Hitotsubashi Journal of Economics 30 (1989) 65-85. © The Hitotsubashi Academy

THEORY OF INTERNALISATION BY MULTINATIONAL CORPORATIONS

KIYOSHI KOJIMA

Abstract

A theorisation of multinational corporation activities or the theory of internalisation developed by the Reading School (England) economists culminated in a sunk-cost model. They stress that to reduce variable cost of transaction through transfer pricing is the source of gain from internalisation through multinational operations. They reach to wrong conclusions such as that MNCs have many advantages from making internal prices lower; that it is justifiable for the MNCs to take monopolistic behaviour with the aim of maximising quasi-rent; and that it is profitable for the MNCs to switch from exporting to overseas production. In contrast, it is shown here that the real gain of internalisation depends on economies of scale of the plant, the firm and the agglomeration which the MNCs design and utilise.

I. *Introduction*

A theorisation of the multinational corporation (MNC) has recently been purused by the Reading School, England, led by John H. Dunning, Alan M. Rugman, Peter J. Buckley, and Mark C. Casson. Their views converge into a theory of internalisation which is, as explained below, essentially based on a sunk-cost model. Relying on the transation cost approach that originates in Ron Coase's classic work [1937], they stress that the practice of internal (or transfer) pricing that enables the MNC to reduce the variable costs of production (or transations in general) is the source of gains from internalisation through multinational operations. However, they forget that the real gains come from the realisation of economies of scale through establishing and operating an efficient hierarchy, the MNC.

This paper first shows that the Reading School's exposition of a theory of internalisation through multinational operations began formally with the Buckley-Casson (B-C) sunk-cost model which consists of fixed and variable costs. The B-C model introduces a criterion for judging whether a given production mode is superior to an alternative one. But the B-C criterion initially derived is inexact (in the sense that it is only a necessary but not a sufficient condition). Rugman then assumes that fixed costs are totally sunk and comes up with a modified version (the Rugman criterion). All this development is reviewed in Section II.

* I owe much to discussions with Terutomo Ozawa of Colorado State University, Norihiko Suzuki, Mitsuhiro Kaneda and Katsura Nakano of International Christian University in preparing this paper.

In section III, a more exact model of fixed costs in which the minimum optimal scale (MOS) is introduced and from which the necessary and sufficient conditions (the Kojima criterion) for a superior production mode is established.

Section IV evaluates how the Reading School conceptualises the gains from internalisation the MNC exploits from its variety of activities overseas by comparing it with the Kojima criterion. It is shown that the real gains from internalisation come not so much from the intra-company mechanisms by which the MNC manipulates its internal prices but from the economies of scale, internal and external alike, that the MNC designs and capitalises on. There is always a limit to internal scale economies but multinationalisation enables the firm to coordinate its operations more effectively with the outside market and interact more profitably with other firms, thereby reaping external scale economies.

It is the MNC's virtue to increase social welfare when it produces useful goods and services at lower costs (that is, to operate under decreasing cost condition), whereas it is the MNC's sin to adopt monopolistic behaviour through internalisation and maximises quasi-rent. Section V looks at this dual character of the MNC.

Internalisationists stress that it is profitable for the MNC to switch from exporting to overseas production. This hypothesis is critically exmained in Section VI. It is shown that such a switch is supportable from a social welfare point of view only if overseas production leads to a superior mode of production—in line with what I call "pro-trade oriented direct foreign investment (DFI)." However, the DFI induced by the existence of tariffs and other trade barriers normally results in an inferior mode of production, facilitating only market-share-seeking monopolistic behaviour—i.e., it becomes "anti-trade oriented."

II. *Buckley-Casson Model*

Figure 1 shows a diagrammatic presentation of a theory of internalisation which was originally presented by Buckley and Casson [1981] and used repeatedly in Casson [1981] and Buckley [1987]. Buckley and Casson assume a Chamberlinian [1933] decreasing cost function as follows:

(1) total cost (TC)$=c(x)=a+bx$; a,b=constant

(2) average cost (AC)$=c(x)/x=a/x+b$,

where x is the volume of outputs (or transactions, sales, purchases, etc.), a is the fixed set-up cost and b is the unit variable cost.

Let there be two modes of production indexes $i=1,2$:

(2.1) $c(x)/x=a_1/x+b_1$,

(2.2) $c(x)/x=a_2/x+b_2$.

We plot in Figure 1(i), the TC curve with a_i being OT_i and b_i being the slope of T_iT_i'. In panel (ii), the constant variable cost (equal to marginal cost, which is also constant) is shown by line M_iM_i', whereas the AC curve by A_iA_i' which decreases and approaches infinitely close to the vertical axis as $x\to 0$ and to the M_iM_i' line as $x\to\infty$.

We can see that if $a_1<a_2$ while $b_1>b_2$, a switching point, E, exists showing that for output level less than $\bar{x}$, mode 1 is more efficient (or superior), while for $x_1\leqq\bar{x}\leqq x_2$, mode 2 is more efficient in terms of both total cost and average cost. This brings us to B-C criterion for a superior mode of production:

The B-C criterion: The mode of production with a higher fixed cost and a lower variable cost (i.e., $a_1 < a_2$ and $b_1 > b_2$) is superior for larger scale of outputs.

However, it will become clear below that the B-C criterion is not general, but is inexact for identifying a superior mode of production. The B-C criterion only applies to one of the nine possible cases[1] that the combinations of a_1, a_2, b_1 and b_2 can make, and, therefore, is a necessary but not a sufficient condition for mode 2 to be superior to mode 1. As will be discussed in the following section, even in the absence of the B-C switching point, mode 2 can still be superior to mode 1.

Furthermore, suppose that the fixed set-up cost, a, consists of investment in non-recoverable assets, both tangible and intangible, and is thus totally sunk. Then, the unit (average) fixed cost, a/x becomes zero, making the AC (and TC) to be dependent solely upon the level of the variable cost, b. Since it is stressed by Rugman that making the variable cost lower through internal (or transfer) pricing is the source of gains from internalisation through multinational operations, let us introduce the following modified criterion:

The Rugman criterion: The lower the variable cost (b), the more superior the mode of production is.

The business approach focuses on how the MNC can make both the fixed cost, a, and the variable cost, b, lower through internalisation but this is only one of the possible gains and derives from a rather superficial source. The real gains come from how well the firm realises technical economies of scale.

Roughly saying, let $a = p_a A$ where A is the firm's assets, tangible as well as intangible, p_a is the price to obtain A, and thus a being the fixed set-up cost as before. Economists who take the business approach stress that the MNCs can make p_a lower. But, economies of scale are created technically (or functionally) depending upon the size of A and the capability and organisational and operational efficiency of the firm involved, etc. They also stress that direct foreign investment (i.e., multinationalisation) uses cheaper foreign labour, hence lowers the variable cost, b. But, the real gains come from economies of scale which are realised by integrated hierarchy, as will be discussed in section IV.

