Abstract

International business literature is replete with empirical studies that attempt to measure the concept of internalisation. These studies use surrogate or proxy measures (e.g., R&D intensity) often based on aggregate statistics to investigate the relationships between internalisation and foreign direct investment (FDI). It is the premise of this paper that such aggregate statistics used in the calculation of proxy measures are inappropriate and should be interpreted with caution. Further, the attempt to operationalise internalisation proves difficult due to its intangible nature and may mean that proxy measures could be unsuitable. In order to advance theory, this paper advocates the use of firm level investigations based on surveys of managerial perceptions which could provide more accurate insights into the relationship between internalisation and FDI.

This paper was presented at the 28th EIBA Conference, Athens Greece, 8th-10th December, 2002

This paper is a work in progress. Material in the paper cannot be used without permission of the author.
ARE PROXIES VALID MEASURES OF INTERNALISATION?

INTRODUCTION

Internalisation theory provides an explanation of the growth of the multinational enterprise (MNE) and gives insights into the reasons for foreign direct investment (FDI). This theory (Buckley and Casson, 1976; Hennart, 1982; Casson, 1983) has been a dominant theme in the last 25 years of international business literature in relation to the growth of the MNE and FDI. The pre-eminence of internalisation as an explanation for the growth of the MNE (McManus, 1972; Buckley and Casson, 1976; Hennart, 1982; Rugman, 1982; Buckley and Casson, 1985; Casson, 1987) cannot be underestimated. Rugman (1981) argued that internalisation is a general all encompassing theory which can explain FDI. Dunning (1980; 1988; 1993) acknowledged the importance of internalisation theory by establishing it as one of the tenets of the eclectic paradigm, but argued that it explains only in part the FDI phenomenon. The internalisation path “sets out to explain the growth of international production as a market replacing activity” (Dunning, 1988).

According to Dunning (1973; 1980; 1988; 1993), a firm must possess certain advantageous conditions related to ownership (O) (firm specific advantages) that are endogenous to the firm, location (L) advantages which are external to a firm, and internalisation (I) incentives which encourage a firm to internalise operations for production to replace the need to utilize markets. Alone, these tenets are not sufficient to explain the multinational firm’s engagement in foreign production. However, all three dimensions are necessary if the role of the MNE in international production is to be understood. Rugman (1981) argued, that the concepts of ownership and location as proposed by Dunning (1980; 1988; 1993) are accounted for in the theory of internalisation.

Buckley (1988) suggested that the ownership advantage as posited by Dunning (1980; 1988; 1993) results in double counting. The ownership third of Dunning’s OLI triumvirate Buckley argues should be encompassed by internalisation advantages. For Buckley (1988), the firm seeks to carry out a strategic move by internalising the market and thus exploits this advantage in competition to others. This results in the growth of the firm (Buckley and Casson, 1976). Dunning (1995) on the other hand argues these ownership advantages are endogenous variables already belonging to the MNE, rather than exogenous as espoused by internalisation theory. In other words Dunning counters Buckley’s (1988) argument that the ownership advantage should be included within the internalisation domain arguing that ownership advantages are already possessed by firms, whereas internalisation advantages result from firms exploiting market imperfections and internalising them into firm advantages.

This paper will examine how internalisation has been measured, its relationship to transaction cost economics and the difficulties evident in studies, which attempt to measure both these constructs. The use of proxy measures to evaluate internalisation leads to a discussion on how transaction costs are perceived at the firm level. The paper suggests that surveys and interviews of managers should be conducted to provide further insights at the firm level into the thinking behind internalisation decisions. Such quantitative and qualitative studies should further our understanding of the internalisation construct.

