high level equilibrium, $n = n_H$, and zero level equilibrium, $n = 0$, are both stable.

To interpret this result, particularly in the context of economic development, it would be useful to imagine the following dynamic scenario. In this story, $n = n_L$ can be thought of as a threshold level, or a critical mass of firms necessary for the economy to achieve productivity growth. Below the threshold, the limited availability of specialized inputs forces the consumer goods industry to use a labor intensive technology. This in turn implies the small market size, which induces firms to exit from the intermediate inputs

sector. This circularity between the market size and the degree of specialization, -- "the division of labor is limited by the extent of market"--, forces the economy to gravitate toward n = 0. On the other hand, above the threshold, the very fact that the relation is circular makes a cumulative advance possible. Over time, productivity growth is achieved through increasing availability of specialized inputs and more indirect methods of production. This story can in fact be formulated rigorously in a dynamic general equilibrium model of monopolistic competition, as demonstrated in Ciccone and Matsuyama (1992).

3.C. Welfare

In the absence of taxes and subsidies, market equilibrium allocation is generally inefficient in the Dixit-Stiglitz model. Furthermore, when multiple equilibria exist in the long run, their welfare ranking is unambiguous. To see this, suppose without further loss of generality that the upper-tier utility function V be linear homogeneous. Then,

$$W = V(X,H) = V(\alpha(P)/P, 1-\alpha(P))\, I \quad . \tag{19}$$

Note that, as seen from (15), s = 0 and τ = 1 imply I = L, independent of n. Thus, differentiating (19) with respect to n yields

$$\frac{1}{W}\frac{dW}{dn} = -\frac{V_1}{V}\frac{\alpha(P)}{P^2}\frac{dP}{dn} = \frac{\alpha(P)}{(\sigma-1)n} > 0 \ ,$$

where use has been made of the envelope theorem and $V = V_1(\alpha/P) + V_2(1-\alpha) = V_1/P = V_2$. Hence, there is a positive relation between the welfare and the number of active firms in equilibrium. For the situation depicted in Figure 3b, the high level equilibrium is more efficient than the middle equilibrium, which in turn more efficient than the zero level equilibrium. This result provides a sharp contrast with the previous model, where the optimality of long run equilibria has

a close connection with their stability. The positive feedback now introduces the possibility of a stable equilibrium, which is less efficient than a unstable one. In the context of increasing specialization and economic progress, this result can be interpreted as the existence of an underdevelopment trap, or the vicious circle of poverty.

Let us now turn to the role of taxes and subsidies in correcting distortions. The inefficiency of the (unique) short run equilibrium is entirely due to the monopoly pricing, and hence easy to remedy. The efficient allocation in the short run can be found as the solution to the problem: for a given n, to choose X and H to maximize V(X,H) subject to the resource constraint, (13). As shown in (13), the shadow price of the composite is $an^{1/(1-\sigma)}$, and hence the unique solution can be implemented as the market equilibrium by setting $\tau = 1 - 1/\sigma$. That is, to countervail the monopoly mark up, the efficiency requires the subsidy on the use of labor in manufacturing the differentiated goods (or, equivalently the tax on the use of labor in manufacturing the competitive good).

In the long run optimization problem, the number of firms (and products) is chosen to minimize the right hand side of (13). The first order condition is

$$f \geq \frac{n^{\frac{1}{1-\sigma}} aX}{(\sigma-1)n} .$$

Note that the right hand side is equal to $PX/\tau\sigma n = \pi/\tau$ in a market equilibrium, hence the first order condition with respect to the number of firms can be realized by setting $s = 1 - \tau$. This implies that, when the equilibrium is unique, the optimal allocation can be achieved by setting $s = 1/\sigma$, and that, if the government makes no attempt to alter the relative price ($\tau = 1$), no entry subsidy should be provided ($s = 0$). The latter reproduces the result of Dixit and Stiglitz (1977, Section I.C.), which states that the market equilibrium

allocation is the constraint optimum.

The situation is more complicated when $\sigma < \epsilon(\bullet)$, as the equilibrium fails to be unique in general. We already know about the problem of implementing the constraint optimum, the optimal allocation subject to $\tau = 1$. The above argument then implies no government intervention, and hence the situation is already given in Figure 3b. The high level equilibrium is the constraint optimum, but the two other equilibria exist; one of them, the less efficient one, is stable.

To see the difficulty involved in implementing the optimal allocation, let us assume that the government has already corrected the static distortions by setting $\tau = 1 - 1/\sigma$. Then, the no-profitable-entry condition (17) becomes

$$\frac{\sigma f}{L}\left[1 - s + \frac{\alpha(P)}{\sigma}\right] \geq \frac{\alpha(P)}{n} \quad \text{where} \quad P = n^{\frac{1}{1-\sigma}} a \ .$$

The left hand side is increasing in n, while the right hand is hump-shaped. Figure 4 shows this condition for the single-peak market share function, with $s = 1/\sigma$. As in the case of no government intervention, which is depicted by the dotted curves in the figure, there are three equilibria with a sufficiently small $\sigma f/L$; the zero level and high level equilibria are stable, and the middle equilibrium is unstable. An increase in the entry subsidy shifts down the upward-sloping curve and hence increases (reduces) the number of firms in the high (middle) equilibrium. This relation is given in the lower half of Figures 5. The welfare level of these equilibria, generated by changing s, only depends on n, but not on s; when $\tau = 1 - 1/\sigma$, $I = L - nf$ and hence the entry subsidy affects the aggregate income only through its effect on n. Differentiating (19) where $I = L - nf$ yields

$$\frac{1}{W}\frac{dW}{dn} = \frac{\alpha(P)}{(\sigma-1)n} - \frac{f}{I} = \frac{\sigma}{(\sigma-1)I}\left[\pi - \left(1 - \frac{1}{\sigma}\right)f\right]$$