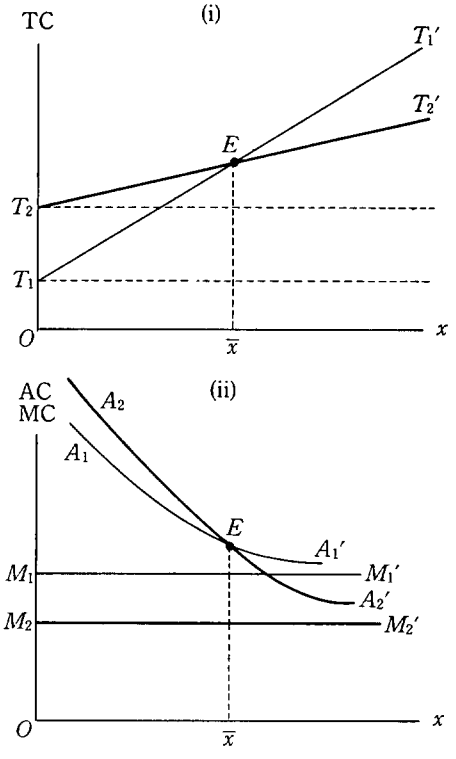
In reference to equations (2.1) and (2.2), Buckley and Casson recognising the importance of economies of scale state that:

> For any given mode, average variable cost is constant (and so equal to marginal cost, which is also constant). Superficially, this assumption rules out economies of scale in production. But if we are prepared to identify different modes according to the size or type of plant to be used, then economies of scale can be included by giving the mode 'production with large plant' higher fixed costs and lower variable costs than the mode 'production with small plant.' [Buckley and Casson 1981, p. 78].

With such specific economies of scale in mind, the B-C criterion is derived. However, since the above cost functions show that the unit fixed cost, a_i/x, decreases infinitely approaching to the constant variable cost, b_i, the degree of economies of scale which is realisable is ob-

[1] The nine cases are: $b_1 > b_2$, $b_1 = b_2$ or $b_1 < b_2$ for $a_1 < a_2$; $b_1 > b_2$, $b_1 = b_2$ or $b_1 < b_2$ for $a_1 > a_2$; and $b_1 > b_2$, $b_1 = b_2$ or $b_1 < b_2$ for $a_1 = a_2$.

FIGURE 1.



scure, and the superiority of cost functions is shown by the level of b_i, as in the Rugman criterion. Therefore, we had better to devise a measure for identifying the degree of economies of scale. This is the task of the next section.

III. *A Model of Sunk Cost with Minimum Optimal Scale*

Let us suppose the following cost function:

$$\text{(3)} \quad \text{TC}=c(x)=\begin{cases} a+bx & \text{if } x\leqq x^* \\ (a/x^*+b)x & \text{if } x>x^* \end{cases} \quad (a,b,x^*=\text{constant})$$

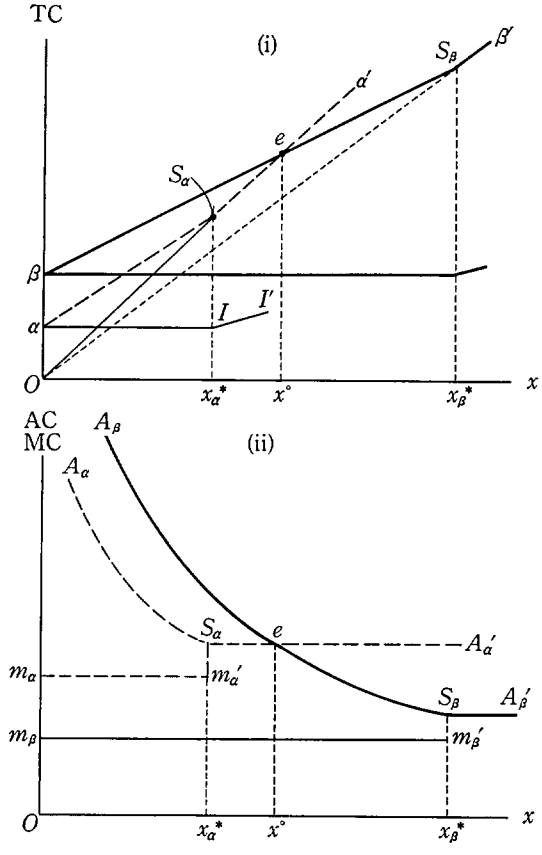
$$\text{(4)} \quad \text{AC}=c(x)/x=\begin{cases} a/x+b & \text{if } x\leqq x^* \\ a/x^*+b & \text{if } x>x^* \end{cases}$$

where x^* is a minimum optimal scale (MOS) of output (or operations in general) under a given mode of production. Up to x^*, technical economies of scale are effective and fixed cost per unit of output, a/x, decreases continuously. Beyond x^*, economies of scale exhaust and unit fixed cost, a/x^*, becomes constant.[2]

[2] The unit fixed cost may also increase beyond the MOS output level due to diseconomies of scale, but such a case is not explicitly discussed in this paper.

We found, after the completion of the draft of this paper, that Dixit [1980] used a similar cost function to our model.

FIGURE 2.



Now, let us compare two modes of production, $i=\alpha,\beta$. Then, minimum AC (which equals marginal cost) are as follows:

(5) $b_\alpha^*=a_\alpha/x_\alpha^*+b_\alpha$

(6) $b_\beta^*=a_\beta/x_\beta^*+b_\beta$ (All variables are constants.)

Minimum AC depends upon (1) variable cost, b_i, which is independently determined by economic conditions, and (2) minimum unit fixed cost, a_i/x_i^*, or the realized degree of economies of scale which is determined by technical efficiency of the firm's assets, tangible as well as intangible, and the size of MOS.

Let us explain mode α in Figure 2. $O\alpha$ is the fixed set-up cost, a, and the slope of line αS_α is the unit variable cost, b_α. At MOS, x_α^*, the line αS_α links to line $S_\alpha\alpha'$ which is an extension of OS_α, the slope of which is the minimum AC, $b_\alpha^*=a_\alpha/x_\alpha^*+b_\alpha$. The minimum unit fixed cost, a_α/x_α^*, is shown by the slope of line OII' and adding this to the slope of line αS_α results in the slope of line $S_\alpha\alpha'$. The TC curve is thus drawn as $O\alpha S_\alpha\alpha'$.

In panel (ii) of Figure 2, as shown by curve $A_\alpha S_\alpha$, AC decreases up to the MOS output, x_α^*. It then becomes constant, as line $S_\alpha A_\alpha'$ shows. Direct unit variable cost, b_α, is the height of line $m_\alpha m_\alpha'$. To this, minimum unit fixed cost, a_α/x_α^*, which is shown by the length

of $S_\alpha m_\alpha'$, is added at the MOS point, making minimum AC to be the height of $S_\alpha A_\alpha'$.

Complication comes from the fact that there are two kinds of fixed set-up costs, that is, $a=\bar{a}+\bar{\bar{a}}$.[3] While $\bar{a}=p_1\bar{A}$ is a nonrecoverable set-up cost, which is a once-and-for-all cost incurred as soon as the mode is adopted (for example, the set-up cost of tangible assets $\bar{A}$), $\bar{\bar{a}}=p_2\bar{\bar{A}}$ is a recurrent cost (that is, independent of the rate of output) which results from indivisibilities of factor inputs hired in connection with the firm's operation (for example, the salary of managers and other overhead personnels, $\bar{\bar{A}}$). Although both tangible assets $\bar{A}$, and managerial staff, $\bar{\bar{A}}$, create economies of scale up to the MOS output level, $\bar{a}$ is sunk, while $\bar{\bar{a}}$ is not and instead is recurrently needed.