INTERNALISATION

The essence of internalisation theory is the acknowledgement of imperfection within the market, which prevents the efficient operation of the international market in trade and investment (Buckley and Casson, 1976; Hennart, 1982; Casson, 1983). These imperfections may arise as a result of exogenous variables (externalities) in the goods or factor markets. These externalities can take the form of government induced regulations and controls, for example tariffs, or can be attributed to other types of market failure (natural externalities) such as the lack of information or knowledge. As a result, multinational enterprises (MNEs) attempt to overcome these externalities by internalising their operations. The MNE can take advantage of a unique feature by being able to maintain control over productive activities outside its national boundaries (Dunning and Rugman, 1985),
Stephen Hymer in his original dissertation (written in 1960, and published in 1976) was one of the first to show why firms engage in international production via an analysis of the MNE based upon industrial organization theory. According to Hymer (1976), Kindleberger (1969) and Caves (1971), the MNE came into existence due to ‘market imperfections’. These imperfections were ‘structural’ in nature and led to a divergence from perfect competition in the final product market. This divergence resulted from the control of ownership advantages of factors such as proprietary technology, privileged access to inputs, scale economies, control of distribution systems, and product differentiation (Bain, 1956). This meant that on the basis of these imperfections, MNEs would seek to consolidate and internalise the relationship between licensor and licensee by establishing monopolist type advantages through the vertical integration of a potential licensee (Hymer, 1976). Despite significant gains to be made in international production by internalising operations, in terms of cost reduction, improving product quality, and fostering innovation, for Hymer the fact that “... the firm internalises or supersedes the market...” (1976 p.48), highlights the issue that economic activity organised internally by the MNE is an opportunity for the MNE to further advance its monopoly power.

According to Dunning and Rugman (1985) and Casson (1987), Hymer (1976) failed to differentiate between market imperfections of the structural type (Bain, 1956) and transaction-cost type imperfections (Williamson, 1975). The transaction cost type imperfections arise naturally and are assumed to be exogenous to the MNE. The MNE responds to these transaction costs by internalising them through an internal market leading to greater efficiency. Once internalised, these transaction costs do not necessarily lead to an increase in “rent” by the MNE. However, internalising these transaction costs can result in savings for the MNE. Therefore these potential cost savings provide the impetus for MNEs to expand their operations via the internalisation of these costs.

Hymer’s failure according to Casson (1987) to distinguish clearly between the two different types of imperfections meant that he failed to relate the discussion to Coase’s (1937) theory of the firm. Transaction cost economics underpins internalisation theory and helps provide an understanding as to the establishment and development of the MNE. In order to explain the existence of the MNE it is necessary to first consider and understand transaction cost theory (Casson, 1987).

The theory of the firm developed by Coase (1937) argued that the operation of markets is not costless. The firm is an organising unit that supplants the price mechanism. Domestic firms seek to avoid the regular market by using internal prices to overcome excessive costs of an outside market. Such costs include seeking buyers and sellers, and costs involved in negotiating, co-ordinating, monitoring, and enforcing contracts. Firms avoid or disengage from the market by internalising these transaction costs where the market is non-existent or when it is cheaper for the firm to undertake these transactions internally rather than via the market mechanism. Casson (1983) refers to Coase’s analysis on internalisation as a landmark in the development of the institutional theory of the firm.

In distinguishing themselves from Hymer’s argument, and providing their own insights into transaction cost theory and the MNE, McManus (1972), Buckley and Casson (1976) and Hennart (1982), argued that ‘market imperfections’ are inherent or ‘natural’ consequences of dealing in a market because neoclassical assumptions of perfect knowledge and perfect competition cannot be realised (Dunning and Rugman, 1985). According to Buckley and Casson (1976), the MNE as a mechanism will seek to by pass these imperfections by internalising operations. The firm will exploit a market opportunity through internal operations rather than distance transactions such as licensing or franchising. This analysis gave rise to the more general internalisation approach that provides an explanation for the growth of the MNE and foreign direct investment.

In a perfectly competitive environment, the price system would organise the market. With perfect knowledge and individual honesty interdependencies between buyer and seller, transactions would be costless. In reality however such transactions in the market place do cost something due to bounded rationality and opportunism (Hennart, 1991).

