Foreign Direct Investment – A General Equilibrium Approach*

Sithanonxay Suvannaphakdy

(Received on October 27, 2010)

Abstract

Based on a general equilibrium approach, this paper presents a review of five theoretical models of foreign direct investment (FDI). These theoretical models are: MacDouglas–Kemp model (1), horizontal FDI model (2), vertical FDI model (3), knowledge-capital model (4), and gravity model (5). The paper shows that these five prominent theoretical models emerged in direct supplement from the earliest model to the latest one. Therefore, they can be represented by a single model, this model being *The Gravity Model of FDI*.

Keywords: FDI, general equilibrium, literature survey

1. Introduction

Since the past few decades, the growing economic globalization has induced business activities to flow across national borders. An important component of the flows is in the form of FDI. The questions of why and how the FDI exists have been posed from various conceptual standpoints and have therefore attracted numerous researches.

Several theoretical models have been developed to explain the emergence of FDI. This paper presents a review of five theoretical models in order to answer the following research question: what is the best theoretical model, addressing the existence of bilateral FDI flows from one country to another, based on a general equilibrium approach? These theoretical models are:[†] MacDouglas–Kemp model (1), horizontal FDI model (2), vertical FDI model (3), knowledge-capital model (4), and gravity model (5).

The remainder of the paper is organized as follows. In the second section, we provide an extensive review on FDI theory. The conclusion is given in the third section.

2. Theoretical Foundation of FDI: General Equilibrium Approach

2.1 *FDI theory based on the MacDouglas–Kemp model*

The early theoretical model, MacDouglas–Kemp model, was stemmed from the Hecksher-Ohlin model in order to explain FDI as part of international-capital trade. The model was constructed based on theoretical models by MacDougall (1960) and Kemp (1964). Assumptions of the model included full employment, perfect competition, constant returns to scale, one good, and two factors of production. FDI existed when there were differences in capital returns across countries. Capital was shifted from a relatively capital-rich country to a relatively capital-scarce country. The process continued until factor price equalization was achieved. Nonetheless, countries' welfare could be further improved by manipulating capital returns and capital flows through imposing taxes on internationally mobile capital.

The extension of MacDouglas–Kemp model was seen in Aliber (1970), who argued that FDI emerged due to the differences in capital endowments and currency risks. The differences in capital endowments and currency risks resulted in differences in capital returns across countries since premium included in interest rate was charged based on the expected currency depreciation. Firms from countries with more stable currencies could borrow money at a lower interest rate than host country firms because of their lower risk structure. Hence, this generated an incentive for foreign firm to invest in the host country.

In summary, the theoretical literature on the MacDouglas–Kemp model has explained FDI to be determined by a combination of differences in capital returns across countries, differences in capital endowments, and currency risks.

2.2 *FDI theory based on the models of horizontal FDI*

The models of horizontal FDI explained the emergence of multinational enterprises (MNEs), in which firms implemented the same activities in many countries. The incentive of horizontal MNEs could be the desire to locate production facilities close to customers and avoid trade costs (proximity-concentration hypothesis). Ownership advantages and internalization were the two broad categories that contributed to the literature of the models of horizontal FDI. Some of these literature include Hymer (1976), Kindleberger (1969), Krugman (1983),

Markusen (1984), Ethier (1986), Horstmann and Markusen (1987), Markusen and Venables (1998), and Markusen and Venables (2000).

Due to the assumption of perfect competition in the Heckscher-Ohlin model, Hymer (1976) and Kindleberger (1969) claimed that the FDI theory based on the Heckscher-Ohlin model had limited ability to explain FDI flows. In their view, FDI was supposed to be related to the theory of MNEs, which are, by definition, large firms with market power. To enter foreign markets, foreign firms required ownership advantages such as product differentiation, managerial expertise, new technology or patents, the existence of internal or external economies of scale or government interference to offset the impediments of foreign market penetration so as to compete with local firms.

Krugman (1983) incorporated the models of horizontal and vertical MNEs into the new trade theory (trade theory with imperfect competition). The combination brought about the models of horizontal and vertical MNEs. For the model of horizontal MNEs, the author claimed that horizontal MNEs emerged as a response to product differentiation. To address this claim, the model of horizontal MNEs was formulated based on a product differentiation model with an assumption that the fixed costs were not tied to the location of production. The model showed that what motivated countries to engage in international trade was the possession of different technologies by each country in the form of the knowledge of how to produce different products. Trade of this knowledge among nations could occur either directly through technology transfer within MNEs; or it could occur indirectly through trade in commodities representing special technological advantages possessed by each country. Which option would be chosen relied on the costs; transportation costs stimulated direct technology transfer, while costs of MNE operation boosted trade.