Now, we can postulate a new criterion for a superior mode of production:

> *The Kojima criterion:* The lower the minimum average cost and the larger the minimum optimal scale of plant are, the superior the mode of production is.

In other words, if

(7) $\quad a_\alpha/x_\alpha^*+b_\alpha>a_\beta/x_\beta^*+b_\beta$

is satisfied, mode β is superior to mode α. In Figure 2(i), this is shown by the slope of line $O\beta'$ being less steep than that of line $O\alpha'$.

In order to satisfy equation (7), there is a critical level of output:

(8) $\quad x^\circ=a_\beta/(b_\alpha^*-b_\beta)=a_\beta/[(a_\alpha/x_\alpha^*+b_\alpha)-b_\beta]$

x° is the level of output produced with the same TC (and AC) by both modes α and β. This critical output level, x°, corresponds to a new switching point, e, where line βS_β crosses line $O\alpha'$ from above as long as $b_\alpha^*>b_\beta$. (This new switching point, e, differs from the B-C's E in Figure 1, where two variable cost lines cross with each other when $b_\alpha>b_\beta$.) Then, given the condition of $b_\alpha^*>b_\beta$, x_β^* (mode β's MOS) becomes larger than the critical rate of outputs, x°,[4] hence, the greater x_β^* is relative to x_α^* (mode α's MOS), the greater the realised economies of scale are, and so the lower the minimum AC is.

It should be stressed here that what makes a mode of production superior is the extent to which technical economies of scale are realised with a larger MOS.

IV. *Gains from Internalisation through Multinational Operations*

We see that the real gains from internalisation through multinational operations depend predominantly upon how well technical economies of scale[5] are realised. We sup-

[3] Buckley and Casson [1981, p. 76] distinguished these two kinds of fixed costs. Dunning [1988a, p. 2] distinguished between the asset-ownership (0_a) advantages and transaction-ownership (0_t) advantages of multinational enterprises, which roughly correspond to our two concepts.

[4] In the B-C case, x° becomes smaller than x_α^* and therefore, x_β^* can become smaller than x° but should be larger than x_α^*.

If $b_\alpha^*<b_\beta$, and also if $b_\alpha^*=b_\beta$, mode β cannot at all be superior to mode α, and these cases are omitted from our consideration.

[5] Instead of "economies of scale," it may be sometimes said that the real gains are created from "economies of scope" or "economies of network," since all the three are based on a production function of increasing returns to scale with somewhat different specifications from each other. In the case of utilizing information and telecommunication, economies of scope or network are more adequate.

pose a formal, but not exactly specified, production function:

(9) $x=f(\bar{A}, \bar{\bar{A}}, L, R)$

where x stands for the level of production, $\bar{A}$ for the firm's assets, $\bar{\bar{A}}$ for managers and other overhead personnels, L for labour directly engaged in production, and R for raw materials and other intermediate inputs. The production function, f, results in increasing returns to scale (or economies of scale) depending upon the quality of each input, production technology, optimum combination of these inputs, and managerial efficiency.

We obtain the following cost function:

(10) $\mathrm{TC}=(p_1\bar{A}+p_2\bar{\bar{A}})+[p_3\tilde{L}(x)+p_4\tilde{R}(x)]=a+bx$

where $\tilde{L}$ and $\tilde{R}$ are the equilibrium values of L and R respectively.

Reading School economits focus on how internalisation through multinational operations can make each price lower than that available through market transactions. Such a cost-accounting approach, however, may bring about an artificial, and somewhat anti-social gain.

In the following section, the above two views on the sources of gains from internalisation are illustrated in terms of some important issues of MNCs.

Hiring Corporate Staff

The business approach [see, for example, (Williamson 1975)] says that the hiring of corporate staff on a long-term contract basis (or with tenure) is itself an internalisation. (a) It reduces the transaction costs of recurrent day-by-day negotiations which are necessary to aquire workers from the market. (b) Corporate wages and salaries are set by internal pricing and become cheaper than that in an open labour market which is often under strong pressure from trade unions.

I doubt the above is true. It is most important for a firm to employ an optimal number of capable and promising personnels with a good mix of skills and ages. Doing this is risky and expensive in the recruiting and training involved. The real gains of economies of scale come from capable staff's 'learning by doing'; their loyalty to the company, and their dedication to quality control and company growth. The internal wages and salaries are even higher than or equivalent to the rates in the market or other companies; otherwise the promising staff would easily be lost to competitors.

It is difficult to keep an optimal number and mix of staff since it becomes excessive easily if the business slacks off. It is therefore usual to use part-time staff and sub-contracting arrangements as cushions. In other words, there is a limit to internalisation from the point of view of optimal economies of scale, and there is a need for dependence on, and coordination with, the outside market and other firms.

Firm-Specific Technology

Casson [1979] and others say that it is very costly and risky to acquire new production technology (more generally, information) from an open market because the market is imperfect or even missing, thus making the transaction costs high. But they do not show that the firm is able to innovate internally the needed mix of technology with a lower cost than buying a licence. It seems to me that licence fee is high because the licencer charges monopolistic price and not because of the lack of market. To buy licences is still cheaper than internal R&D efforts, particularly for late-comer small firms and enterprises in devel-

oping countries. Firms with large scale R&D (basic know-how, modern equipment and excellent researchers) produce a continuous stream of innovations, i.e., economies of scale (or scope) in R&D, which are the real gains from internalisation of innovation. There is a limit to this, however, since it is too expensive to innovate every kind of technology the firm needs. Besides the availability of a certain technology may be delayed.

Moreover, Rugman [1981] and others stressed that, as compared to the high licence fee, internal technology (or information, brand names, etc.) is transferred at lower internal price or even with no charge. This is concerned with $\bar{a}=p_2\bar{A}$, i.e., the unit fixed cost of the recurrent operating type, but not with $\tilde{a}=p_1\tilde{A}$. The operating costs incurred in giving incentives to R&D workers and in maintaining and renovating equipment and facilities are not necessarily cheaper than the licence fee.

Location-Specific Advantages

Dunning [1981, pp. 80–81] added a third source of gains from multinationalisation, location-specific advantages. Direct foreign investments—i.e., activities of "multinationalisation"—avail themselves of cheap labour, raw materials and other intermediate goods. This relates to $p_3\tilde{L}+p_4\tilde{R}=bx$ in equation (10) or direct variable costs. These costs are reduced by internal sourcing from the company's upstream plants set up at least cost supply locations.

However, it should be recognised that the real gains from multinationalisation depend, for example, upon how superior an automobile assembly plant of an affiliate company abroad is, and how large the economies of scale it can realise under the prevailing conditions in both the product and the factor markets. Overseas production results in a superior mode of production but it is not always so. (Further discussion will follow in Section VI).

Reading School economists mentioned other gains from multinationalisation. (1) Overseas production which makes it possible to bypass tariffs and other trade barriers and to save transportation costs—this is another minimisation of transaction costs. (2) Aliber [1970] mentioned that the MNC can take advantage of cheaper funds (this relates to $p_1\tilde{A}+p_2\bar{A}=a$). (3) Manipulating transfer pricing and making use of tax havens, the MNC can minimise its tax burden and maximise its profit. (4) Using liquid funds, the MNC can earn extra exchange and other financial gains. These are, however, rather secondary, transactional gains from multinationalisation. Some of them are of anti-social consequences that result from the MNC's monopolistic behaviour[6] (as will be discussed in the next section).