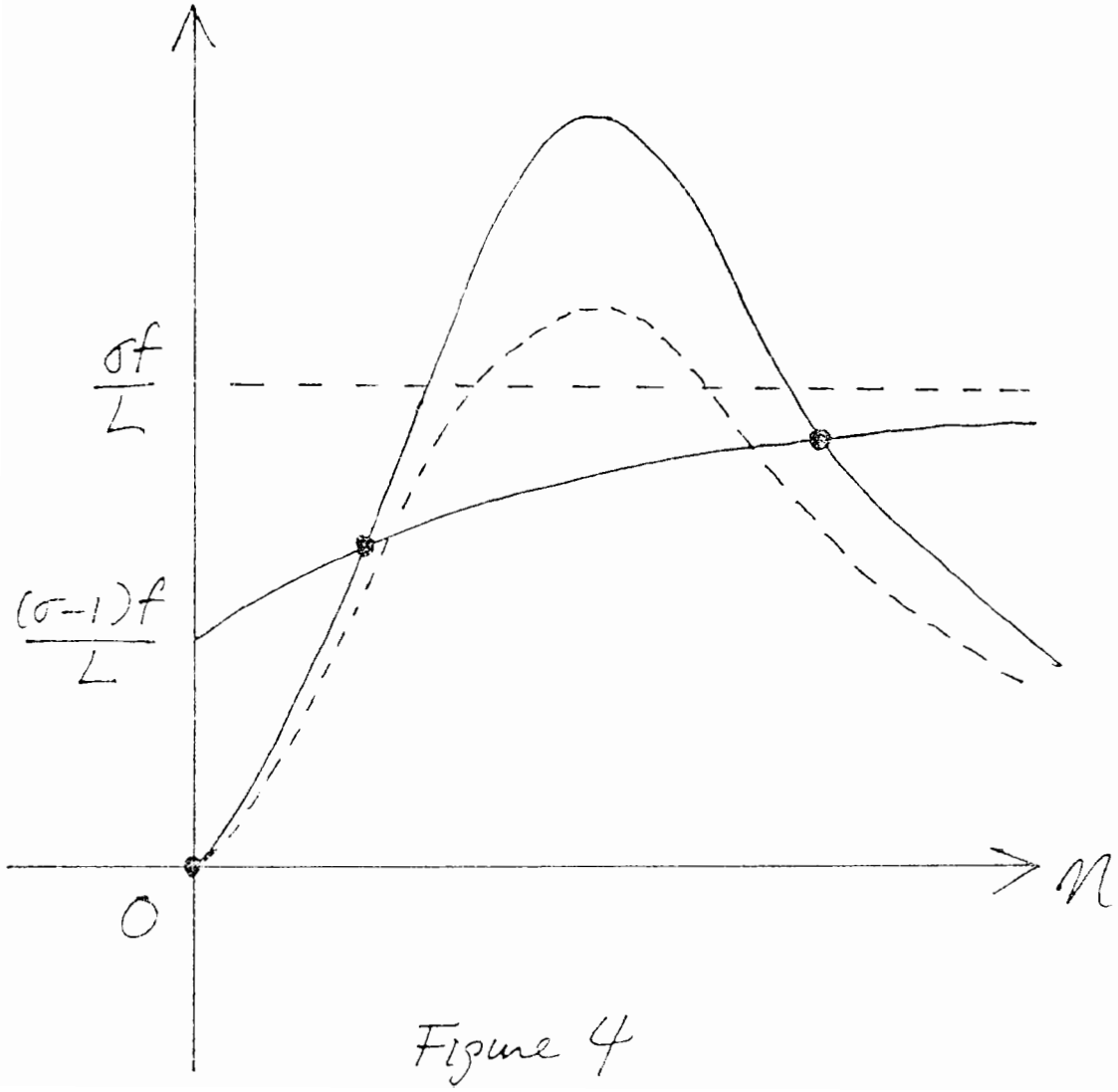

Figure 4

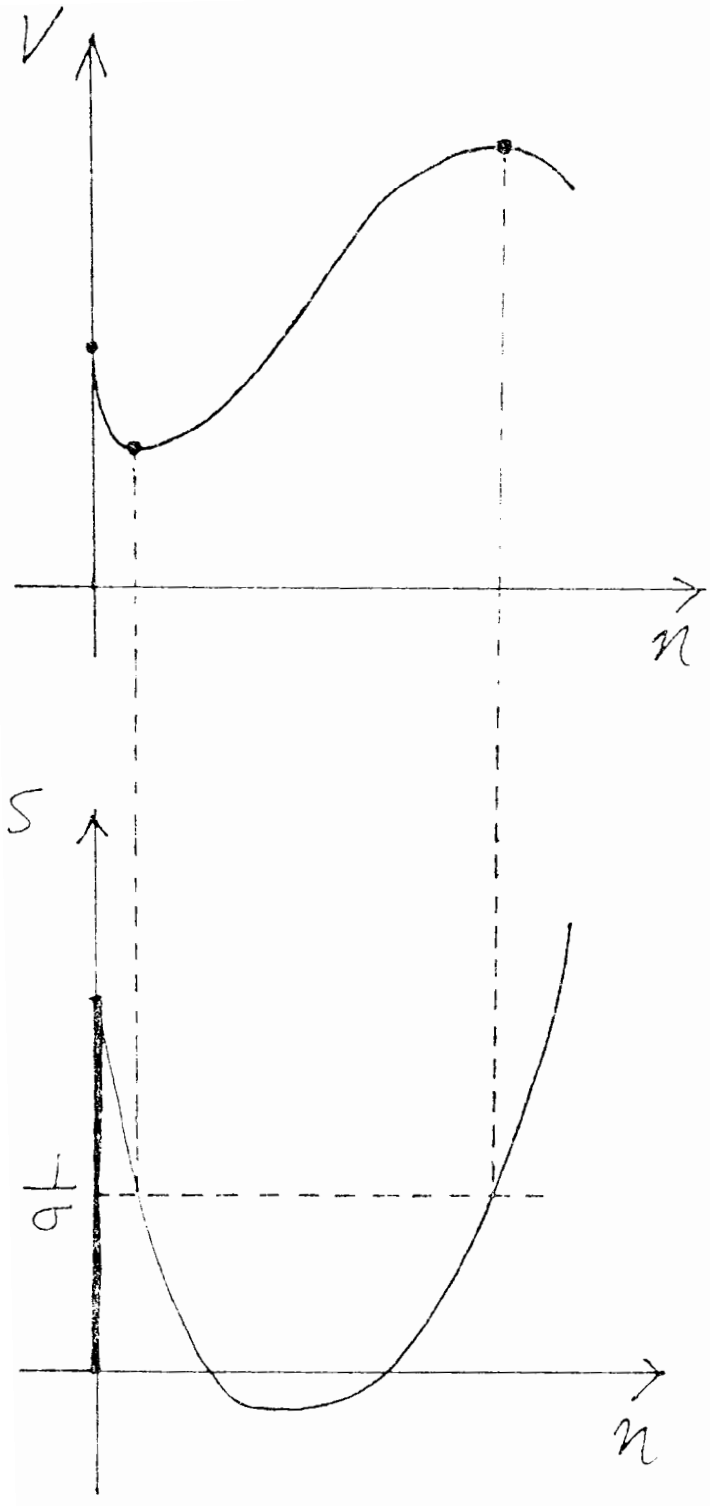

Figure 5a

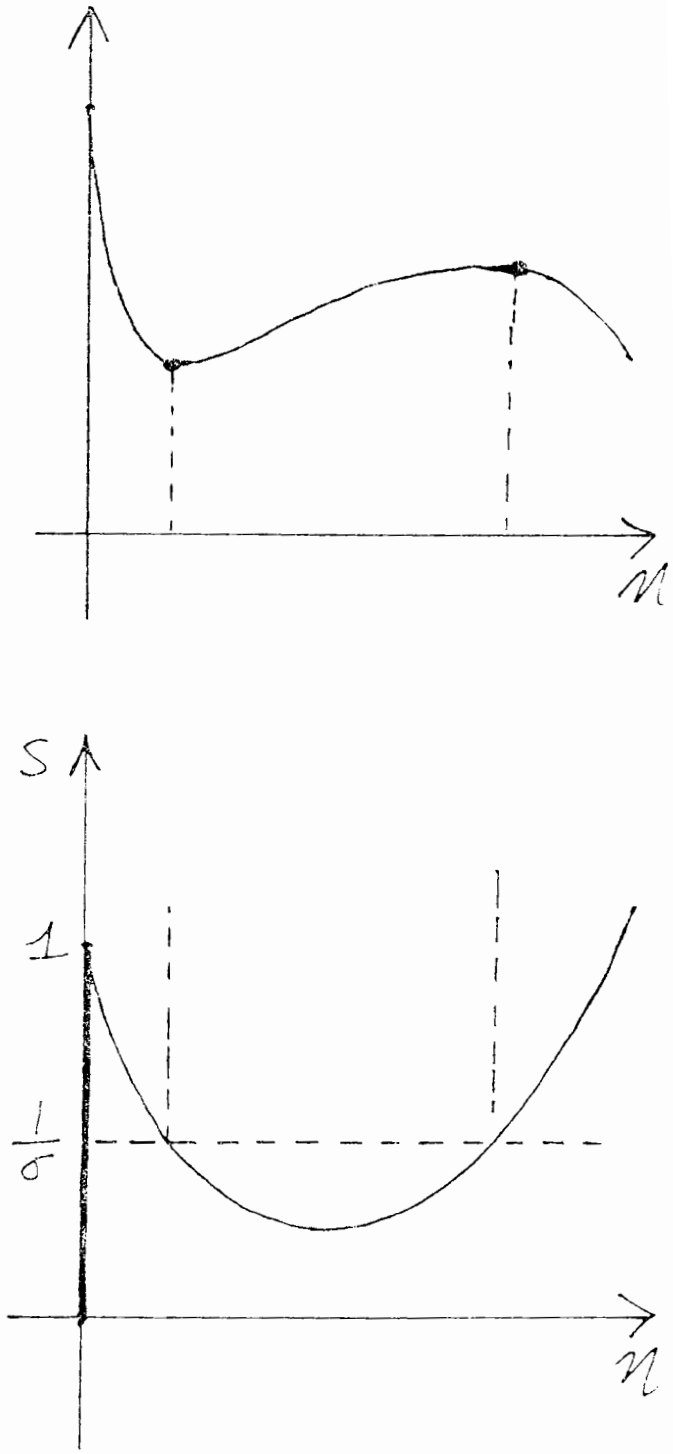

Figure 5b

and, since $\pi = (1-s)f$ whenever $n > 0$,

$$\frac{1}{W}\frac{dW}{dn} = \frac{(1-\sigma s)f}{(\sigma-1)L} \quad . \tag{20}$$

This relation is depicted in the upper half of Figures 5. Figure 5a shows the case in which the optimal allocation can be achieved in the high level equilibrium with $\tau = 1 - 1/\sigma$ and $s = 1/\sigma$, while Figure 5b shows the case in which the optimal allocation requires $n = 0$. As seen in (20), adding more firms incrementally could enhance the welfare if and only if $s < 1/\sigma$. This means that an increase in the entry subsidy improves (reduces) welfare in a stable (an unstable) equilibrium when $s < 1/\sigma$, while it reduces (improves) welfare in a stable (an unstable) equilibrium when $s > 1/\sigma$. At $s = 1/\sigma$, there are generally multiple equilibria and the level of welfare reaches a local maximum (minimum) in a stable (an unstable) equilibrium.

