Theoretical and Applied Economics
Volume XXXI (2024), No. 3(640), Autumn, pp. 119-136

The relevance of theories of foreign direct investment to mergers and acquisitions in Africa

Magdalene Kasyoka WILSON
University of Johannesburg, South Africa
mwilson@uj.ac.za

Abstract. *There is no single theory that explains all the activities of multinational enterprises (MNEs). In this article, I briefly discuss selected theories relevant to developing countries and from a historical perspective. This theoretical literature review explains why international production takes place, explains the Mergers and Acquisitions (M&A) mode of entry for FDI, and demonstrates the relevance of the selected theories to M&A in Africa. The article concludes that developments in theories of FDI since the beginning of the 21st century no longer view FDI in aggregate form. To understand trends of FDI, researchers must draw from theories that consider entry mode-choice.*

Keywords: foreign direct investment, international business, mergers and acquisitions, modes of entry, Africa.

JEL Classification: F21, F23, C35.

1. Introduction

Foreign Direct Investment (FDI) and international production are mainly carried out by multinational enterprises (MNEs). Ownership alone is not what matters for FDI, but also control of value-adding activities that distinguish FDI from portfolio investment. The International Monetary Fund (IMF) and the United Nations Conference on Trade and Development (UNCTAD) suggest a minimum of 10% of ownership in equity capital or voting power to ensure control of an enterprise. In this article, and following Dunning (2009), I use international production and FDI interchangeably and assume that FDI is carried out by MNEs.

International capital movement in the 19th century and up to the end of the Second World War was mainly portfolio investment. However, the post-war period witnessed a substantial growth in FDI and activities of MNEs, especially in the United States of America (USA) and Europe as documented by Buckley and Casson (1976, pp.2-31). The neo-classical economic theory at the time could not explain the phenomenon of MNEs and a whole new approach was required. In the mid- to late- 1950s Stephen Hymer set out to explain FDI and MNEs in his PhD thesis, completed in 1960. Since then, there have been various theories of FDI.

Theories of international production draw from several fields in economics; among them international trade, industrial organisation and microeconomic theory of the firm. It is established in the literature that there is no single theory that explains all FDI and activities of MNEs. Therefore, in this article, I briefly discuss some selected theories that are more relevant to developing countries and follow the historical approach of Ietto-Gillies (2012).

The objectives of this article are to explain why FDI takes place; to explain the M&A mode of entry for FDI; and to point out the relevance of the selected theories in an African context. The article is organised as follows. Sections 2 and 3 examine theories of FDI and MNEs based on the industrial organisation approach. These include market power theory and internalisation theory, which have their roots in the Coasian theory of the firm. Section 4 discusses the eclectic paradigm, which integrates earlier theories, and concludes the purely industrial organization approach to FDI.

Section 5 deals with approaches based on international trade; that is, the new trade theory that views FDI in terms of horizontal and vertical forms. Section 6 briefly discusses more recent theories, which combine elements of industrial organization and international trade approaches. These include the assignment theory of FDI, which considers the two entry-mode choice for FDI, and the theory of M&A that specifically addresses the case of M&A entry mode. I conclude the article in Section 7.

2. Market power theory

Stephen Hymer was the first to present a theory of FDI and MNEs in his PhD dissertation (Hymer, 1960). Previously, FDI was not considered a separate field of study requiring its

own theory; instead, FDI was analysed using the theory of portfolio investment by Mundell (1960). Hymer (1960) set out to distinguish the two international capital flows: portfolio investment and FDI. The author argues that whereas portfolio investment is explained by interest rate theory, FDI is not. The reason is that interest rate theory does not explain *control* of MNEs, and therefore cannot explain FDI.

Hymer (1960) criticises the theory of portfolio investment for not being able to explain some stylised facts of FDI and activities of MNEs. In particular, the theory does not explain bi-directional flows of FDI. According to Cantwell (1991) portfolio investment theory explains less FDI between countries with similar factor proportions. Hymer (1960) argues that market imperfections are difficult to explain using the theory of portfolio investment and, therefore, the theory does not explain FDI. Hymer (1960) is credited for being the first economist to explain 'why FDI?' and 'why MNEs?' (Dunning and Pitelis, 2008) making him the father of modern theories of FDI and MNEs.

Hymer's (1960) theory is mainly concerned with why MNEs take up investment opportunities in foreign countries, other than their own countries, regardless of the costs of doing so, such as language and cultural barriers. The author suggests two main motivations for FDI. The first is, to remove competition with rival firms in foreign countries. The second is, the existence of specific advantages that the firm can profitably exploit abroad.

The motivation of removal of competition with rival firms in foreign markets was inspired by the theory of firm developed by Bain (1956), who associated incumbent oligopolists with collusive behaviour. When firms try to enter a new market and rival firms are already operating in that market, then a conflict results. A solution to this conflict would be to agree to share the market with rival firms (that is, collude or merge) or take direct control of the market (that is, acquire). Ietto-Gillies (2012) notes that the latter strategy of acquisition and control of foreign market results in further increase in market power for the acquiring firm.

The second motivation of FDI, the possession of firm-specific advantages, is directly related to the existence of market imperfections. The author observed that some firms have considerable advantages over others in particular activities. For such firms, it is more beneficial to take control (of these activities) to obtain the best returns for their skills and abilities. The author adds "unequal abilities of firms is a sufficient condition for foreign operations" (Hymer, 1960, pp.46). However, it is not a necessary condition, as Dunning and Pitelis (2008) note, because a firm can choose alternative ways to enter a foreign market such as licensing or exporting. The main reason a firm will choose direct production is market imperfections.