Market Failure and Transfer Pricing

Thus, what the Reading School economists say boils down to the following: Through internalisation and multinationalisation the cost of transactions (or production) is minimised by two factors: (a) Some of the set-up fixed investment costs are sunk. A typical example is that internally innovated technology (or information more generally) is used within the firm at lower price or without charge. This makes internally procured intermediate inputs cheaper through transfer pricing. (b) There are, especially in international markets, many market failures or imperfections. Multinationalisation makes it possible

[6] Kojima [1978] distinguishes "pseudo-economies of scale" from "genuine economies of scale." The former contributes to private profit but not to social welfare.

to set internal prices suitable to the firm's activities. In other words, in terms of our model, internalisation makes the variable cost, b, lower, by treating the set-up fixed cost as entirely sunk. Therefore, the Rugman criterion essentially depends upon "transfer pricing."

> Transfer prices are not arbitrary numbers but are the correct internal administrative prices required to make internalisation function. It is meaningless to examine transfer prices on their own, or to attempt to compare them to non-existent arm's length prices. Instead, the MNE should be allowed to use whatever transfer prices it cares to. Only its performance in producing final goods is of interest to consumer and governments. [Rugman 1981, p. 85].

This is, as said before, a superficial gain from internalisation. There is no genuine gain unless the firm realises technical economies of scale. Moreover, transfer prices are not necessarily cheaper than market prices, and all the set-up fixed investment costs can not be considered sunk.

Scale Economies of the MNC

A plant, factory, mine, plantation, headquarters office or the like is located at a place where suitable workers with lower wages, abundant raw materials of good quality, and sufficient infrastructure are available. It is designed with the minimum optimal scale and appropriate technology under the given circumstances so as to realise economies of scale (or decreasing cost) as large as possible. This is the scale economies exploited by the MNC.

The MNC can realise further economies of scale through vertical (or horizontal or conglomerate) integration of production processes and sales, as Dunning [1981, pp. 80–81] mentioned: product or process diversification, ability to take advantage of division of labour and specialisation, economies of joint supply, to control supplies of inputs with appropriate quality and quantity at best timing, and so forth. Besides, the integrated firm can use common overhead fixed assets, tangible as well as intangible, such that the average operating cost decreases up to its minimum optimal scale of transaction.

It is most important for the MNC to design scale economies in such a way (a) that MOS can be properly attained; (b) that each plant within the firm operates at the MOS and hence at the minimum average cost; and (c) that each integrated plant's capacity should produce the best proportionality that the input-output relations require. These requirements pose a limit to the extent of internalisation by MNC.

In addition to these, transportation and communication costs between the headquarters of the MNC and its scattered production plants and sales offices are the important elements to be taken into consideration.

External Economies

So far internal economies of scale (or scope) within a firm have been examined, but there are also various kinds of external economies. The trade-off between internal and external economies of scale determines the extent of internalisation through multinational operation, on the one hand, and the extent of dependence on externalities which are provided by the outside market and of cooperation with other firms, on the other.

Intra-Industry External Economies

Suppose that, for example, the automobile industry consists of assemblers A, B, and C each with a minimum optimal scale. When the total demand for cars increases, new technology suitable to a larger MOS is innovated and training of a larger number of workers is provided at the industry level. These create internal economies for the industry but external economies for each firm. These external economies do not affect the relative competitiveness of each firm and thus put each firm into monopolistic (or contestable) competition which is compatible with perfect competition.

When the total demand for cars becomes large enough for a number of assemblers to operate with large MOSs, the independent makers of specific parts and accessories are established with larger MOSs and supply their products to all assemblers with lower prices than the assemblers' internal production price. This is another type of external economies to the assemblers which is called pecuniary externalities.

When many suppliers of different parts and accessories agglomerate in a certain region centering around an assembler, external economies accumulate, in the form of secure supply of parts (saving inventories of the assembler) and reduce transportation costs. These benefits are called regional externalities or economies of agglomeration: they are further enhanced by the consolidation of business infrastructure such as general trading firms, banks, transportation and telecommunication companies, etc.

Inter-Industry External Economies

It is not economical even for a large car assembler to have its own steel mill whose MOS would be as lagre as 5 million tons per year. When a country sets up a modern steel mill which can provides excellent quality special steel at decreasing cost to the automobile industry and other users, it creates enormous inter-industry external economies or linkage externalities. It can also be said that the cheap supply of I.C. (internal circuit) of high quality plays the same role as steel does.

It is another concern whether or not an automobile firm, for example, has its own trading firm and financial institution, or whether or not an oil company has its own tankers. It is not economical for smaller firms to internalise such business infrastructures, and even for larger firms it is often more economical to use specialised outside facilities which can provide cheaper services with greater economies of scale. Efficiency in transportation and communications depends upon economies of scope created by a network of specialised large enterprises.

Well equipped infrastructure creates national scale economies which are also enhanced by such public goods as education and research, stable government and its policies, domestic and external, especially on money, trade, investment, price and wage levels, tax and tariffs, as well as foreign exchange rates.

Lastly, international public goods such as political and military security, international aid, and well functioning international monetary and trade regimes promote international scale economies.

To conclude, there is a trade-off for a firm between internalisation and dependence on market and/or on other firms and such a trade-off is primarily determined by the comparison of internal and external economies of scale. It is our concern that whether or not overseas production deprives external economies of scale from which home production

benefits (as will be discussed in Section VI).

V. *Monopoly or Competition*

Internalisation through multinational operations aims at decreasing costs through establishing an efficient hierarchy and better management. It is most advantageous for the firm to have monopolistic behaviour under decreasing cost conditions. Therefore, the theory of internalisation contributes to justify monopolistic behaviour.[7] However, in order to maximise social benefits, the firm should behave as competitively as in perfect competition so as to realise its economies of scale fully, expanding production up to and beyond the MOS, and to supply the products at minimum average cost. This is the social contribution of the firm whereas monopolistic behaviour is its sin.

Monopolistic Behaviour

In Figure 3, curve AA' shows the decreasing average cost, line mm^* representing the direct variable cost, and line SA' representing the minimum average cost consisting of Sm^* of operational fixed cost and m^*x^* of direct variable cost (these being the same as in Figure 2(ii)). In addition, curve DD_1 and MR_1 represent the demand for the firm concerned and marginal revenue respectively.

The firm prefers to limit its rate of production to x_1 with which direct variable (marginal) cost is equal to marginal revenue at point m_1. It sells its products at price as high as P_1, obtaining maximum quasi-rent (or monopoly profit) as large as P_1m_1 per product.

Total quasi-rent is area of triangle Dmm_1 which is equivalent to area Pmm_1P_1. Quasi-rent represents excess profit when the set-up fixed cost (a in equation (8)) is thought to be totally sunk and only the direct variable (marginal) cost (area mOx_1m_1) is accounted as expenditure. The firm dares to sell its products at price which covers only the direct variable (marginal) cost in order to increase its market share or when market situation is bad and unstable in situations such as volatile changes in exchange rates.

There are, however, the unsunk and recurrent fixed operating costs ($\bar{a}$) which is shown as area Bmm_1C. Because of this, the average cost for producing x_1 is at point C, and the net profit or the difference of selling price over average cost is P_1C per unit (total net profit is area $PBCP_1$).