Some remarks are in order. First, whenever multiple equilibria exist in the absence of taxes and subsidies, there are also multiple equilibria when $s = 1/\sigma$ and $\tau = 1 - 1/\sigma$. This is because introducing the taxes and subsidies increases the right hand side of (17) and reduces its left-hand side, as shown in Figure 4. An attempt to correct distortions in fact makes a multiplicity of equilibria more likely. Such a case arises when $\sigma f/L$ is large so that $n = 0$ is the only equilibrium in the absence of taxes and subsidies. Setting $s = 1/\sigma$ and $\tau = 1 - 1/\sigma$ then may generate new equilibria that are less efficient, the possibility illustrated in Figure 5b. Second, the above exercise shows that the welfare ranking of market equilibria are quite sensitive to the extent to which the government corrects the static distortion. In the absence of taxes and subsidies, the number of active firms is a perfect indicator for the welfare levels of different equilibria. Once the static distortion is corrected,

however, the stability of equilibria suggests the local optimality of equilibria, as in the model of the previous section. Third, as discussed in the previous section, the simple corrective taxes and subsidies are not enough to implement the optimal allocation. Even when the static distortions are corrected and any stable equilibrium is a local optimum, the problem of coordination would still remain.

Before proceeding, it would be useful to point out that much of the results in this section would carry over even if $\epsilon(\bullet) > \sigma$ only over a limited range. What matters is that the share of differentiated goods occasionally rises rapidly with n. To illustrate this, suppose that $V(X,H) = \text{Max}\ \{ X^{\alpha}H^{1-\alpha}\ ,\ X^{1-\alpha}H^{\alpha}\}$, with $\alpha < 1/2$. In the context of specialization and economic progress, this example means that the consumer goods industry has access to two Cobb-Douglas technologies. Alternatively, this example can be interpreted as a model of a small open economy, one similar to the model by Rodriguez (1992). There are two tradeable consumer goods sectors. Each sector employs nontradeable intermediate inputs and labor using a Cobb-Douglas technology; the two industries differ in their factor share of intermediate inputs. The economy is small in that the relative price of the two consumer goods is determined in the world market and equal to one. In this example, $\alpha(P) = 1 - \alpha$, if $P < 1$; $\epsilon\ [\alpha,\ 1-\alpha]$ if $P = 1$; $\alpha(P) = \alpha$ if $P > 1$. Hence, as n increases, the market share of each firm declines in general, but jumps up at the critical value of n, satisfying $P = n^{1/(1-\sigma)}a\tau/(1-1/\sigma) = 1$. The existence of such a threshold would generate multiple long run equilibria. As interpreted in the context of a small open economy, this possibility suggests that international trade may lead two identical economies, in terms of their underlying structures, to adopt completely different production patterns.

4. Complements and Substitutes

In all models considered above, differentiated products are either all substitutes to each other or all complements to each other. In this section, I extend the model of section 2, so that a product could be a substitute to some but a complement to others at the same time. There are two types of entrepreneurs in the economy; X and Y. Entrepreneurs of type X can supply X-products in either group 1 or group 2, but not in both. Let n^X_i denote the fraction of X-entrepreneurs that supply in group i, so that $n^X_1 + n^X_2 = 1$. Similarly, n^Y_i represents the fraction of Y-entrepreneurs that supplies Y-products in group i and hence $n^Y_1 + n^Y_2 = 1$. The distribution of entrepreneurs is summarized by a point in the unit square, $(n^X_1, n^Y_1) \in [0,1]^2$.

The agent's preferences are now given by a three-tier CES function,

$$W = \left[V_1^{1-\frac{1}{\gamma}} + V_2^{1-\frac{1}{\gamma}} \right]^{\frac{\gamma}{\gamma-1}},$$

where

$$V_i = \left[X_i^{1-\frac{1}{\epsilon}} + Y_i^{1-\frac{1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}, \quad (i=1,2)$$

and

$$X_i = \left[\int_0^{n_i^X} [x_i(z)]^{1-\frac{1}{\sigma}} dz \right]^{\frac{\sigma}{\sigma-1}}, \quad Y_i = \left[\int_0^{n_i^Y} [y_i(z)]^{1-\frac{1}{\sigma}} dz \right]^{\frac{\sigma}{\sigma-1}}, \qquad (i=1,2)$$

with $\sigma > 1$. The budget constraint is

$$\sum_{i=1}^{2} \left[\int_0^{n_i^X} p_i(z)\, x_i(z)\, dz + \int_0^{n_i^Y} q_i(z)\, y_i(z)\, dz \right] \leq I$$

where $p_i(z)$ and $q_i(z)$ denote the prices of X-product of variety z in group i and Y-product of variety z in group i, respectively. The demand for each product is

$$x_i(z) = \left[\frac{P_i^{\sigma-\epsilon}R_i^{\epsilon-\gamma}I}{R_1^{1-\gamma}+R_2^{1-\gamma}}\right][p_i(z)]^{-\sigma}, \qquad y_i(z) = \left[\frac{Q_i^{\sigma-\epsilon}R_i^{\epsilon-\gamma}I}{R_1^{1-\gamma}+R_2^{1-\gamma}}\right][q_i(z)]^{-\sigma} \tag{21}$$

where

$$P_i = \left[\int_0^{n_i^X}[p_i(z)]^{1-\sigma}dz\right]^{\frac{1}{1-\sigma}}, \qquad Q_i = \left[\int_0^{n_i^Y}[q_i(z)]^{1-\sigma}dz\right]^{\frac{1}{1-\sigma}}, \tag{22}$$

and

$$R_i = \left[P_i^{1-\epsilon}+Q_i^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, \tag{23}$$