To conclude this section, Hymer (1960) made a 'seminal and lasting contribution' to the theory of FDI and MNEs (Dunning and Pitelis, 2008). The author identified two broad motives of FDI and MNEs. These motives are the removal of competition with rival firms and the possession of firm-specific advantages. The specific-advantages motive laid the foundation of the internalisation theory of FDI and inspired scholars such as Dunning (1980a) to come up with ownership advantages in the eclectic paradigm of FDI.

Hymer's (1960) argument that rivalry between competing firms in foreign markets can be reduced or even eliminated through M&A is still relevant in modern-day strategic asset seeking motive of firms, even in Africa. Often, we find firms competing in the same market, using advantages such as marketing and advertising to increase their market share. Such firms could benefit, for instance, by cutting down on costs in management and marketing, if they merged or acquired one another. However, unlike during the author's time, institutions such as competition tribunals have been established to control increasing market power by such firms.

Hymer's (1960) market power theory associates the existence of firms, particularly MNEs, to market imperfections. The same author explains that MNEs develop firm-specific advantages, which are best exploited within the firm, and are used to compete in foreign countries. What is also different from the author's time is the kind or type of competitive advantages that foreign firms use. For instance, emerging market MNEs, some of which are from Africa, use firm-specific advantages such as knowledge of local environment, cultural attributes, and other business practices to compete with more established firms from developed countries.

3. Internalisation theory

The basic insights into the internalisation theory of international production come from the theory of firm by Coase (1937), who observed that economic theory assumes markets, that is, the price mechanism to be the only means to allocate resources. However, the author argues that in business, activities are co-ordinated by managers of firms. Therefore, markets and firms are alternative mechanisms of resource allocation. Whereas Coase (1937) examines imperfections in domestic markets, Hymer (1960) explains imperfections in international markets. Hymer (1960) suggests that some advantages are generated within a firm, and that it is more beneficial for MNEs to internalise these advantages. However, the same author did not further explain the aspect of internalisation, which was transformed into a fully-fledged theory of FDI and MNEs by Buckley and Casson (1976).

Buckley and Casson (1976) use Coasian ideas of market imperfections to explain why MNEs internalise intermediate markets. The authors argue that the price mechanism cannot efficiently facilitate transactions involving, firstly, intermediate products and secondly, knowledge. Regarding intermediate products, the authors explain that output from firms are not just final goods and services but semi-processed materials and most importantly "types of knowledge and expertise, embodied patents, human capital" Buckley and Casson (1976, pp.33). The authors assert that the market for certain intermediate products is either difficult to organise or imperfect, and this difficulty creates an incentive to bypass the external market and use an internal one, that is, internalise the exchange. According to Rugman (1986) the main argument in the internalisation theory is that MNEs use hierarchical structures to avoid inefficient or inoperable market systems.

Buckley and Casson (1976) regard the market for knowledge to be the strongest case for internalisation. The authors explain that knowledge is a public good within the firm, whose transaction costs in the external market are much higher than those of internally selling knowledge, and, as a result, an internal market is preferred. Internal information within a firm is one type of knowledge that is difficult to sell in the external market.

Grubel (2014) explains that a bank cannot sell information about a client because the bank would not get a fair price, because of information asymmetries. The client would not like this either! The author argues that the main source of competitive advantage by subsidiaries of multinational banks is the ability to draw on information and personal contacts from the parent firm at a low marginal cost. Tschoegl (1987) argues that in multinational banking, certain factors such as economies of scale and knowledge of local conditions cannot be sold unless there is a well-functioning market. Unfortunately, certain markets such as that of knowledge do not exist because of information asymmetries. Rugman (1986) reinforces the argument of non-existence of external markets for knowledge by regarding internalisation of information as the main advantage of the multinational bank.

The other type of knowledge emphasised by Buckley and Casson (1976) is research and development (R&D) which the authors define in very broad terms to include technical- as well as market-oriented R&D. The authors argue that modern business activities not only involve routine production of goods and services but particularly important is expertise in marketing, R&D, training of labour, mobilisation, and management of finance among other factors. There is a strong incentive to internalise R&D to cover the costs of generating and transmitting this kind of knowledge, hence the internal market within a firm is considered the best option. However, Auerbach et al. (1993) criticise the underlying assumption of internalisation that a market always exists, and a manager of a firm has a choice between the external market and internal one. The point is that some products may not have markets outside the firm.

Internalisation explains international production when market imperfections occur not only in domestic markets but also across national boundaries. As Buckley and Casson (1976, pp.45) explain "there is a special reason for believing that internalisation of the knowledge market will generate a high degree of multinationality of firms. Because knowledge is a public good which is easily transmitted across national boundaries, its exploitation is logically an international operation." However, Dunning and Lundan (2008) criticise internalisation for not fully developing the international perspective of the firm. The authors argue that the theory of internalisation takes the aspect of internationalisation for granted and focuses too much on the organisation of the firm. According to Ietto-Gillies (2012) the biggest criticism of internalisation theory comes from empirical evidence in the last three decades, which has brought up the phenomenon of externalisation; that is, outsourcing and downsizing, at both domestic and international levels.