---

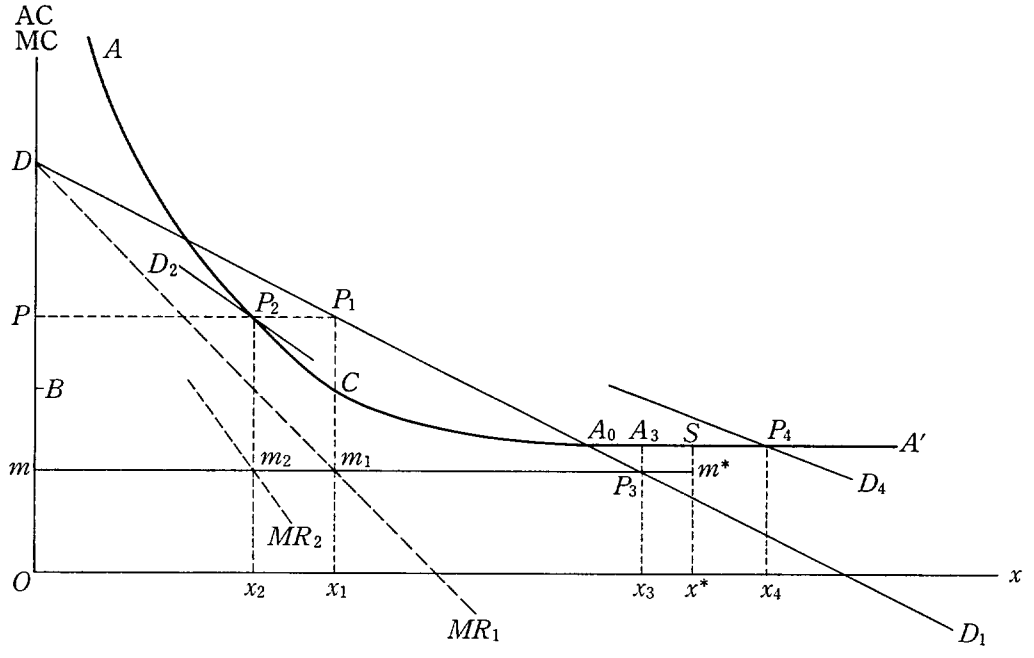
[7] First of all, Rugman [1981, p. 32] says:

> A major theme is that the multinational enterprise is ultimately an economic creature. It exists to produce and market goods and services from which it can earn profits. I found that it is not the function of the MNE to transfer technology, to act as a development agency or to redistribute income.

Buckley [1987, p. 24], on the other hand, points out both the contribution and sin of MNCs as follows:

> The theory recognises both welfare gains and welfare losses in the establishment and growth of multinational firms. Welfare gains arise where the replacement of an imperfect external market results in a superior allocation of resources internally. . . .
>
> Welfare losses arise where MNEs maximise monopoly profits by restricting the output of (high technology) goods and services and where vertical integration is used as a barrier to entry. Farther, MNEs may reduce social efficiency because they provide a more suitable mechanism for exploiting an international monopoly than does a cartel, i.e., by internalising a collusive agreement MNEs make the enforcement of collusion more effective.

FIGURE 3.



In essence, to limit sales to the volume at which marginal cost is equal to marginal revenue maximises monopoly profit. The lower the direct variable cost, b, is the larger the total monopoly profit (area of triangle Dmm_1) is. Therefore, how to make the direct variable cost, b, lower through transfer pricing is of the primary concern for the internalisationists who, in this way, justify monopolistic behaviour as the MNC's advantages.

Monopolistic Competition

When it becomes known that firm A is earning a net profit, its rival firms B, C, etc., enter the market with the same or somewhat differentiated products. The demand faced by firm A contracts from curve D_1 to D_2, and so does the marginal revenue from curve MR_1 to MR_2.

The firm is forced to lower its level of output from x_1 to x_2 at which direct variable cost is equal to the new marginal revenue at point m_2. This is still a monopolistic behaviour to maximise quasi-rent. But, because of competition with new entrants, the firm must sell at P_2 which is equal to average cost, with no net profit remaining. This way of pricing is compatible with perfect competition. Such behaviour is called monopolistic competition [Chamberlin (1933)] or contestable market [Helpman and Krugman (1985)].[8]

Maximisation of Social Benefits

Let us define social benefits as consumer's surplus. Its maximum is shown by triangle

[8] Further complicated discussions are necessary, though omitted here, to deal with the strategies of monopoly, duopoly, oligopoly, etc.

DmP_3 when price, P_3, (under a given demand curve DD_1 as before) is set to be equal to direct variable (marginal) cost and the output (and sales) is expanded to x_3. In this situation, quasi-rent (monopoly profit) becomes nil. This is called "marginal cost pricing principle" [Hotelling 1929]. For the firm, at point A_0, the break-even point, its net profit becomes zero, since price equals average cost. If the firm is forced to sell output x_3, it incurs a loss that corresponds to area $A_3P_3 \times mP_3$, and consequently the firm needs a subsidy from government of at least that amount. If the subsidy is accounted as negative social benefit, the net social benefit becomes smaller by triangle $A_0P_3A_3$ than at point A_0.

Thus, even under decreasing cost conditions, the firm should expand its output and sales much more than the level at which monopoly profit is the maximum in order to create larger social benefits. And this can be enforced when competition in the market is fierce. It is the firm's self interest to make its market share as large as possible even if it reduces or entirely loses its monopoly profit which is after all by nature anti-social.

Free Competition Behaviour

Let us use the term "free competition" instead of "perfect competition," since the latter requires too rigorous a definition to be useful and invites much confusion for internalisationists. For example, the existence of transportation costs, tariffs, international differential taxes and wages, etc. does not bring about "one price for one commodity" throughout the world. Internalisationists say that there are market failure or market imperfections (or distortions). Free competition requires that price is equal to average (=marginal) cost of sales for every firm, and leaving no marginal profit.[9] Under free competition, equilibrium is attained when the firm reaches and goes beyond the MOS output level with the minimum average (=marginal) cost. Thus the firm contributes to the maximisation of social benefits.

In Figure 3, it is supposed that the demand for the firm concerned increased from curve D_1 to D_4 due to, for example, the rise of domestic income level and/or the development of an export market. Demand curve D_4 cuts the minimum AC=MC line at P_4 with output x_4 which is beyond the MOS level of x^*.

To repeat, the real gains from internalisation through multinational operations depend upon the degree of economies of scale realised which in turn is the function of the attainable MOS, but not due to arbitrary transfer pricing and monopolistic behaviour. To attain a suitable MOS, the existence of a sufficiently large market is very important. The pattern of market structure also affects the size of demand for individual firms. If the MOS is just suitable in size for only one firm to exist in the market, its monopoly should be allowed so long as the firm adopts free competition behaviour. This is one of the central issues in the anti-trust law [Williamson (1975)].

VI. *Trade vs. Overseas Production*

Along the line of Vernon's product cycle theory [Vernon (1966)], a firm innovates a new

[9] This is a static equilibrium definition. To consider the dynamic situation, it is better to apply an uniform make-up ratio to every firm as the full-cost principle advocates.

product, first cultivates its domestic market under decreasing cost conditions, develops export, and then switches to its overseas production (i.e., DFI); in this way it expands through internalisation via multinational operations. Whether or not and why the switch is preferable is a key issue raised by Buckley and Casson [1981], who identify production mode 1 as exporting and mode 2 as DFI by Figure 1.[10] Since their presentation is somewhat misleading, I intend here to make it more exact by using our model of sunk-cost with MOS. We show that only when overseas production becomes a superior mode to home production, the former is preferable.