for i = 1, 2. Thus, all producers face the market demand for a constant price elasticity, σ. If one unit of output requires one unit of the primary factor, taken to be the numeraire, then $p_i(z) = q_i(z) = \sigma/(\sigma-1)$ and hence, from (21)-(23),

$$P_i = \left[n_i^X\right]^{\frac{1}{1-\sigma}}\left(\frac{\sigma}{\sigma-1}\right), \qquad Q_i = \left[n_i^Y\right]^{\frac{1}{1-\sigma}}\left(\frac{\sigma}{\sigma-1}\right)$$

and, hence, the relative profit for X-entrepreneurs depends solely on the distribution of entrepreneurs:

$$\begin{aligned}\frac{\pi_1^X}{\pi_2^X} &= \frac{p_1x_1/\sigma}{p_2x_2/\sigma} = \left[\frac{P_1}{P_2}\right]^{\sigma-\epsilon}\left[\frac{R_1}{R_2}\right]^{\epsilon-\gamma}\\ &= \left[\frac{P_1}{P_2}\right]^{\sigma-\epsilon}\left[\frac{P_1^{1-\epsilon}+Q_1^{1-\epsilon}}{P_2^{1-\epsilon}+Q_2^{1-\epsilon}}\right]^{\frac{\epsilon-\gamma}{1-\epsilon}}\\ &= \left[\frac{n_1^X}{n_2^X}\right]^{\frac{\sigma-\epsilon}{1-\sigma}}\left[\frac{\left(n_1^X\right)^{\frac{1-\epsilon}{1-\sigma}}+\left(n_1^Y\right)^{\frac{1-\epsilon}{1-\sigma}}}{\left(n_2^X\right)^{\frac{1-\epsilon}{1-\sigma}}+\left(n_2^Y\right)^{\frac{1-\epsilon}{1-\sigma}}}\right]^{\frac{\epsilon-\gamma}{1-\epsilon}}\end{aligned} \tag{24}$$

Similarly,

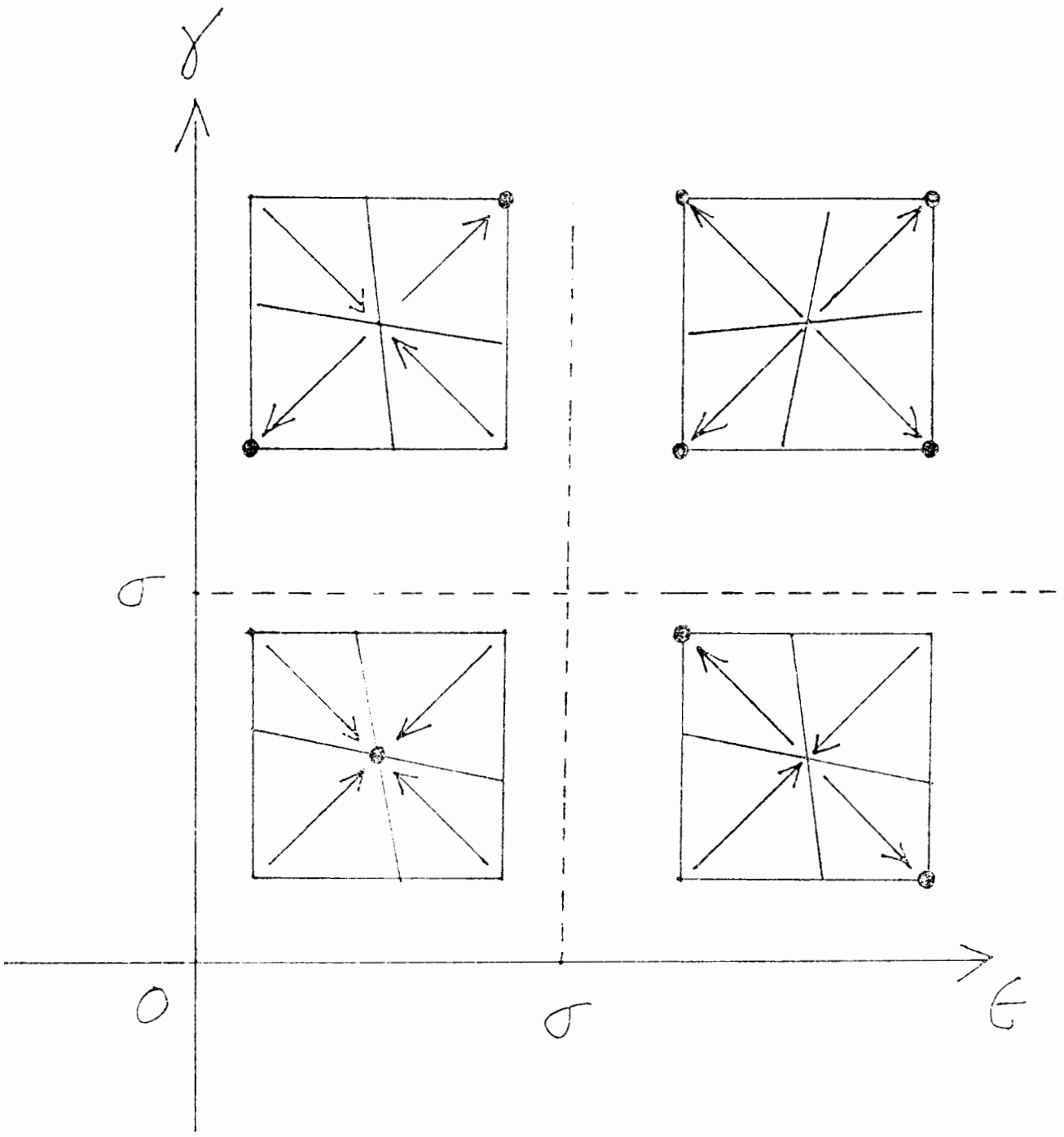

Figure 6.