Markusen (1984) argued that horizontal MNEs emerged due to the presence of multi-plant economies of scale. 'Multi-plant economies' was referred to technical advantages owned by a single owner of two or more production facilities in an industry. Sources of multi-plant economies were from firm-specific assets which included research and development (R&D), engineering, marketing expertise, and management. These assets could serve as a public good to various production facilities of a firm at a very low cost. Marketing strategies could, for instance, be supplied to additional plants without reducing the value of them in existing plants. As a result, the presence of multi-plant economies of scale generated an incentive for the firm to produce the same goods by setting up plants in both domestic and foreign countries.

Ethier (1986) argued that horizontal MNEs emerged as a response to imperfections in con-

tracting under uncertainty.   To clarify this argument, the author combined the internalization decision into a general-equilibrium-trade model based on specific factor endowments with a differentiated manufacturing sector.   Internalization was regarded as the interaction between agents in two informational issues: public good nature of information and the size of the dispersion of information flows.   In order to examine these issues, the author constructed a model which contained two variables, research effort and product quality corresponding to the public good nature of information and the size of the dispersion of information flows, respectively. Using relative factor endowments and the degree of intrinsic uncertainty facing agents as basic parameters of the model, the author showed that each parameter was positively related to the existence of multinationals; similarity in relative factor endowments could lead to the emergence of multinationals, and the firm decided to internalize in order to respond to imperfections in contracting under uncertainty.

Horstmann and Markusen (1987) argued that horizontal MNEs emerged in industries that had higher firm-specific and tariff/transportationation costs compared with plant scale economies.   To clarify this, the authors extended the models of Markusen (1984) and Helpman (1984) to explain the existence or non-existence of horizontal multinationals.   The model contained three key elements: firm-specific costs, transportationation costs, and plant scale economies.   The combination of firm-specific costs and transportationation costs generated incentives for MNE's branch-plant production.   Plant scale economies generated incentives for centralization and serving foreign markets through exports.   As the firm-specific costs and transportationation costs became higher than plant scale economies, the firm might decide to locate its production plants close to the markets.

Markusen and Venables (1998) claimed that similarity in country size, similarity in relative factor endowments, and growth in world income encouraged horizontal MNEs to play more crucial role relative to national enterprises (NEs).   The authors illustrated this claim by constructing a general-equilibrium oligopoly model of horizontal MNEs, which allowed both NEs and horizontal MNEs to emerge endogenously.   The model consisted of two countries (home country and foreign country), two homogeneous goods, two factors of production (labor and resources), transportation costs, plant-specific fixed costs, and firm-specific fixed costs.   Labor was mobile between industries but internationally immobile.   Resources were a specific factor used only in certain industry.   There were four firm types which might or might not exist in equilibrium depending on similarities in country size and in relative factor endowments: NEs located in the home country, NEs located in the foreign country, MNEs headquartered in the

home country, and MNEs headquartered in the foreign country. The results from comparative statics analysis of the model showed that a ceteris paribus increase in the world income led to a rise in markup revenues of the MNEs than that of the NEs due to the presence of the transportation costs; and differences between countries in size and in relative endowments were disadvantageous to MNEs relative to NEs. The numerical simulations of the general equilibrium model also confirmed the same results.

Markusen and Venables (2000) argued that the presence of positive trade costs induced the emergence of MNEs. To explain this issue, the authors generalized a general-equilibrium monopolistic-competition model developed in Helpman-Krugman (1985) by allowing trade costs and endogenous MNEs. Ingredients of the model were the $2 \times 2 \times 2$ Hechscher-Ohlin structure, monopolistic competition, iceberg trade costs, and MNEs. Using the generalized model, the authors showed that trade costs modified the trade pattern, generated incentives for factor mobility which might cause agglomeration of activity, and might result in the emergence of MNEs.

In summary, the theoretical literature on the horizontal FDI has explained FDI to be determined by a combination of product differentiation, the presence of firm-specific costs, transportation costs, tariffs, similarity in country size, similarity in relative factor endowment, and growth in world income.