Home Production vs. Overseas Production

Figure 4 is the same as Figures 2 and 3 but with additions of overseas production and an export market. In the upper panel of Figure 4, $OASA''$ is the TC curve of home production inclusive of supply to domestic and foreign markets (or exporting),[11] where OA is a set-up fixed investment cost (a_α in equation (5)), the slope of line AS is the direct variable cost (b_α) and the slope of SA'' (or OA'') line is the minimum average (=marginal) cost attained (b_α*).

Let us suppose that overseas production is planned to start at point O' where home production has already expanded to this level (panel i-b). For outputs beyond O', TC of home production is shown by curve $O'A'SA''$, whereas TC of overseas production would be $O'BS'B'$ which is clearly inferior to the former judged by any criteria, whether by B-C's, Rugman's or Kojima's. The reason for this will be examined in detail below.

In the lower panel of Figure 4, as in Figure 3, the average cost of overseas production decreases along curve BS', reaching minimum average (=marginal) cost (b_β* in equation (6)) shown by line $S'B'$, and direct variable cost (b_β) by line P_1m_β. On the other hand, the average cost of home production of the firm decreases along curve AS, reaching the minimum average (=marginal) cost (b_α* in equation (5)) shown by line SA', and direct variable cost (b_α) shown by line mm*. In the case of exporting, an extra t percent sales cost caused by transportation charges, tariffs, etc. is added to the average cost curve, ASA', giving rise to dotted curve $A_tA_t'A_t''$, the export supply curve. A total demand curve, D_1, for the firm concerned consists of home country demand, D_h, and foreign demand, D_f, which is derived as the difference between curves D_1 and D_h. Therefore, the firm exports an amount P_1P_1' at price P_1 where quasi-rent is maximised but net profit is zero.

It is now clearly shown in panel (ii-b) that overseas production should not take place for its very high average cost (curve $BS'B'$) is not able to meet the foreign demand (curve D_f). Instead of resorting to DFI, the firm had better to expand home production, decreasing the average cost further and export at point P_2 with a larger volume and at lower supply price, thereby making the firm's market share larger. This is possible because economies of scale are still increasing and the average cost is decreasing.

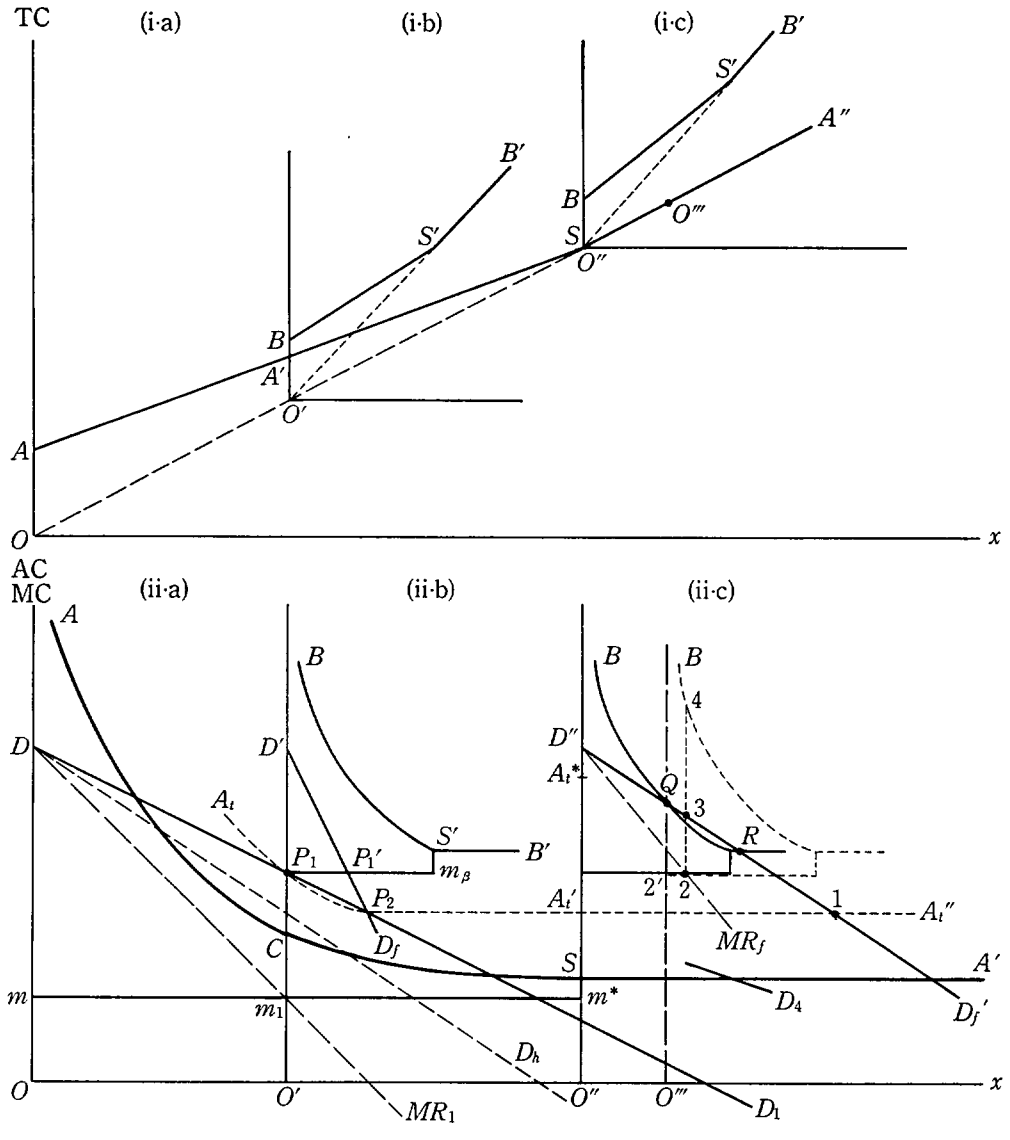
Transfer of Inferior Production Mode

We have in mind a case such as a Japanese automobile company (say, Toyota) deciding

[10] They compare with licensing production too, which is omitted in this paper in order to avoid further complications.

[11] Buckley and Casson [1981] compare exporting with overseas production, but this is misleading because exporting is only a part of total home production, especially in manufacturing.

FIGURE 4.



to set up an assembly plant in the U.S.A. (NUMMI) because of the voluntary export restraint (VER) enforced since 1981. (More generally, we want to treat DFI in manufacturing industries among advanced countries). We can conceive a number of reasons why such overseas production cannot be superior to home production.

(1) *Direct variable cost* for overseas production is more expensive than that for home production, that is $b_\alpha < b_\beta$, as against what Buckley and Casson presupposed. Concerning equation (10), (a) efficiency wage is apparently higher in U.S. than in Japan and, moreover,

labour unions' pressure is stronger in the U.S. (b) A portion of intermediate inputs (parts, components and accessories) is procured at cheaper transfer prices from the parent company and/or integrated overseas affiliates. However, production in the U.S. cannot avoid transportation charges and tariffs on the intermediate inputs imported from Japan. Moreover, the advantage of transfer pricing is often smaller in overseas production than in home production because of the strong demand to raise the local content ratio. (c) Some local procurements are available at cheaper cost in the U.S. than in Japan (for example, soybean for soy sauce factory, or scraps for a small electric steel mill set up to bypass the VER on Japanese steel exports), but mostly the reverse is true. Moreover, the external economies and agglomerated economies from which home production has benefits are lost in overseas production.