$$\frac{\pi_1^Y}{\pi_2^Y} = \left[\frac{n_1^Y}{n_2^Y}\right]^{\frac{\sigma-\epsilon}{1-\sigma}} \left[\frac{(n_1^X)^{\frac{1-\epsilon}{1-\sigma}} + (n_1^Y)^{\frac{1-\epsilon}{1-\sigma}}}{(n_2^X)^{\frac{1-\epsilon}{1-\sigma}} + (n_2^Y)^{\frac{1-\epsilon}{1-\sigma}}}\right]^{\frac{\epsilon-\gamma}{1-\epsilon}} \tag{25}$$

The long run equilibrium is described by the distribution of entrepreneurs, (n^X_1, n^Y_1), in which no entrepreneurs gain from reallocating. Clearly, the center point in the unit square, (n^X_1, n^Y_1) = (1/2,1/2), is an equilibrium, but its stability depends on the parameter values, as shown in Figure 6. Given the relative magnitudes of the three parameters, σ, ϵ, and γ, the figure displays a unit square, in which how the incentive to reallocate depends on the distribution of entrepreneurs is indicated by arrows. As one moves clockwise on the parameter space, the locus along which $\pi^X_1 = \pi^X_2$ rotates clockwise, and the locus along which $\pi^Y_1 = \pi^Y_2$ rotates counter-clockwise.[11] As seen in the figure, the middle equilibrium, (1/2, 1/2), is stable if $\sigma > \epsilon, \gamma$; (0,0) and (1,1) are stable equilibria if $\gamma > \sigma > \epsilon$; (0,0), (1,0), (1,1) and (0,1) are stable equilibria if $\epsilon, \gamma > \sigma$; (1,0) and (0,1) are stable equilibria if $\gamma < \sigma < \epsilon$.

To grasp the intuition behind this result, suppose that $n^X_1 = n^Y_1 = n_1$. Then, (24) and (25) become

$$\frac{\pi_1^X}{\pi_2^X} = \frac{\pi_1^Y}{\pi_2^Y} = \left[\frac{n_1}{1-n_1}\right]^{\frac{\sigma-\gamma}{1-\sigma}}$$

Hence, the middle equilibrium is stable if $\gamma < \sigma$ and unstable if $\gamma > \sigma$ along the diagonal of the unit square, $n^X_1 = n^Y_1$. This is to say that, when Group 1 products and Group 2 products are close substitutes, all entrepreneurs tend to cluster together into the same group. Suppose now that $n_1 = n^X_1 = 1 - n^Y_1$, then

---

[11]Some algebra can show that the slope of $\pi^X_1 = \pi^X_2$ at (1/2,1/2) is equal to $(\epsilon+\gamma-2\sigma)/(\epsilon-\gamma)$, and that of $\pi^Y_1 = \pi^Y_2$ is equal to $(\epsilon-\gamma)/(\epsilon+\gamma-2\sigma)$, and that the two loci intersect only at (1/2,1/2), as long as $\epsilon \neq \sigma$ and $\gamma \neq \sigma$.

$$\frac{\pi_1^X}{\pi_2^X} = \left[\frac{n_1}{1-n_1}\right]^{\frac{\sigma-\epsilon}{1-\sigma}} = \frac{\pi_2^Y}{\pi_1^Y}$$

which shows that the middle equilibrium is stable if $\epsilon < \sigma$ and unstable if $\epsilon > \sigma$ along the other diagonal, $n^X_1 + n^Y_1 = 1$. Thus, when X-products and Y-products are close substitutes, X-entrepreneurs and Y-entrepreneurs tend to avoid each other in their group choice.

When the intragroup elasticity is larger than the two intergroup elasticities ($\sigma > \gamma, \epsilon$), all products are substitutes and the center point is stable. On the other hand, if it is smaller ($\sigma < \gamma, \epsilon$), then all corner points are stable equilibria. These results essentially replicate those of section 2. Some new possibilities arise when the intragroup elasticity falls between the two intergroup elasticities.

1) Shopping districts ($\gamma > \sigma > \epsilon$). Suppose that X-entrepreneurs plan to open restaurants and Y-entrepreneurs retail stores in either one of the two shopping districts, East and West. A restaurant and a retail store are highly complementary to each other ($\epsilon < \sigma$), so that they tend to seek each other. On the other hand, the two shopping districts are similar ($\gamma > \sigma$), and hence consumers are responsive to differences in the selection offered in the two places. This creates a tendency for restaurants and retail stores to cluster together in the same place.

This example explains, unlike one given in section 2, that even retail stores that sell similar products cluster together. The presence of other stores

that supply complementary products is responsible for this result.[12] This is in fact seems to be the explanation suggested by Chamberlin in his review (1962, Appendix C) of Hotelling's (1929) spatial competition model. After pointing out that the result of Hotelling, which asserts that the two stores selling identical products will locate next to each other, is not robust, Chamberlin argued "It is obviously for the convenience of buyers that <u>different</u> products be sold in proximity to each other, and herein lies the explanation of most of the concentration actually found in retail trading.... The "shopping district" combines on a grand scale the two principles of grouping (a) widely different products, and (b) many varieties of each (1962, pp.262-263; Emphasis in the original)."