According to Williamson (1975), transaction costs are those which favour hierarchical organization (internalisation) instead of markets. These costs occur because agents have bounded rationality and are
opportunistic. Transaction cost theory suggests that markets will fail due to bounded rationality (i.e. the lack of perfect knowledge which means that agents cannot foresee all possible circumstances to incorporate in the contract) (Simon, 1955) and opportunism (i.e. agents cannot be trusted, in that they make decisions based on self-interest, thus making the contract difficult to enforce) (Williamson, 1975). Market failures create transaction costs (Hennart, 1991). When transaction costs are perceived to be high companies prefer to internalise these costs rather than engage in arms-length trade.

The tacit nature of firm specific knowledge makes it difficult to transfer to outside partners. A firm is unable to articulate and quantify all the knowledge it possesses, making it potentially costly to transmit between firms (Anderson and Gatignon, 1986; Buckley and Casson, 1996). The transfer through contractual means proves difficult in terms of pricing and value (Teece, 1977; Williamson, 1985). A paradoxical situation arises. The more information revealed to convince the buyer to purchase, the less the value of that information, as the buyer now possesses knowledge that he has obtained for free (Buckley and Casson, 1976). Thus a firm specific competitive advantage such as knowledge (Casson, 1982; Caves, 1982; Dunning, 1993) once held exclusively by a firm is potentially expropriated by opportunistic behaviour. For this reason, when tacitness is high external market mechanisms become unsuitable to transfer intangibles assets such as knowledge. In such circumstances, internalised modes are preferred (Madhok, 1998).

Internalisation theory is thus very closely related to transaction cost theory (Rugman, 1981). Internalisation considers the internal operations of a multinational enterprise and thus takes into account the global arena, while transaction cost economics considers the growth of any company regardless of operating domain (Madhok, 1998). Rugman (1986) argues that internalisation is the transaction cost theory of the multinational. In other words both theories are interconnected and compatible. By seeking to understand the transaction costs phenomena within a firm, we gain a valuable insight into the internalisation process within a MNE.

The major significance of internalisation for the MNE is not that it explains the existence of the firm (transaction cost theory), but that it explains its multi-plant operation over space (Casson, 1982). The importance of vertical integration for multi-plant operations should not be underestimated. Casson (1982) suggested that there are economies and efficiencies to be gained by internalising operations. These include, long-term contracts through more efficient governance structures, R&D to prevent the dissipation of know-how which is unpatentable; tax differentials and foreign exchange controls, which create incentives for, transfer pricing. In addition, internalisation allows the firm to control quality by integrating backwards and internalising the process to maintain required standards.

MEASUREMENT OF INTERNALISATION

Despite a plethora of writing on the issue of transaction costs, their measurement has not been widely discussed (Buckley and Chapman, 1997). Transactions costs in the main are made up of factors (e.g., dissipation of knowledge) often outside the domain of quantification and straddle other costs normally associated with production or marketing (Buckley and Chapman, 1997), making them very difficult to specify. By extension a similar difficulty applies to internalisation theory in terms of measuring internal firm specific advantages.

Proxy measures have been used as a means for measuring internalisation. There are many studies in the international business literature which use proxies to measure variables that cannot be directly measured (Vernon, 1966; Horst, 1972; Dunning, 1980; Grubaugh, 1987; Kumar, 1987; Ray, 1990; Yu, 1990). Proxies are measures of an observable construct which establishes the value of a different unobservable trait. The proxies do not stand-alone but approximate an unmeasurable construct so that analysis can be undertaken. As measures, proxies are concerned with information that is conceptually quantifiable (Carley, 1981). Many researchers have sought to empirically define and accurately measure the internalisation advantage through the use of proxy measures. This has proved difficult due to two key factors. First, the intangible nature of factors such as the negotiation, maintenance and enforcement of contracts, which may help explain the internalisation advantage are difficult to quantify and second, it is difficult to empirically isolate the internalisation advantage from the ownership advantage (Denekamp, 1995). The two constructs
(internalisation advantages and ownership advantages) are therefore not mutually exclusive (Casson, 1987; Buckley, 1988; Love and Lage-Hidalgo, 1999).