As a whole, direct variable cost in the U.S. is much higher than in Japan for automobile assemblers, even if the recent depreciation of the dollar has reduced the gap to a considerable extent.

(2) *Overhead operation cost*, $p_2\bar{A}/x^*=a/x^*$ in equation (10), is also much higher under overseas (U.S.) production than under home (Japan) production. (a) The set-up fixed cost, a, is larger in overseas production than in home production at point O', as shown in panel (i-b) of Figure 4 since much of the latter has already been sunk at that production level. But how the difference affects cost calculation is uncertain, for it depends upon how much of the fixed cost is sunk. (b) A more important element is the overhead operation cost. Compensations for Japanese expatriates and local managers definitely cost more in the U.S. than in Japan. Interest charges on set-up capital, once cheaper in the U.S., are now much higher than in Japan. Rent on factory estate and fuel and electricity charges are somewhat cheaper in U.S. but these do not weigh so heavily in the total overhead operation cost.

(3) *Comparative advantages.* The sum of (1) and (2) above is the minimum average (=marginal) cost, $b^*=a/x^*+b$, which is more expensive for overseas production than for home production in absolute terms when exchange rates and other circumstances are taken into accounts. But we have to find comparative advantages for Toyota (say firm A) in relation to Japanese rival companies (say firm B) as well as to the local Big Three (say firm C). Even if some advantages are secured by firm A from overseas production, these advantages are equally applicable to Japanese rival firm B. Besides, local firm C should be able to utilise better local location-specific advantages.

(4) *MOS.* In setting-up an overseas manufacturing plant "scaling-down" is often inevitable, making its MOS smaller than that at home. (e.g., Toyota's porduction capacity in Japan is 5 million cars per year while the NUMMI's in the U.S. is 0.2 million.) As shown in panel (i-b), scaling down makes overseas production inferior to home production as the "Kojima criterion" identifies. In order to establish a superior overseas production for firm A, its MOS not only has to be larger than that of its own home production, but also larger than that of Japanese rival firm B at home and that of each of American Big Three. But this is impossible. The only way to make the overseas plant competitive is to have it specialise in the production of only certain types of cars, thereby attaining an adequate MOS. This is the case of the Honda's American factory. This idea is explored in a theory of "agreed specialisation" [Kojima, (1987)].

Barriers to Trade and Monopolistic Incentives

Overseas production which transfers an inferior mode would not be successful from the point of view of firm concerned and would not be recommendable from our macro-economic approach. However, in practice, so many DFIs of this type between advanced industrialised countries have been promoted. Is there any justifiable reason for this?

In panels (i-c) and (ii-c) of Figure 4, we draw another situation where overseas production is decided to start after home production has reached its MOS at point O'' where economies of scale are fully realised and minimum average cost will decrease no more. It is assumed, as before, that foreign demands have increased (curve D_f') and its marginal revenue to MR_f, and accordingly the total demand curve for the firm concerned has expanded from curve D_1 to D_4. With the assumed level of tariffs (and transportation cost), the firm exports a volume represented by distance $A_t'-1$. The cost curves for overseas production remain unchanged.

Now, let us first suppose that tariffs (inclusive of transportation cost) are raised prohibitively high so as to make the export supply price to be at A_t*, higher than the prevailing market price and consequently making exports by the firm impossible. The firm is then forced to begin overseas production at point O''. If its average cost curve (B) crosses the foreign demand cyrve (D_f') within the range of the tariff wall, as shown in panel (ii-c), the firm's overseas production is able to sell to this foreign market either at point Q with quasi-rent, or at point R with larger volume at minimum average (=marginal) cost. In this case, the overseas production is induced by tariffs and works as export-substitution with a smaller supply.[12]

Alternatively, it is supposed that the firm's export quota under VER arrangements is limited to be as large as distance O''-O''', and overseas production is enforced at point O''', as illustrated by the dotted cost curves. This average cost curve does not meet the foreign demand curve, D_f'. But the firm can adopt monopolistic behaviour. At point 2, direct variable cost coincides with marginal revenue. The firm limits its production by distance 2'-2, selling at price 3, and obtains quasi-rent (monopoly profit) as large as distance 3-2 per unit, but incurs net loss as large as distance 4-3 per unit. This case tells us that in order to expand market share more than export quota, it is inevitable for the MNC to seek the level of overseas production that provides some monopoly profit. The total supply, i.e., the sum of export quota and overseas production, is again smaller than that under free trade.

Thus, overseas production, even if it is the transfer of an inferior mode, could survive behind the tariff wall, and moreover, earn monopoly profit. Therefore, such DFI is motivated to preempt the foreign market. It is, however, feasible only when the selling price is sufficiently raised due to high tariffs and/or quota restriction, or by a hidden collusion between the firm concerned and rival companies, Japanese as well as local. The result of this is a sacrifice of the welfare of American consumers who are forced to buy a smaller amount at a higher price.[13] It should also be mentioned, though often neglected, that a loss of or a decrease in exports forces the firm to contract its home production and lose its economies of scale.

It should be remembered that these adverse effects of overseas production are brought

[12] This case is fully discussed in [Kojima (1977)].

[13] The effect of VER is examined in [Kojima, (1981c)] and [Tarr and Morkre (1984)].

about by the transfer of the inferior production mode, even if it appears to be profitable as a monopolistic behaviour.

Pro-Trade Oriented DFI

In contrast, overseas production resulting in the transfer of a superior production mode makes the firm choose to abandon home production and to import the products (and/or export from an offshore plant to third countries). This is the case of "pro-trade oriented DFI."[14]

Buckley and Casson intend to explain the shift from export to overseas production by Figure 1 which, in fact, shows the profitability of offshore sourcing. In the more exact model of Figure 2, let production mode α be home production and mode β be overseas production. Up to the switching point, e, while the volume of output is small, home production is cheaper than overseas production; whereas if outputs (and demands) become larger than the point, overseas production is less expensive than home production. Therefore, the firm had better to stop home production and shift to import from its overseas affiliate. In doing so, overseas production should become superior with a larger MOS to the abandoned home production, as the "Kojima criterion" requires.

Japan is endowed with very limited natural resources, such as in oil, coal, iron ore, bauxite, etc. The MOS of overseas production becomes larger and more efficient with the abundant endowments of better quality factors abroad. Overseas production realises greater economies of scale, and lowering its average cost and forcing home production to close down and import natural resources back to Japan. A similar pattern can apply for plantation-type agribusiness products such as rubber, fish, banana, chicken and beef.

Production of labour intensive manufactures in developing countries is advantageous for the Japanese firms for two reasons. (a) Since the efficiency wage level there is very low, the direct variable cost is low. (b) In developing countries, the same machine as installed in Japan can be operated by two or three shifts of workers, making the MOS larger and the mode of production more labour intensive, enhancing the benefits of low-cost labour.

Spinning mill must attain a fairly large MOS (say 100,000 spindles per factory) while weaving and clothing are operated with small scale and simple technology. Therefore, Japanese samll firms went first to developing countries, followed by the spinning mill set up by large firms. A similar process is taking place in the automobile industry; it begins with the offshore procurement of some parts and accessories and is followed by assembly plants.