To conclude this section, I note that internalisation theory is developed from Coasian transaction cost theory of the firm. The theory capitalises on the weakness of imperfections of external markets and suggests the alternative as internal markets within firms. The

theory's focus on intangible assets of knowledge and information, sharpens our understanding of how MNEs generate firm-specific advantages, a concept developed earlier by Hymer (1960).

The main challenge with internalisation theory has been how to test the theory empirically. Buckley (1988) acknowledges that the theory cannot be tested in its general form, and only more restricted forms can be tested. The testing could most probably take place at industry level. Even so, challenges remain on how to measure variables such as transaction costs. Nevertheless, the theory is still useful in explaining multinational banking and markets for primary products such as food, minerals and oil. This makes internalisation theory much more relevant at explaining FDI in Africa which is well-endowed with minerals and agricultural products, where internal markets perform much better than external ones.

Internalisation theory contributes to the understanding of MNEs and FDI, especially investment in knowledge-intensive sectors. This kind of investment is mainly found in developed countries. The theory has also been incorporated in imperfect competition models of the new trade theory by Helpman and Krugman (1985) and Markusen (1984) as explained later in Section 5.

4. The eclectic paradigm

The eclectic paradigm is the most comprehensive and successful approach to international production and FDI (Cantwell 1991). The theory draws from previous approaches based on imperfect markets (Buckley and Casson, 1976; Coase, 1937; Hymer, 1960) and adds a third dimension of location. The paradigm asserts that the extent and pattern of international production is a function of a combination of three key factors: ownership advantages (O), locational advantages (L), and internalisation advantages (I) commonly referred to as the OLI advantage (Dunning 1980a).

Dunning (2001, pp.175) defines ownership advantages as "any kind of income-generating asset that allows firms to engage in foreign production." These advantages come from a firm's privileged ownership or access to assets that are exclusive or specific to the enterprise (Dunning 1980b). Dunning builds on earlier work by Hymer (1960) who argues that "for firms to own and control foreign value adding facilities they must possess some kind of cost or marketing advantages, sufficient to outweigh the disadvantages faced in competing with indigenous firms in the country of production" (Dunning 1988, pp.28). Perhaps Calvet (1981, pp.44) explains ownership advantage more succinctly: "foreign firms must possess a countervailing advantage over local firms" for direct investment to be viable.

Ownership advantages are specific to the firm and include tangible or Ricardian-type factor endowments like natural resources, capital and manpower, as well as intangible assets such as information, technology, marketing, organisational and entrepreneurial skills, and access to intermediate and final products (Tolentino, 2001). These ownership advantages

may be used domestically or internationally and lead to reduction in a firm's production costs, hence allowing firms to compete in foreign markets (Nayak and Choudhury, 2014). Multinational firms from developed countries mostly use financial capital and proprietary technologies as their competitive advantage, particularly in host developing countries.

Internalisation advantages (I) "relate to the way firms organise the generation and use of resources and capabilities" (Dunning 2001, pp.175). These advantages arise from the failure of external markets because of high transaction costs. Hence firms must rely on hierarchies to effectively co-ordinate their value adding activities across national boundaries. Such activities would otherwise be dispersed between different firms and countries. In this way, an advantage arises from them being co-ordinated from the parent multinational firm. Internalisation advantages explain why firms choose to exploit their ownership advantages internally, other than to acquire and/or sell them through the open market (Dunning, 2001). So, the alternative to external markets is hierarchies within a firm.

From the above discussion of both O and I advantages, it can be observed that a multinational firm not only needs to possess superior resources such as technology as Hymer (1960) argued, but MNEs must also "have the desire and ability to internalise the advantages which result from their possession" (Calvet, 1981: pp.50). Once a firm has acquired these two advantages, it can then exploit them beyond national boundaries, that is, to internationalise. This brings us to the third component of the eclectic paradigm - the addition by Dunning (1980a).

Location advantages are external to the firm and relate to the geographical and political space within which firms operate. The advantages include the cultural- legal- political- and institutional environment, government legislation, and policies (Dunning 1991 in Dunning (1995). As can be noted, locational advantages combine both home- and host- country incentives of multinational firms to invest abroad. For instance, regarding home advantage, a developing country multinational firm might have a location or geographical advantage over a developed country firm in a host developing country through similarities in culture, institutions, and business strategy among other things. As for political advantage, host governments might encourage foreign investment in or by certain countries through legislation; or these governments may give tax incentives or, create the necessary infrastructure such as industrial development zones. Therefore, location is one of the key factors that explain the geographical distribution of multinational firms. Even though a decision to invest in a particular country is firm specific, depending on firm strategy, MNEs tend to invest in a particular country or region as a group; that is, collectively. Knickerbocker (1973) called this 'a bandwagon effect', a kind of herd behavior. This phenomenon could partly explain the tendency of MNEs to flock to the East Asian region in the 1990s and to Africa in the 2000s during the commodity boom.

What location advantage attracts FDI into regions such as East Asia or Africa? Before I tackle this question, there is a need to emphasise the regional dimension of location. According to Sethi et al. (2003), MNEs may not evaluate a location based on a country but on a region, because countries in a region may have similar cultures, political and economic

systems, and development level. Therefore, location takes a wider dimension, encompassing more than one country. A region may have a configuration of locational advantages that makes it attractive to FDI. For instance, in the 1990s, many countries in Asia liberalised their economic policies, particularly regulations on FDI. In addition, the East Asian region has well-trained labour force, good infrastructure, rapidly growing economies, and a supportive banking system, and all these factors make the region attractive to investment especially efficiency-seeking type. Furthermore, relatively low wages of abundant skilled labour is important. Rapid and sustained economic growth in the region further added to the drive to invest in East Asia. FDI in Europe and Latin America has followed this kind of regional dimension to exploit the advantages of economic integration and regional similarities (Sethi et al. 2003).