A fundamental issue is whether the proxies provide an adequate measure of internalisation from which to infer behaviour. This paper argues that the use of proxy measures is often problematic in seeking to elucidate internalisation theory.

Empirical studies in international business because of their reliance on objective measures have found it difficult to develop appropriate proxies to identify the costs or the benefits associated with internal versus external action which a company may take (Agarwal and Ramaswami, 1992). In the past, measures used have led to confusion or operational problems (Dunning, 1980). The difficulty often lies in trying to measure firm specific assets that are intangible. These include, R&D, brand reputation, process innovation, or tacit information (Teece, 1977; Williamson, 1985) The nature of these assets make them difficult to value and protect in the market, thus needing to internalise to protect such firm specific advantages.

Boerner and Macher (2001) raise the concern that most empirical studies which examine transaction cost economics in international business employ advertising intensity and R&D intensity as measures of asset specificity. According to Boerner and Macher (2001) the use of such accounting proxies may be inappropriate as they confuse the effects on governance choice and as such, make results obtained difficult to interpret. Empirical studies have used a wide range of proxies to capture the asset specificity of intangible assets. Three of these proxy measures namely advertising intensity, legal intensity, and R&D intensity are discussed below.

The advertising intensity (advertising/sales) proxy has been used to approximate, management skill, and marketing skills. These are intangible resources which provide for tacit knowledge creating monopolistic advantages (Caves, 1971). Internationally they are disseminated through FDI. Advertising intensity is often used as a proxy to try and capture this tacit knowledge (Grubaugh, 1987; Kumar, 1987). This measure represents the degree of a firm’s intangible assets, in that the money which is spent on advertising and marketing generates firm specific assets in the form of brand recognition and product differentiation. An internalisation advantage is said to exist because the contracting of these assets outside the firm may see them dissipate any advantage held by the firm (Denekamp, 1995).

In his study Denekamp (1995) used the number of lawyers in an industry to proximate for legal intensity. This proxy would indicate the presence of an internalisation advantage in an industry by reflecting contractual difficulties that are presumed to be associated with intangible assets. In other words there is an inference that the relative percentage of lawyers present in an industry is proportional to the level of contracting difficulties that are present in the industry as a whole (Denekamp, 1995). Therefore, high values for the legal intensity proxy signifies the existence of contractual difficulties in an industry indicating a higher degree of intangible assets which need to be protected internally within the firm, suggesting the existence of an internalisation advantage.

R&D intensity is used in many empirical studies in the international business literature as a proxy to indicate the existence of internalisation advantages (Trevino and Grosse, 2002). R&D expenditure/sales, at the industry level is often used to indicate the firm specific assets that are created from research and development activity (Vernon, 1966; Horst, 1972; Dunning, 1980; Grubaugh, 1987; Kumar, 1987; Ray, 1990; Yu, 1990). Because of the notion of asymmetric information in R&D intensive products, the market mechanism is said to break down due to the tacit nature of knowledge. In other words, high degrees of R&D intensity indicate the existence of intangible assets which may take the form of information, patents, and technological know-how which need to be protected and internalised within the firm’s operations thus creating an internalisation advantage.

Hughes (1988) who specifically looked at the measurement of R&D intensity identified a number of problems in interpreting this measure. The validity of any proxy measure depends in part on the reliability and accuracy of the indicator used. In her research, Hughes (1988) identified that firm characteristics in terms of their internal diversification highlighted such reliability problems with the R&D intensity measure.
The extent of a firm’s diversification as reflected in its internal structure impacts on the R&D intensity measure. If the internal structure is endogenous with respect to one industry, then the R&D measure will accurately reflect the industry. However, if the firm has a diverse structure and range of products attributable to different industries, then the observed industry’s R&D intensity will be a composite quantification of the firm’s diverse R&D. Hughes (1988) found that in her analysis of industry classifications in many countries, the principal product classification of the firm determines the industry classification. R&D expenditure will be allocated to that industry despite the fact the MNE may engage in R&D in a number of product areas. In other words, measuring the R&D expenditure and apportioning it to one industry may not reflect where the R&D expenditure has actually taken place, thus compromising the R&D intensity measure.