2.3 *FDI theory based on the models of vertical FDI*

The models of vertical FDI explained the emergence of MNEs, in which firms geographically segmented stages of production. The fundamental idea behind vertical MNEs was that they were able to benefit from factor-price differences in the world economy, by relocating a final production plant with unskilled-labor-intensive activity in a relatively unskilled-labor-abundant country. The implication that the emergence of vertical MNEs was to exploit foreign cheap labor brought about a positive relationship between vertical MNEs and cross-country differences of factor proportions. Some of these literature included Krugman (1983), Helpman (1984), Helpman (1985), Helpman and Krugman (1985), and Zhang and Markusen (1999).

For the model of vertical MNEs, Krugman (1983) claimed that vertical MNEs emerged in order to eliminate the distortion of input prices supplied by other firms. To illustrate this claim, the author formulated the model of vertical MNEs based on Perry's (1978) model of vertical integration due to monopsony. The model assumed that the firm could geographically fragment its production process. Since an attempt to maintain a lower price of raw material by

a monopsonistic down-stream firm distorted the production decision of its foreign suppliers, its foreign suppliers would supply too little.   Using the constructed model, the author showed that the distortion could be eliminated if the firm was vertically integrated.   The author also showed that profits of the vertically integrated firm were larger than the combined profits of independent upstream and downstream firms.

Helpman (1984) argued that vertical MNEs emerged in order to exploit differences in factor rewards across countries.   To illustrate this claim, the author combined the idea of firm-specific input into a general-equilibrium-monopolistic trade model.   It was assumed that technology of production was characterized by increasing returns to scale and firm-specific inputs could serve plants located in foreign countries.   Since the model assumed that there was no transportation costs, firm would concentrate its production in a single plant due to increasing returns.   As differences in factor prices existed across countries, firm would become a vertical MNE by relocating its production plant to a lower cost country, say a lower wage country, so as to maximize its profits.

The extension of Helpman's (1984) model was found in Helpman (1985) and Helpman and Krugman (1985, Chapter 13).   Helpman (1985) integrated the horizontally and vertically integrated firms from Helpman (1983) into the theory of trade and direct investment in Helpman (1984).   The combination resulted in a theory of international trade in which horizontally and vertically integrated firms, which have plants in more than one country, trade in finished goods, intermediate inputs, and invisibles.

Helpman and Krugman (1985, Chapter 13) explained trade patterns associated with the emergence of multinationals by combining the general equilibrium trade framework into the model of vertically integrated firms derived in Helpman (1983).   Using the extended model, the authors showed that the existence of multinationals made trade flows to consist of three components: inter-sectoral trade, intra-industry trade, and intra-firm trade.   Trade volume was determined by the relative country size and the difference in relative factor endowments.   However, the effect of the relative country size on trade volume would be lower, the larger the role of multinationals in the world economy.

Zhang and Markusen (1999) claimed that FDI into a small least-developed country (LDC) approached zero as difference in factor endowments between the developed country (DC) and small LDC became extremely large.   To explain this issue, the authors extended the work of Markusen and Venables (1996, 1997, and 1998).   The extension focused on the direct factor requirements, which involved MNEs' needs for local skilled-labor such as managers, techni-

cians, engineers, and accountants.　The extended model was a type of vertical MNEs in which there were FDI flows from DC into LDC in order to exploit lower wage rate in relatively unskilled-labor-abundant countries.

The model developed in Zhang and Markusen (1999) was a two-country (skilled-labor-abundant country representing DC and unskilled-labor-abundant country representing LDC), two-goods, and two-factor (skilled labor and unskilled labor) model.　One kind of goods was produced with skilled and unskilled labor under constant returns to scale and perfect competition.　Another kind of goods was produced in two stages under increasing returns and imperfect competition.　In the first stage, the intermediate good was produced with only skilled labor.　In the second stage, the final product was produced with unskilled labor and the intermediate goods.　Both intermediate and final goods could be produced by two types of firms: the first type was NEs located in DC; and the second type was the vertical MNEs which located their headquarters in the DC and located their production plants in the LDC.

As the differences in factor prices existed across countries, firms in the DC would become vertical MNEs in order to seek a lower cost country for MNEs operation.　However, the model predicted that FDI into a small LDC approached zero as difference in factor endowments between the DC and small LDC became extremely large.　The reason was that most output produced by the potential plant would be transportationed back to the DC, thus incurring very high aggregate transportation costs relative to a situation in which all output was produced in the DC. The model also predicted that country size played no role for the emergence of vertical MNEs as the goods had low transportation costs and small scale economies.

In summary, the theoretical literature on the vertical FDI has explained FDI to be determined by a combination of the distortion of input prices, cross-country differences in factor endowments and in factor prices, transportation costs, and scale economies of the goods produced.