The eclectic paradigm stresses the importance of each of the OLI advantages. Dunning (1998) argues that the three advantages work together like a three-legged stool that would be balanced only if the three legs were in place. Therefore, one cannot isolate an advantage such as O and claim that it is the most critical because the advantage only works in the presence of the two others. The author made this argument in response to a comment by Ethier (1986, pp.803) who claimed that internalisation is the "Caeser of the OLI triumvirate". Dunning (2001) further asserts that the way the three advantages combine varies across industries, regions or countries. Therefore, it is necessary to identify the configuration of advantages that give a specific country or industry, and a region such as Africa competitive advantages in foreign owned production.

The paradigm is *eclectic* because it partly draws from different strands of economic theory, even though the author insists that the paradigm is not bound by any doctrine (Dunning, 1986), and partly because the paradigm explains elements of alternative approaches to international production, which are direct investment, exports and contractual resource transfers or licensing (Dunning 1980b). As such, the paradigm should not be treated as a general theory because of its generality. As explained by Dunning (1988, pp.1), the eclectic paradigm should not be misunderstood as another general theory “precisely because of its generality, the eclectic paradigm only has limited power to explain or predict particular kinds of international production; even less, the behaviour of individual enterprises.”

Since 1990 Dunning (1995) and Rugman and Verbeke (2004) have attempted to adapt the OLI framework to accommodate developments such as international M&A, international joint ventures and collaborative alliances and the so-called ‘newcomers’(Ietto-Gillies, 2012). However, the eclectic paradigm does not explain some aspects of international production, for instance emerging markets MNEs that exhibit different characteristics from the ones the eclectic paradigm was initially hypothesised from.

Aykut and Goldstein (2007) and Bonaglia et al. (2007) argue that the OLI paradigm as proposed by Dunning (1980a) is based on the experiences of large, successful firms mainly from North America, Europe and Japan, which are well established and which can “easily find resources and the capabilities to expand internationally if they wished to do so” (Bonaglia et al. 2007, pp.4). However, emerging market MNEs do not have OLI advantages

such as proprietary technology, financial capital, brands, or experienced management. Indeed, MNEs from emerging markets internationalise to take advantage of the opportunities created by globalisation, to acquire the resources that enable them to grow (Aykut and Goldstein, 2007). Developing country MNEs do not have much time to develop advantages like their developed country counterparts. The MNEs cannot wait until they are large enough to internationalise and must grow quickly because protection at home is eroded by market liberalisation (Aykut and Goldstein, 2007).

As such, emerging market MNEs have developed a unique way of inter- nationalising, which is contrary to the traditional perspective. Even though some of the MNEs may be small in size, they make extensive use of linkages with other firms globally. Lau (1992) supports the idea that the size and the level of technology that developing country MNEs have is not comparable to those that developed countries of North America and Western Europe have. Therefore, the theory of FDI based on experiences of developed country MNEs may not necessarily apply to developing country MNEs.

To summarise this section, the main contribution by Dunning (1980b) is that the author brought together separate elements of previous complementary theories that attempted to explain international production. The eclectic paradigm takes Hymer’s (1960) specific advantage, combines the advantage with internalisation theory, and adds a new element of location of FDI. In this way, the OLI theory puts together three broad determinants of FDI. According to Dunning (2001), the three elements of the theory must work together. The paradigm identifies a set of factors for each of the motives for FDI; for instance, ownership advantage determines which firms will pursue foreign production. Location advantage indicates where the production will take place, and a firm will avoid market imperfections by internalising its ownership advantages.

5. The new trade theory and FDI

Since 1990, FDI has grown faster than international trade and plays an important role of facilitating trade. Navaretti et al. (2004) observed that, almost a third of world trade in intermediate goods is intra-firm – between subsidiaries of firms – and intra-industry. This observation indicates how trade and FDI have become inter-dependent and, therefore, the need to examine this relationship.

The mathematical modelling of monopolistic competition by Dixit and Stiglitz (1977) resulted in many applications in economic theory, especially in international trade theory, and specifically led to the development of the new trade theory. Previous attempts to incorporate imperfect competition in international trade models had been constrained by the assumption of constant returns to scale. However, Dixit and Stiglitz’s 1977 model enabled the analysis of increasing returns to scale and product differentiation which are important elements in understanding intra-industry trade (Brakman and Garretsen, 2008) especially in manufactured goods between developed countries.

The argument behind the new trade theory is that specialisation and trade are because of economies of scale and comparative advantage (Krugman,1985). Economies of scale are of two kinds: those internal to the firm and those external to the firm. Internal economies relate to the firm whereas external economies relate to the industry. Internal and external economies can combine with each other and accelerate the process of specialisation and concentration of industries in specific areas, resulting in spatial agglomeration effects. For instance, in Yangtze River Delta in Taipei, the Chinese government established modern infrastructure such as highways, ports and airports in the Delta, which attracted many technology firms that produce personal computers in the region creating economies of scale. The concentration of similar industries reduced service link costs, as firms in the area could easily obtain parts and components required in production (Thorbecke and Salike, 2013). This spatial agglomeration also attracted business partners in related industries, which promoted pooling skills and technologies together and further reducing costs and creating economies of scale.