The pro-trade oriented DFI or offshore sourcing is highly beneficial in promoting industrualization in the host country and in expanding more gainful trade between the investing and the host economies.

---

[14] Pro-trade oriented DFI is defined as follows [Kojima and Ozawa (1984), p. 6]:

Proposition 1: Countries gain from trade and maximize their economic welfare when they export comparatively-advantaged goods and import comparatively-disadvantaged goods.

Proposition 2: Countries gain even more from expanded trade when superior entrepreneurial endowments are transferred through multinational corporations from the home countries' comparatively-disadvantaged industries in such a way as to improve the efficiency of comparatively-advantaged industries in the host countries.

VII. *Conclusion*

The theory of internalisation explored by Reading School economists can be simplified into a sunk-cost model with MOS. They stress that gains from internalisation through multinational operations come from the practice of transfer pricing that lowers the variable cost. This view justifies the MNC's monopolistic behaviour.

However, overseas production or DFI is profitable for the firm concerned and is simultaneously beneficial to social welfare only when it transfers a superior production mode abroad, realising greater economies of scale and lowering average cost, a process through which the real gains from internalisation can be attained. It is most important to examine whether or not overseas production results in the transfer of a superior production mode and, moreover, to assess whether or not DFI is pro-trade (or anti-trade) oriented. It is misleading and dangerous to support all types of internalisation through multinational operations merely from the point of view of monopolistic gains and without asking these critical questions.

INTERNATIONAL CHRISTIAN UNIVERSITY AND HITOTSUBASHI UNIVERSITY

REFEKENCES

Aliber, Robert Z. (1970), "A Theory of Direct Foreign Investment," in Charles P. Kindleberger (ed.), *The International Corporation: A Symposium*, M.I.T. Press.

Beamish, Paul W. (1988), *Multinational Joint Ventures in Developing Countries*, Routledge, London and New York.

Buckley, Peter J. and Mark Casson (1981), "The Optimal Timing of a Foreign Direct Investment," *Economic Journal* 91 (March), 75–87, reprinted in their (1985) *The Economic Theory of The Multinational Enterprise*, Macmillan, Chapter 5.

Buckley, Peter J. (1983a), "Macroeconomic versus International Business Approach to Direct Foreign Investment: A Comment on Professor Kojima's Interpretation," *Hitotsubashi Journal of Economics* 24 (No. 1, June).

Buckley, Peter J. (1983b), "New Theories of International Business: Some Unsolved Issues," in M.C. Casson, (ed.), *The Growth of International Business*, George Allen & Unwin.

Buckley, Peter J. (1987), *The Theory of Multinational Enterprise*, ACTA Universitatis Upsaliensis.

Casson, Mark (1979), *Alternatives to the Multinational Enterprise*, Macmillan.

Casson, Mark (1981), "Preface" to Alan M. Rugman, *Inside the Multinationals*, Croom Helm, London.

Chamberlin, E.A. (1933), *The Theory of Monopolistic Competition: A Reorientation of the Theory of Value*, Harvard University Press.

Coase, R.H. (1937), "The Nature of the Firm," *Economica*, New Series 4 (November).

Dixit, Avinash (1980), "The Role of Investment on Entry-Deterence," *The Economic Journal* 90 (March).

Dunning, John H. (1980), "Explaining Changing Patterns of International Production: In

Support of the Eclectic Theory," *Oxford Bulletin of Economics and Statistics* 41 (No. 4, November), reprinted in his (1981) *International Production and the Multinational Enterprise*, Chapter 5.

Dunning, John H. (1977), "Trade, Location of Economic Activity and the MNE: A Search for an Eclectic Approach," Betil Ohlin et al. (eds.), *The International Allocation of Economic Activity*, Macmillan.

Dunning, John H. (1988a), "The Eclectic Paradigm of International Production: A Restatement and Some Possible Extensions," *Journal of International Business Studies* 19 (No. 1, Spring).

Dunning, John H. (1988b), *Multinationals, Technology and Competitiveness*, Unwin Hyman, London.

Grunwald, Joseph and Flamm Kenneth (1985), *The Global Factory: Foreign Assembly in International Trade*, The Brookings Institution, Washington, D.C.

Helpman, Elhanan and Paul R. Krugman, *Market Structure and Foreign Trade*, The MIT Press.

Hennart, Jean-Francois (1986), "What is Internalization?," *Weltwirtshaftliches Archiv* 122 (Heft 4).

Hotelling, Harold (1929), "Stability in Competition," *Economic Journal* 34.

Hymer, Stephen (1976), *The International Operations of National Firms: A Study of Direct Foreign Investment*, The MIT Press.

Kojima, Kiyoshi (1977), "Direct Foreign Investment between Advanced Industrialized Countries," *Hitotsubashi Journal of Economics* 18 (No. 1, June).

Kojima, Kiyoshi (1978), " Giant Multinational Corporations: Merit and Defects," *Hitotsubashi Journal of Economics* 18 (No. 2, February).

Kojima, Kiyoshi (1981), "Effects of Voluntary Export Restraints of Japanese Cars to the United States" (in Japanese), *Hitotsubashi Ronso* (November).

Kiyoshi Kojima (1982), "Macroeconomic versus International Business Approach to Direct Foreign Investment," *Hitotsubashi Journal of Economics* 23 (No. 1, June).

Kojima, Kiyoshi and Terutomo Ozawa (1984), "Micro- and Macro-economic Models of Direct Foreign Investment: Toward a Synthesis," *Hitotsubashi Journal of Economics* 25 (No. 1, June).

Kojima, Kiyoshi (1987), "Agreed Specialization and Cross Direct Investment," *Hitotsubashi Journal of Economics* 28 (No. 2, December).

Linder, Staffan Berenstam (1961), *An Essay on Trade and Transformation*, John Wiley and Sons.

Marris, Robin and Dennis C. Mueller (1980), "The Corporation, Competition, and the Invisible Hand," *Journal of Economic Literature* 18 (March).

Parry, Thomas G. (1985), "Internalization as a General Theory of Foreign Direct Investment: A Critique," *Weltwirtschaftliches Archiv* 121 (Heft 3).

Penrose, Edith T. (1959), *The Theory of the Growth of the Firm*, Basil Blackwell.

Richardson, G.B. (1972), "The Organization of Industry," *Economic Journal* 83 (No. 327, September).

Rugman, Alan M. (1980a), "Internalization as a General Theory of Foreign Direct Investment: A Re-Appraisal of the Literature," *Weltwirtschaftliches Archiv* 116 (Heft 2).

Rugman, Alan M. (1980b), *Multinationals in Canada: Theory, Performance, and Economic*

Impact, Martinus Nijhoff Publishing, Boston.
Rugman, Alan M. (1981), *Inside the Multinationals*, Croom Helm, London.
Rugman, Alan M. (1985), "Internalization is still a General Theory of Foreign Direct Investment," *Weltwirtschaftliches Archiv* 121 (Heft 3).
Tarr, D.G. and M.E. Morkre (1984), *Aggregate Costs to the United States of Tariff and Quotas on Imports*, Federal Trade Commission, Washington, D.C.
Vernon, Raymond (1966), "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics* 80.
Williamson, Oliver E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, The Free Press, New York.