2) <u>Regional economies ($\gamma < \sigma < \epsilon$)</u>. Suppose that there are two regions, North and South, and two industries, i = 1 and 2. Each industry produces a single final good, employing a variety of nontradeable producer services. These services are highly specific to the industry, and the services designed for industry 1 are good substitutes to each other than to those designed for industry 2 ($\gamma < \sigma$). On the other hand, the good produced by the i-th industry in North, X_i, is a very close substitute to the good produced by the same industry in South, Y_i ($\sigma < \epsilon$). Finally, X-entrepreneurs, who live in North, and Y-entrepreneurs, who live in South, choose which industry to supply services. Since the two regions produce close substitutes in both industries, each region has a comparative advantage in the industry that enjoys access to a larger number of suppliers. Entrepreneurs in turn choose to supply in the industry that has

---

[12]Wolinsky (1983) and Dudey (1990), on the other hand, provide the explanation based on the imperfection information of consumers.

comparative advantage. This positive feedback effects lead to a complete specialization. This pattern of industrial localization, once emerged, becomes self-perpetuating.

In this example, the outcome is quite striking if compared to the situation in which the two regions do not trade. In autarky, the North residents are barred from consuming Y_i and the South residents from consuming X_i. Then, from the analysis in section 2, $\gamma < \sigma$ implies that there is a unique equilibrium, in which one half of entrepreneurs chooses industry 1 and the other half chooses industry 2 in both regions. In other words, the middle equilibrium is stable. Removing the trade barrier makes this equilibrium unstable. Under occasional perturbations, each region gains comparative advantage in one industry, and then the positive feedback effect through the reallocation of entrepreneurs leads to a complete specialization of each region, despite there is no inherent difference between the two regions.

5. Concluding Remarks.

The standard practice in the monopolistic competition literature is to assume, often implicitly, that the intergroup elasticity of substitution is smaller than the intragroup elasticity of substitution. What I argued in this paper is that, under certain situations, it is more plausible to assume that the former is larger than the latter. Once the inequality is reversed, then products that belong to the same group become complementary to each other. Hence, entry of new firms will increase the profit of incumbent firms. Because of this positive feedback effect, firms tend to cluster together. This property makes the monopolistic competition model a useful apparatus within which to demonstrate a variety of agglomeration phenomena, phenomena that are difficult to obtain in

the classical paradigm of perfect competition.

This result also helps to clarify the recent work by Fujita (1990, ch. 8.4), Helpman and Krugman (1985, ch. 11), Krugman (1991), Rivera-Batiz (1988). The monopolistic competition model of an open economy yields predictions similar to the classical theory of trade in many respects, as long as the differentiated goods are costlessly tradeable.[13] The above studies demonstrated that the predictions drastically change and geographical concentration occurs if differentiated goods are costly to transport, while migration is free (in the case of differentiated consumption goods) and the final good is tradeable (in the case of differentiated inputs). As the preceding analysis suggests, the reason why predictions drastically change is that high transportation costs effectively group the differentiated goods based on the location and that free migration and free movement of the final good make the intergroup substitution infinite.

Other than being faithful to the spirit of Chamberlin, one reason why the existing literature on monopolistic competition only considers the case of substitutes may be that the model has multiple equilibria in the case of the complements. To some economists a model with multiple equilibria may be unsettling. In particular, it poses a serious problem concerning the validity of comparative static. It is thus understandable if those interested in the policy analysis are tempted to make restrictions that lead to a unique equilibrium. However, it should be remembered that the mere fact that certain parameter values ensure the uniqueness of equilibrium does not mean that they are more plausible than those implying multiplicity. Nor should one conclude that a model with multiple equilibria cannot yield useful predictions. For the fact

---

[13]In fact, much of Helpman and Krugman's (1985) book is devoted to the demonstration of how far the insights of the classical theory of trade carry over in the presence of imperfect competition and increasing returns.

that the multiplicity results in certain cases and not in others itself has some predictive content. That is precisely the nature of predictions given in the analysis above. The condition under which agglomeration and clustering take place is an important thing to know; it can also be tested empirically. One should not deny the possibility of agglomeration simply because one cannot tell where the firms agglomerate.

References:

Anderson, Simon P., Andre de Palma, and Jacques-Francois Thisse, "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach," Review of Economic Studies 56 (January 1989) 21-35.

Arthur, W. Brian, "Self-Reinforcing Mechanisms in Economics," in The Economy as an Evolving Complex System, edited by Philip W. Anderson, Kenneth J. Arrow, and David Pines, Addison-Wesley Company, 1988.

Arthur, W. Brian, "Competing Technologies, Increasing Returns, and Lock-in by Historical Events," Economic Journal 99 (March 1989): 116-131.

Arthur, W. Brian, "'Silicon Valley' Locational Clusters: When Do Increasing Returns Imply Monopoly?" Mathematical Social Sciences, 19 (February 1990): 235-251. a).

Arthur, W. Brian, "Positive Feedbacks in the Economy," Scientific American, February 1990, 92-99. b).

Azariadis, Costas, and Allan Drazen, "Threshold Externalities in Economic Development," 105 (May 1990): 501-526.

Beath, John, and Yannis Katsoulacos, The Economic Theory of Product Differentiation, (Cambridge: Cambridge University Press), 1991.

Brown, Stephen, "Retail Location Theory: The Legacy of Harold Hotelling," Journal of Retailing, 65 (Winter 1989): 450-470.

Bulow, Jeremy, John Geanakoplos, and Paul Klemperer, "Multimarket Oligopoly: Strategic Substitutes and Complements," Journal of Political Economy, 93 (June 1985): 488-511.

Chamberlin, Edward Hastings, The Theory of Monopolistic Competition, The Eighth Edition, (Cambridge: Harvard University Press), 1962.

Ciccone, Antonio, and Kiminori Matsuyama, "Start-up Costs and Pecuniary Externalities as Barriers to Economic Development," Working Papers in Economics, E-92-14, Hoover Institution, Stanford University, 1992.

Dixit, Avinash, and Victor Norman, The Theory of International Trade, (Cambridge: Cambridge University Press), 1980.

Dixit, Avinash, and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," American Economic Review 67 (June 1977): 297-308.

Dudey, Marc, "Competition by Choice: The Effect of Consumer Research on Firm Location Decisions," American Economic Review, 80 (December 1990): 1092-1104.