There are two main strategies for FDI, horizontal and vertical. Horizontal FDI occurs when a parent firm replicates production units in other countries, whereas vertical FDI occurs when a parent firm fragments production into various stages, with each stage of production occurring in a different country, to take advantage of lower production costs. For instance, a MNE would produce a high-skill intensive product in the USA and assign production of labour-intensive input such as assembly in a low-wage country. This fragmentation of production and location in various countries enables the MNE to take advantage of, or arbitrage, on differences in labour costs.

Since the emergence of the new trade theory in the 1980s, there has been a renewed interest in re-examining theories of international production. Helpman (1985), Markusen (1984) and other authors, have attempted to integrate theories of international trade and FDI. The authors were inspired by the ideas of Hymer (1960), Buckley and Casson (1976) and Dunning (1980a) to explain international production within the framework of imperfect markets. Furthermore, Helpman (1985) and Markusen (1984) were inspired by Dixit and Stiglitz (1977) to apply monopolistic competition to international trade under the new trade theory. I therefore consider theories of FDI and MNEs in the context of the new trade theory.

5.1. The new trade theory and vertical FDI

Helpman (1983) and Helpman (1985) develops a general equilibrium model that incorporates the role of MNEs and FDI in international trade. The author makes assumptions about internal economies of scale that are generated within the firm because of utilisation of shared inputs. The joint inputs, which Helpman (1985) calls ‘headquarter services’, can be used by other firms within the organisation or transferred to such firms located in different countries at a low cost. This exchange of services leads to intra-firm trade in intangibles and vertical integration of firms within a country or across countries. The headquarter services are specific to the organisation and include services such as management, marketing, and R&D.

Krugman (1985) explains the role of MNEs in the new trade theory. Like Buckley and Casson (1976), the author argues that technology is a joint input that cannot be traded in the external market due to transaction costs. As such, services of joint inputs, or headquarter services, must be internalised within the organisation through transfer to different firms, resulting in intra- firm trade. Owing to market imperfections, firms cannot export joint inputs or license them. The firm must produce inputs directly in different countries. Hence Krugman’s (1985) model leads to vertical fragmentation of production across countries as well as intra-firm and intra-industry trade. Like Hymer (1960), the author views the main characteristic of MNEs as control of activities across borders. The author argues: “what the new models make clear, above all, is that a multinational enterprise is not a type of factor mobility. It represents an extension of control, not necessarily movement of capital." (Krugman 1985, pp. 34).

It can be argued that Helpman (1983), Krugman (1985) and Helpman and Krugman (1985) models of vertical integration of production, based on factor endowments, explain the geographical location of FDI in developing countries. The models assume factor endowments between different countries, and separate production of intermediate and final goods, which predict the production of labour-intensive products in developing countries and production of capital-intensive products in developed countries. The new trade theory also assumes that MNEs are headquartered in developed countries; for this reason, the main source of FDI is developed countries.

FDI in East Asia can partially be explained by Helpman and Krugman’s (1985) vertical integration of firms. In terms of source of FDI, East Asia has attracted high levels of investment from Europe, Japan, and the USA. Regarding factor endowment, the region has a skilled labour force, with factor prices relatively lower than in home developed countries. Production in the region is in labour-intensive products such as textiles and clothing, or/and in labour-intensive assembly operations like electronics. Thorbecke and Salike (2013) show that trade within the East Asian region is mainly in electronic parts and components (intermediate inputs) and in final goods such as computers and office equipment from the region to the rest of the world. This vertical FDI takes place through fragmentation of production, such as that of a personal computer, in different countries in the region depending on factor endowment. This production and the assembly of products have created intricate regional value chains in East Asia.

It appears that countries in East Asia are grouped into production blocks depending on their capabilities, that is, labour intensive and technology intensive. For instance, more advanced economies such as Japan, Korea, Singapore, and Taiwan belong to a block that produces high-technology intermediate products which are then exported to assembly operations in relatively more labour-intensive economies such as China, Malaysia, Thailand and Indonesia (Blattner, 2005; Thorbecke and Salike, 2013). The final goods are then exported to the rest of the world. This grouping of countries is a clear example of vertical integration of MNEs in the East Asian region and the creation of value-adding chains in the region.

Vertical FDI has also been observed in other world regions such as the North American Free Trade Area (NAFTA) between USA, Canada and Mexico. Hanson et al. (2005) find that USA MNEs fragment production depending on factor prices, trade costs and host country characteristics. Regarding factor prices, the authors find that parent companies export intermediate inputs to affiliates located in countries with relatively less skilled labour such as Mexico. Though East Asia is quite distant, USA firms still export intermediate inputs in industries such as electronics for further processing because of relatively low wages and open markets in East Asia. Another finding from the study is that intra-firm trade in intermediate inputs also takes place with countries that have lower trade costs, Canada for instance. Hanson et al. (2005) argue that vertical FDI is more prevalent in certain industries – machinery, transport equipment and electronics (especially computers). Characteristics of these industries enable production to be divided and separated into various stages such as design, component production and assembly. Factor inputs required in the production process are also different, for instance, design and component production require high-skilled labour whereas assembly of product requires low skilled labour. Owing to these different factor-input requirements, firms locate production in different countries to take advantage of factor prices.