2.4　*FDI theory based on the knowledge-capital model*

The knowledge-capital model was referred to a model in which firm fixed costs in its technology of production were described by relatively low costs of geographically fragmenting headquarters and a single plant, skilled-labor intensity of firm fixed costs relative to production, and the jointness of firm fixed costs across multiple plants.　It helped to predict both horizontal and vertical FDI.　Some of these literature include Markusen (1997) and Markusen (2002, Chapter 8).

Markusen (1997) argued that abundant factors in certain industries of a relatively skilled-labor-scarce economy could be exploited through knowledge-intensive producer services provided by direct investment.   To clarify this issue, the author constructed a theoretical model which consisted of two homogeneous goods, two factors of production (unskilled labor and skilled labor), and two countries.   One kind of goods was produced with constant returns under perfect competition, and another kind of goods was produced with increasing returns at both the firm and plant level.   There were six possible firm types, each defined by the locations of plants and headquarters.   These firm types included horizontal MNEs, vertical MNEs, and national firms.   In equilibrium, only certain types of firms existed depending on difference in country size, difference in relative factor endowments, transportation costs, and costs of MNE operation.

Using the constructed-general-equilibrium model, Markusen (1997) numerically simulated four scenarios.   The first scenario with high protection of trade and FDI was used as the base case.   In the base case, there were only national firms, but most of them located in a large and skilled-labor-abundant country.   The second scenario was trade liberalization which showed that all national firms located in a small and unskilled-labor-abundant country were driven out. The third scenario was the investment liberalization which showed that horizontal MNEs existed.   The emergence of horizontal MNEs expanded production into a small and unskilled-labor-abundant country, and raised the real and relative price of skilled labor in both the home and host countries.   The fourth scenario was the liberalization in both trade and investment which showed that the vertical MNEs existed as countries differed in relative endowments.

Extension of the analysis of the knowledge-capital model was found in Markusen (2002, Chapter 8) who investigated the relationship between affiliate production and trade in goods produced with increasing returns to scale, the relationship between factor prices and the degree of liberalization in trade and investment barriers, and welfare effects of liberalization in the case of asymmetric countries (differences in GDPs and in relative factor endowments).   Using the numerical simulations of the general-equilibrium model, these issues were considered in any four scenarios, differing in the degree of restrictions on trade and restrictions on the existence of multinational firms.   These scenarios were no liberalization (high trade costs and prohibition of FDI), trade liberalization (low trade costs and prohibition of FDI), investment liberalization (high trade costs and permission of FDI), and full liberalization (low trade costs and permission of FDI).   The first scenario with no liberalization was used as a base case.

Firstly, the result on the relationship between affiliate production and trade showed that in

the case of trade liberalization, lower trade costs encouraged vertical FDI, while discouraged horizontal FDI; the result of investment liberalization was that trade and horizontal FDI were substitutes; and the result of the full liberalization was that trade and vertical FDI were complements. Secondly, the effects of investment liberalization on factor prices had a pro-skilled-labor bias. That is, resources in both countries were transferred to a more skilled-labor-intensive industry. Thirdly, the effects of investment liberalization on welfare were that the small country was assured of gains from investment liberalization, while the larger country might lose for some differences in size and in relative endowments. By considering different degrees of trade costs, the author then concluded that the host countries were ensured of gains, whereas the parent countries could lose from investment liberalization.

In summary, the theoretical literature on the knowledge-capital model has explained FDI to be determined by a combination of differences in relative factor endowments, differences in country size, transportation costs, costs of MNE operation, and tariff barriers.

2.5 *FDI theory based on the gravity model*

The gravity models of trade and FDI predicted and explained bilateral trade flows and bilateral FDI flows (stocks) in terms of the economic size and distance between trading partners. The theoretical rationale behind the application of the gravity model to FDI was found in Bergstrand and Egger (2007, 2010).

Based on the results of the empirical gravity equations of trade and FDI, Bergstrand and Egger (2007) claimed that theoretical rationale for estimating the gravity models for bilateral flows of trade, FDI, and foreign affiliate sales (FAS) could be derived. To explain this idea, the authors introduced a third internationally-mobile factor (physical capital) and a third country (the rest of the world, ROW) to the standard $2 \times 2 \times 2$ "knowledge-capital" model of MNEs in Markusen (2002). The extended model became a three-country (home country, foreign country, and ROW), three-factor (skilled labor, unskilled labor, and physical capital), and two-good (homogeneous goods and differentiated goods) general equilibrium model of multinational and national firms ($3 \times 3 \times 2$ model). Using the $3 \times 3 \times 2$ model, the authors first showed that the presence of the physical capital—combined with an assumption that the establishment of headquarters in any countries required home country's skilled labor; and the establishment of plants required home country's physical capital—provided explanation for the coexistence of intra-industry trade and intra-industry FDI (as well as FAS) for two identically-sized developed countries for a wide range of parameter values. Second, the presence of both a third country and