The greater the diversification that firms exhibit in both their sales and their R&D, the more misleading it will be to collect R&D at firm level and to categorize it by principal product industry. Since R&D tends to be highly concentrated in large firms and since these large firms are diversified, the size of the error in such statistics is likely to be high (Hughes, 1988). In order to minimise distortions that are possible at the firm level ideally figures produced should be at the product level classification rather than principal product classification and sales at the firm level attributed to the specific classifications.

Another constraint in trying to understand R&D intensity is that the average figures at the industry level include firms which do not engage in any R&D. It thus makes it hard to gain any sense of what are the distributions of R&D intensities across firms within an industry. It may prove misleading to show a relationship between R&D intensity and FDI flows indicating an internalisation advantage if the R&D intensity proxy represents an industry that may have one or two dominant players in R&D and a lot of small player who undertake no R&D. The average industry figure may overstate or understate the extent of the individual firm’s internalisation advantage. As a result the proxy fails to truly reflect the impact if any of the internalisation advantages in a particular firm’s international expansion. Consequently the use of aggregate statistics in constructing any proxy measure may camouflage or inflate the true situation for the individual firm.

These proxy measures should not be disregarded; they do however need to be treated with caution. In the case of R&D intensity industry level data will remain problematic unless the underlying R&D distribution is known. Ideally absolute figures such as R&D expenditure on product classification rather than principal product classification need to be determined. Further, specific levels of sales which can be attributed to R&D need to be obtained as well as individual firm R&D expenditure to avoid reliance on industry averages.

Several researchers in the social sciences field suggest caution with the use of proxy measures. Regression analysis of the measurement of individual socio-economic status through the use of aggregate geographical data indicated that coefficients for income and education at the group level were larger than those at the individual level (Soobader, LeClere, Hadden and Maury, 2001). Again these findings highlight the problem of ‘ecological fallacy’ (Robinson, 1950). This occurs when inferences about individual situations are made based on aggregate data for a group.

INTERNALISATION AT THE FIRM LEVEL

In a perfect world, variables which explain internalisation should be calculated directly through some quantifiable measure. Where possible, accurate firm-specific data gained through primary research would be preferable to aggregate composite figures that often fail to reflect or capture the true state of the firm’s specific advantages. However, even under primary research conditions, the measurement of the internalisation concept is not necessarily an easy task. Agarwal and Ramaswami (1992) point out that indicators of the internalisation advantage have not been properly identified in the international business literature.

Buckley and Chapman (1997) suggest an alternative approach may need to be considered in order to attempt to measure and explain transaction costs. In attempting to measure and explain transaction costs, Buckley and Chapman (1997) interviewed and observed managers in the pharmaceutical and scientific instruments.
industries to investigate the process of decision-making. Managers considered decisions which affected the relationships of the parent company in regard to other companies and the extent to which the parent company needed to internalise or externalise operations.

The observation of managers in making decisions on transaction cost issues proved particularly insightful in terms of the decision making process. Among the issues faced by managers were whether R&D should be contracted out, how and when to contract out production, whether the sales force be internalised or hired, how to initiate and manage co-marketing agreements and whether to pursue economies of scale through mergers (Buckley and Chapman, 1997).