From this subsection of the new trade theory and vertical FDI, I note that Helpman and Krugman (1985) combine elements of ownership – that is, firm-specific advantages – with internalisation elements in the form of utilisation of intermediate and intangible products. To these, the authors add locational aspects to explain the role of MNE in the prediction of trade patterns. Helpman (1983) explains that the role of MNE is to exploit factor endowments in different countries, by locating firms in relatively low-cost countries.

5.2. The new trade theory and horizontal FDI

Markusen (1984) draws from the OLI paradigm to develop a knowledge- capital model of MNE activities. The author emphasises the importance of intangible firm-specific assets which become joint inputs in multi-plant operations. These firm-specific assets are found in "human capital of the employees, patents or other exclusive technical knowledge, copyrights or trademarks, or even more intangible assets such as management, knowhow or the reputation of the firm" (Markusen 1995, pp.174). The author therefore calls these 'knowledge-based assets', which he argues are quite different from physical capital assets, because they can easily be transferred across space and supplied to additional production units at a low cost. The ease of transfer of knowledge-based assets gives them the property of a public good, which can be used in multiple plants at the same time, and in turn increases firm efficiency, by creating economies of scale.

Markusen's (1984) theory explains international production taking place between countries at a similar level of development; such countries would also have similar factor endowments. Knowledge-capital theory predicts horizontal FDI, by which similar plants are set up in different countries to produce the same products. This replication of production in different locations results in intra-industry trade and FDI. In this case, direct production is preferred to exports, because of reduction in transport and trade costs.

The knowledge-capital model has proved quite useful in explaining the direction of FDI. Empirical data shows that most FDI occurs between developed countries that have similar factor endowments. The horizontal FDI model has been empirically tested by Carr et al. (2001), Markusen and Maskus (2002) and Brainard (1993) and results are in support of the knowledge-capital theory. However, the theory does not explain the organisational aspect of FDI; that is, why firms prefer to internalise and not to license or outsource.

The new trade theory approaches to international production have been criticised for not being able to explain the I element in the OLI paradigm, which constitutes the essentials of a theory in FDI. Dunning (1995) argues that the new trade theory ignores the organisational aspect of the firm. The author argues that organising agents such as hierarchies within firms, networks, consumer groups and governments play an important role in organising economic activity. Thus, the market cannot be assumed to be the most effective way of allocating resources. The roles of such agents may sometimes complement or conflict with the market. When such agents are supportive of the market mechanism, they may create an enabling environment for the market to operate effectively. Therefore, the inadequacies of the market mechanism, have resulted in other resource allocation agents.

6. Theory of FDI and M&A

The distinction between the two subsets of FDI, greenfield investment and M&A becomes crucial at this point, as demonstrated by recent developments in FDI theory – discussed below – which have focused on the two modes of direct foreign investment, and on vertical and horizontal M&A.

Nocke and Yeaple (2008) examine the theory of vertical FDI by considering entry-mode choice. In the ‘assignment theory of FDI’ the authors argue that FDI occurs because of country differences in factor endowments. According to the theory, greenfield FDI occurs when firms take advantage of factor price differences by relocating their production from high- to low- cost countries. In contrast, M&A take place not just because of factor price differences but also because of the desire to exploit complementary firm-specific capabilities in host economies, such as distribution networks. This theory has used in the arguments for M&A in Africa by Wilson and Pholo (2019).

Neary (2007) draws from the industrial organisation literature and inter- national trade theory to come up with a theory for cross-border M&A as ‘instruments of comparative advantage.’ Unlike Dixit and Stiglitz (1977), the author proposes a general oligopolistic equilibrium model that allows for strategic interaction between firms. Strategic interaction is a crucial element of M&A activities. Brakman, Garretsen, and Van Marrewijk (2007) test the theory of M&A using sectoral data for five Organisation of Economic Co-operation and Development (OECD) countries from 1980 to 2004. The authors find that acquiring firms that come from low-cost sectors have revealed comparative advantage as measured by the Balassa index. Further, the authors find a positive autocorrelation of mergers within sectors, which suggests that M&A come in waves.

Neary (2009) further argues that the experience of 1990s, when trade liberalisation increased at the same time as FDI, contradicts the proximity- concentration trade-off, which suggests that a fall in trade costs should discourage FDI. The author claims that most FDI is horizontal, and this type of FDI should have decreased in the 1990s. Instead, horizontal FDI was encouraged within economic blocs such as the European Union (EU), by intra-bloc trade liberalisation. The reason is that foreign firms, specifically USA firms, established export platforms within the trading blocs (the EU Single Market), to serve the entire regional market.

Neary (2009) also revisits an earlier model of cross-border M&A in oligopolistic markets, which had been introduced in Neary (2003) and extends the model to explain that trade liberalisation encourages cross-border M&A. The author argues that the strategic motive of firms, a characteristic of M&A, and economic integration, lead to competition within a trade bloc and promote cross-border M&A. Neary (2009) theory has been applied in the analysis of regional integration and M&A in Africa by Wilson and Pholo (2019). The authors find that the size of market size within a regional bloc is important driver of cross-border M&A in Africa.

7. Summary and conclusion

This article reviews a selection of theories of FDI and MNEs from a historical perspective. The article presents a theoretical background to the empirical analysis on determinants of FDI by explaining the activities of MNEs. The first economist to address the question 'why FDI?' and 'why MNEs?' was Hymer (1960) in his PhD thesis. The author argued that international production does not only involve transfer of financial capital but also transfer of 'specific advantages' that MNEs possess, such as technology-, managerial- and marketing expertise. Therefore, MNEs need to own these firm-specific advantages to be able to overcome the many disadvantages of foreign production, such as unfamiliar culture and unfamiliar political and economic environments.