Ethier, Wilfred, J., "National and International Returns to Scale in the Modern

Theory of International Trade," American Economic Review 72 (1982): 389-405.

Faini, Ricardo, "Increasing Returns, Non-traded Inputs and Regional Development," Economic Journal, 94 (June 1984): 308-323.

Farrell, Joseph, and Garth Saloner, "Standardization, Compatibility, and Innovation," Rand Journal of Economics, 16 (Spring 1985), 70-83.

Farrell, Joseph, and Garth Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements and Predation," American Economic Review, 76 (December 1986): 940-955.

Friedman, Milton, "The Methodology of Positive Economics," in Essays in Positive Economics, (Chicago: University of Chicago Press), 1953.

Fujita, Masahisa, Urban Economic Theory: Land Use and City Size, (Cambridge: Cambridge University Press, 1989).

Grossman, Gene M., and Elhanan Helpman, Innovation and Growth in the Global Economy, (Cambridge, MIT Press, 1991).

Hart, Oliver D., "Monopolistic Competition in a Large Economy with Differentiated Commodities," Review of Economic Studies, 46 (January 1979): 1-30.

Hart, Oliver D., "Monopolistic Competition in the spirit of Chamberlin: A General Model," Review of Economic Studies, 52 (October 1985): 529-546.

Helpman, Elhanan, and Paul R. Krugman, Market Structure and Foreign Trade, (Cambridge: The MIT Press, 1985).

Hicks, John R., and R. C. D. Allen, "A Reconsideration of the Theory of Value, Parts I and II," Economica, N.S., February 1934, pp.52-76, and May 1934, pp.196-219.

Hotelling, Harold, "Stability in Competition," Economic Journal, 39 (March 1929): 41-57.

Jones, Larry E., "The Efficiency of Monopolistically Competitive Equilibria in Large Economies: Commodity Differentiation with Gross Substitutes," Journal of Economic Theory, 41 (1987): 356-391.

Kandori, Michihiro, George J. Mailath, and Rafael Rob, "Learning, Mutations, and Long Run Equilibrium in Games," CARESS Working Paper, #91-01, University of Pennsylvania, 1991.

Katz, Michael, and Carl Shapiro, "Network Externalities, Competition, and Compatibility," American Economic Review, 75 (June 1985): 424-440.

Katz, Michael, and Carl Shapiro, "Technology Adoption in the Presence of Network Externalities," Journal of Political Economy, 94 (August 1986): 822-841.

Krugman, Paul R., "The narrow moving band, the dutch disease, and the competitive consequences of Mrs. Thatcher: notes on trade in the presence of dynamic scale economies," Journal of Development Economics, 27 (1987), 41-55.

Krugman, Paul R., Geography and Trade, (Cambridge; MIT press), 1991.

Lucas, Robert E., Jr., "On the Mechanics of Economic Development," Journal of Monetary Economics, 22 (1988): 3-42.

Matsui, Akihiko, and Kiminori Matsuyama, "An Approach to Equilibrium Selection," Working Papers in Economics, E-91-14, Hoover Institution, Stanford University, 1991.

Matsuyama, Kiminori, "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," Quarterly Journal of Economics 106 (May 1991): 617-650.

Matsuyama, Kiminori, "Agricultural Productivity, Comparative Advantage, and Economic Growth," Working Papers in Economics, Hoover Institution, Stanford, 1992a. Also forthcoming in Journal of Economic Theory.

Matsuyama, Kiminori, "The Market Size, Entrepreneurship, and the Big Push," Working Papers in Economics, E-92-13, Hoover Institution, Stanford University. 1992b. Also forthcoming in Journal of the Japanese and International Economies.

Ogaki, Masao, "The Indirect and Direct Substitution Effects," American Economic Review, 80 (December 1990): 1271-1275.

Perloff, J., and S. Salop, "Equilibrium with Product Differentiation," Review of Economic Studies 52 (January 1985): 107-120.

Rivera-Batiz, Francisco L., "Increasing Returns, Monopolistic Competition, and Agglomeration Economies in Consumption and Production, Regional Science and Urban Economics, 18 (1988): 125-153.

Roberts, Kevin, "The Limit Points of Monopolistic Competition," Journal of Economic Theory, 22 (April 1980): 256-278.

Rodriguez, Andres, "The Big Push in a Small Open Economy," unpublished manuscript, Stanford University, 1992.

Romer, Paul M., "Growth Based on Increasing Returns Due to Specialization," American Economic Review Papers and Proceedings 77 (May 1987): 56-62.

Romer, Paul M., "Endogenous Technological Change," Journal of Political Economy, 98 (October 1990, Part 2): S71-S102.

Samuelson, Paul A., "The Monopolistic Competition Revolution," Chapter 5 in Monopolistic Competition Theory: Studies in Impact; Essays in Honor of Edward H. Chamberlin, Edited by Robert E. Kuenne, (John Wiley & Sons,

Inc., New York), 1967.

Samuelson, Paul A., "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," Journal of Economic Literature, 12 (December 1974): 1255-1289.

Sattinger, Michael, "Value of an Additional Firm in Monopolistic Competition," Review of Economic Studies 51 (April 1984): 321-332.

Stigler, George J., "Monopolistic Competition in Retrospect," in Five Lectures on Economic Problems (London; MacMillan Co.), 1949.

Triffin, Robert, Monopolistic Competition and General Equilibrium Theory, (Cambridge: Harvard University Press), 1940.

Wolinsky, Asher, "Retail Trade Concentration Due to Consumers' Imperfect Information," Bell Journal of Economics, 14 (Spring 1983): 275-282.

Young, Alwyn, "Learning by Doing and the Dynamic Effects of International Trade," Quarterly Journal of Economics 106 (May 1991):.