In seeking to assess the ‘viability’ of the managers’ decisions, Buckley and Chapman’s (1997) observations of managers at no stage came across any manager who had sought to numerically justify a rational assessment of the transaction cost issues mentioned above. Managers appeared to be appropriately prepared for making decisions. In making decisions they rely entirely upon informed insight into the industry in which they work. However, managers seemed to make no attempt to quantify these transaction cost issues. Buckley and Chapman (1997) concluded that managers had no understanding of the notion of transaction costs or the discourse around the issue in academic circles. However managers were engaged in decisions and actions which impacted on transaction cost issues. Further, in deciding whether to produce internally or contract out, decisions regarding costs were based on the actual costing of the two options. Issues such as reliability of supply, trust, control, issues of quality enforcement, and motivation were not assessed in any numerical sense. Decisions regarding these transaction cost issues were made on the basis of ‘judgement’, ‘gut-feeling’, ‘intuition’, ‘experience’, and ‘knowledge’ rather than any sense of systematic rational analysis (Buckley and Chapman, 1997). This was the case for a number of transaction cost problems.

As a result of their study Buckley and Chapman (1997) identified the research implications for research in this area. In seeking to consider the impact of transaction costs, company managers appear to make decisions using the information at their disposal. Decisions seem to be based on subjective assessment without any reference to any objective supporting data.

Buckley and Chapman (1997) recognise that not all transaction costs can be quantified. However, the transaction costs identified are those as perceived by the managers. Thus if a manager perceives the future costs of buying a small supply firm to be less than the future costs of continued co-operation with the small supplier, the decision might be made to buy the supplier, internalise the operations and thus reduce the transaction costs. The outcome will still be transaction cost reducing but based on perception rather than any sense of measurement of these costs. “All transaction costs are in an important sense perceptual matters” (Buckley and Chapman, 1997, p.139). In a similar vein, Agarwal and Ramaswami (1992) in their survey sought to measure internalisation directly by asking managers their perception of issues (e.g., costs of writing and enforcing contracts, risk of dissipation of knowledge) affecting internalisation. These studies lend further weight to the argument that in order to gain greater insight into the internalisation phenomena, future research will have to consider managerial perceptions in the decision to invest. Therefore the importance of managerial perceptions in decision-making should not be overlooked (Cyert and March, 1963).

**CONCLUSION**

This paper has made the case for careful consideration in the use of proxy measures which seek to capture the internalisation advantage in international business studies. Given the lack of data when it comes to measuring internalisation, the discarding of the proxy measures may not be a realistic option. However, these measures need to be used judiciously keeping in mind the pitfalls of using constructs like R&D intensity that are often calculated using industry averages. Ideally the use of firm specific indicators would be preferable if the intention is to investigate what occurs at the individual firm level in terms of international expansion.

If we are to continue using proxy measures to try and explain the internalisation advantage, greater robustness is required in terms of the measures used. The use of industry aggregates to determine concepts such as R&D intensity will remain a problem if the underlying R&D distribution is not specified. Firms
need to provide more detailed classifications as to their R&D expenditure according to product lines and related sales rather than composite firm totals. For this to happen it may be that Departments of Statistics in various countries have to start insisting on more detailed classification of R&D expenditure.

Firms may face many difficulties in trying to provide more detailed data due to issues of time and cost. Firms may simply be reluctant to undertake further record keeping, suggesting that an adequate proxy measure for internalisation could be difficult to achieve.

An alternative approach to measurement of the internalisation advantage could involve scholars and researchers finding out directly from firms, by asking managers their perceptions of factors which are likely to impact on internalisation. Managerial perceptions of these factors may be just as useful and ultimately may be a more accurate indicator of internalisation than proxy measures.

Future research should focus on longitudinal studies which seek to capture the efficacy of the managerial decisions in terms of transaction costs over time. Such studies would allow for assessment of whether the initial perception or ‘gut feeling’ resulted in an appropriate decision being made or whether with hindsight a more rational decision could have been made. This raises issues concerning what information would managers require in order to make more robust rational decisions and would outcomes be improved if quantifiable data were available?

As this paper has shown, transaction costs and by inference the factors affecting internalisation, are rarely if ever quantified in the decision making process. It is the task of the international business researcher to undertake the appropriate surveys and ask the relevant questions which will delineate and capture this phenomenon.
REFERENCES