Buckley and Casson (1976) extend the ideas of Coase (1937) on market imperfections in domestic markets to international markets, and come up internalisation theory of FDI and MNEs. The authors explain that imperfections in cross-border markets for intermediate products and knowledge, is the primary reason why MNEs exist – to exploit internal markets within firms. However, for internal markets to be utilised, the benefits of internalisation must outweigh the higher resource, organisation and communication costs. Even though internalisation theory explains knowledge intensive FDI in developed countries, the main challenge of the theory has been how to test it empirically.

To integrate earlier theories of international production into a general theory of FDI, Dunning (1980a) suggests the OLI framework. The paradigm takes the 'specific advantage' suggested by Hymer (1960), combines the advantage with internalisation theory and adds a new element of location of FDI. In this way, the OLI paradigm puts together three broad determinants of FDI.

Theories of international production have also borrowed from international trade theory and these theories tend to focus more on the location of FDI. The location of FDI views international production in two forms; horizontal and vertical FDI, both of which are developed from internalisation theory. Two strands of literature on FDI emerge, with vertical FDI (Helpman, 1985) explaining foreign investment based on factor endowment. However, vertical FDI only explains part of FDI, such as that between developed and developing countries. There is strong evidence that most FDI is horizontal, between countries at similar levels of development and mainly developed countries. Markusen (1995) explains this phenomenon using the knowledge- capital theory, which is supported by much empirical evidence. However, studies such as Hanson et al. (2005) and other empirical evidence suggest that there is no clear-cut distinction between the vertical and horizontal FDI, and that some FDI has elements of both.

Consequently, developments in theories of FDI since the beginning of the 21st century have combined industrial organisation and international trade theories. This literature has maintained the two strands of FDI- horizontal and vertical, and has gone further to consider the two modes of foreign market entry, that is greenfield investment and cross-border M&A.

Therefore, recent theories of FDI no longer view FDI in general form, that is, aggregate FDI. To understand trends of FDI, researchers must draw from the literature that considers entry mode choice for FDI.

References

Auerbach, A.J., Hassett, K. and Slemrod, J., 1993. Taxation and foreign direct investment in the United States: A reconsideration of the evidence. *Studies in International Taxation*, pp. 119-148.

Aykut, D., and Goldstein, A., 2007. Developing country multinationals: South-south investment comes of age*, Working Paper No. 257*, OECD Development Centre, Paris.

Bain, J.S., 1956. *Barriers to new competition: Their character and consequences in manufacturing industries*. Cambridge, MA.

Blattner, T.S., 2005. *What drives foreign direct investment in Southeast Asia? a dynamic panel approach* (tech. rep.). Mimeo, European Central Bank.

Bonaglia, F., Goldstein, A. and Mathews, J.A., 2007. Accelerated internationalization by emerging markets' multinationals: The case of the white goods sector. *Journal of World Business*, *42* (4), pp. 369-383.

Brainard, S.L., 1993. *An empirical assessment of the factor proportions explanation of multinational sales* (tech. rep.). National Bureau of Economic Research.

Brakman, S. and Garretsen, H., 2008. Foreign direct investment and the multinational enterprise: An introduction. *Foreign direct investment and the multinational enterprise, edited by Steven Brakman and Harry Garretsen* (pp. 1-10). The MIT Press, Cambridge MA.

Brakman, S., Garretsen, H., and Van Marrewijk, C., 2007. Cross-border mergers and acquisitions: On revealed comparative advantage and merger waves. *Tinberg Institute Discussion Paper*, 08-013 (2).

Buckley, P.J., 1988. The limits of explanation: Testing the internalization theory of the multinational enterprise. *Journal of International Business Studies*, *19* (2), pp. 181-193.

Buckley, P.J. and Casson, M., 1976. *Future of the multinational enterprise*. Macmillan Press Limited, London & Basingstoke.

Calvet, A.L., 1981. A synthesis of foreign direct investment theories and theories of the multinational firm. *Journal of International Business Studies*, *12* (1), pp. 43-59.

Cantwell, J., 1991. A survey of theories of international production. *The Nature of the Transnational Firm*, pp. 16-63.

Carr, D.L., Markusen, J.R. and Maskus, K.E., 2001. Estimating the knowledge-capital model of the multinational enterprise. *The American Economic Review*, *91* (3), pp. 693-708.

Coase, R.H., 1937. The nature of the firm. *Economica*, *4* (16), pp. 386-405.

Dixit, A.K. and Stiglitz, J.E., 1977. Monopolistic competition and optimum product diversity. *The American Economic Review*, *67* (3), 2 pp. 97-308.

Dunning, J.H., 1980a. Explaining outward direct investment of developing countries: In support of the eclectic theory of international production. *Multinationals from developing countries, edited by K. Kumar and M. Mcleod* (pp. 1-22). Lexington Press, San Francisco.

Dunning, J.H., 1980b. Toward an eclectic theory of international production: Some empirical tests. *Journal of International Business Studies*, *11* (1), pp. 9-31.

Dunning, J.H., 1986. The investment development cycle revisited. *Weltwirtschaftliches Archiv*, *122* (4), pp. 667-676.

Dunning, J.H., 1988. The theory of international production. *The International Trade Journal*, *3* (1), pp. 21-66.

Dunning, J.H., 1995. What's wrong – and right – with trade theory? *The International Trade Journal*, *9* (2), pp. 163-202.

Dunning, J.H., 1998. Location and the multinational enterprise: A neglected factor? *Journal of International Business Studies*, *29* (1), pp. 45-66.

Dunning, J.H., 2001. The eclectic (OLI) paradigm of international production: Past, present and future. *International Journal of the Economics of Business*, *8* (2), pp. 173-190.

Dunning, J.H. and Lundan, S.M., 2008. *Multinational enterprises and the global economy*. Edward Elgar Publishing, Cheltenham.

Dunning, J.H. and Pitelis, C.N., 2008. Stephen Hymer's contribution to inter- national business scholarship: An assessment and extension. *Journal of International Business Studies*, *39* (1), pp. 167-176.

Ethier, W.J., 1986. The multinational firm. *The Quarterly Journal of Economics*, *101* (4), pp. 805-833.

Grubel, H.G., 2014. A theory of multinational banking. *PSL Quarterly Review*, *30* (123).

Hanson, G.H., Mataloni Jr, R.J. and Slaughter, M.J., 2005. Vertical production networks in multinational firms. *Review of Economics and Statistics*, *87* (4), pp. 664-678.

Helpman, E., 1983. *The multiproduct firm: Horizontal and vertical integration*. Working Paper No. 335, MIT Press, Cambridge MA.

Helpman, E., 1985. Multinational corporations and trade structure. *The Review of Economic Studies*, *52* (3), pp. 443-457.

Helpman, E. and Krugman, P.R., 1985. *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy*. MIT press, Cambridge MA.

Hymer, S., 1960. *The international operations of national firms: A study of direct foreign investment*. Massachusetts Institute of Technology, Cambridge MA [PhD dissertation published by the MIT in 1976].

Ietto-Gillies, G., 2012. *Transnational corporations and international pro- duction: Concepts, theories, and effects*. Edward Elgar Publishing, Cheltenham & Northampton MA.

Knickerbocker, F.T., 1973. Oligopolistic reaction and multinational enterprise. *Thunderbird International Business Review*, *15* (2), pp. 7-9.

Krugman, P.R., 1985. *Increasing returns and the theory of international trade*. National Bureau of Economic Research, Cambridge MA.

Lau, H.-F., 1992. Internationalization, internalization, or a new theory for small, low-technology multinational enterprise? *European Journal of Marketing*, *26* (10), pp. 17-31.

Markusen, J.R., 1984. Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics*, *16* (3), pp. 205-226.

Markusen, J.R., 1995. The boundaries of multinational enterprises and the theory of international trade. *The Journal of Economic Perspectives*, *9* (2), pp. 169-189.

Markusen, J.R. and Maskus, K.E., 2002. Discriminating among alternative theories of the multinational enterprise. *Review of International Economics*, *10* (4), pp. 694-707.

Mundell, R.A., 1960. The pure theory of international trade. *The American Economic Review*, *50* (1), pp. 67-110.

Navaretti, G.B., Venables, A. and Barry, F., 2004. *Multinational firms in the world economy*. Princeton University Press, Princeton.

Nayak, D. and Choudhury, R.N., 2014. *A selective review of foreign direct investment theories*. Working Paper No. 143, Asia-Pacific Research; Training Network on Trade (ARTNeT), Bangkok.

Neary, J.P., 2003. Globalization and market structure. *Journal of the European Economic Association*, *1* (2-3), pp. 245-271.

Neary, J.P., 2007. Cross-border mergers as instruments of comparative advantage. *The Review of Economic Studies*, *74* (4), pp. 1229-1257.

Neary, J.P., 2008. Foreign direct investment: The OLI framework. *Princeton Encyclopaedia of the World Economy*.

Neary, J.P., 2009. Trade costs and foreign direct investment. *International Review of Economics & Finance*, *18* (2), pp. 207-218.

Nocke, V. and Yeaple, S. 2007. Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity. *Journal of International Economics*, *72* (2), pp. 336-365.

Nocke, V. and Yeaple, S., 2008. An assignment theory of foreign direct investment. *The Review of Economic Studies*, *75* (2), pp. 529-557.

Rugman, A.M., 1986. New theories of the multinational enterprise: An assessment of internationalisation theory. *Bulletin of Economic Research*, *38* (2), pp. 101-118.

Rugman, A.M. and Verbeke, A., 2004. A perspective on regional and global strategies of multinational enterprises. *Journal of International Business Studies*, *35* (1), pp. 3-18.

Sethi, D., Guisinger, S.E., Phelan, S.E. and Berg, D.M., 2003. Trends in foreign direct investment flows: A theoretical and empirical analysis. *Journal of International Business Studies*, *34* (4), pp. 315-326.

Thorbecke, W. and Salike, N., 2013. *Foreign direct investment in East Asia*. Citeseer.

Tolentino, P.E., 2001. From a theory to a paradigm: Examining the eclectic paradigm as a framework in international economics. *International Journal of the Economics of Business*, *8* (2), pp. 191-209.

Tschoegl, A.E., 1987. International retail banking as a strategy: An assessment. *Journal of International Business Studies*, *18* (2), pp. 67-88.

Wilson, M.K. and Pholo, A.B., 2019. Regional integration and cross-border mergers and acquisitions in Africa. *Journal of Economic Integration*, *34*(1), pp. 109-132.