physical-capital mobility in the $3 \times 3 \times 2$ model explained the relationship between a pair of countries' GDP similarity and the complementarity of bilateral trade, FDI, and FAS. The reason behind this idea was that the "complementarity effect" generated by the endogenous relative prices of physical-to-human capital interacting with the three countries' GDP potentially counterbalanced the "substitution effect" associated with exogenous trade-to-investment costs. To evaluate numerically the relationship the bilateral flows – trade, FDI, and FAS – and GDP sizes and similarities, the authors applied the numerical version of the general equilibrium model and the regression models. The results of both methods confirmed that the economic determinants of bilateral trade, FDI, and FAS should be "well-approximate" by gravity equations.

The extended version of the "knowledge-and-physical-capital" model was found in Bergstrand and Egger (2010) who argued that bilateral flows of final goods trade, of intermediate goods trade, and of outward FDI could be explained by the gravity equation. To illustrate this issue, the authors included intermediate goods into the $3 \times 3 \times 2$ "knowledge-and-physical-capital" model. The extended model became a three-country (home country, foreign country, and ROW), three-factor (skilled labor, unskilled labor, and physical capital), and three-goods (homogeneous goods, differentiated goods, and intermediate goods) general equilibrium model of multinational and national firms. Based on the numerical simulations of the $3 \times 3 \times 3$ general equilibrium model and the regression models, the authors formulated a theoretical rationale for estimating simultaneously gravity equations for bilateral flows of final goods trade, of intermediate goods trade, and of bilateral outward FDI. In order to test the performance of the theoretical gravity equations of these bilateral flows, the authors applied these gravity equations to estimate the empirical data. The results confirmed that they explained very well the empirical final goods trade, intermediate goods trade, and FDI flows from one country to another.

In summary, the theoretical literature on the gravity model has explained FDI to be determined by a combination of country size and distance between trading partners.

3. Conclusion

Several theoretical models have been constructed to explain the existence of FDI and its location decision. The MacDouglas–Kemp model explained FDI as part of international-capital trade due to differences in returns on capital, but the model was criticized because of its assumption of perfect competition. Under the assumption of imperfect competition, FDI was explained by the models of horizontal FDI using the proximity-concentration hypothesis,

whereas it was explained by the models of vertical FDI using the factor-proportions hypothesis. The knowledge-capital model provided a generalized model which allowed both forms of horizontal and vertical FDI to exist as its special cases. For empirical purpose, the knowledge-capital model has been modified to be a theoretical gravity model of FDI.

There are many factors determining FDI. These factors include differences in capital returns across countries, differences in capital endowments, currency risks, market size, growth, trade barriers, investment barriers, factor costs, plant economies of scale, and transportation costs.

Based on the literature surveyed, five prominent theoretical models emerged in direct supplement from the earliest model to the latest model. These are *The MacDouglas–Kemp Model, The Horizontal FDI Model, The vertical FDI Model, The Knowledge-Capital Model*, and *The Gravity Model of FDI*. The prominent authorities in the field of international economics posited the following theoretical models on the general equilibrium theory of FDI. When critiqued, these theoretical models can be represented by a single model, this model being *The Gravity Model of FDI*.

Acknowledgement

Bibliography

Aliber, R. Z. (1970). A theory of direct foreign investment. In C. Kindleberger (Ed.), *The international corporation.* Cambridge, MA: MIT Press.

Bergstrand, J. H. & Egger, P. (2010). A general equilibrium theory for estimating gravity equations of bilateral FDI, final goods trade, and intermediate trade flows. In P. A. van Bergeijk & S. Brakman (Eds.), *The gravity model in international trade* (pp. 29–70). Cambridge: Cambridge University Press.

Bergstrand, J. H. & Egger, P. (2007). A knowledge-and-physical-capital model of international trade flows, foreign direct investment and multinational enterprises. *Journal of International Economics*, *73*, 278–308.

Ethier, W. J. (1986). The multinational firm. *The Quarterly Journal of Economics*, *101* (4), 805–834.

Faeth, I. (2009). Determinants of foreign direct investment - A tale of nine theoretical models. *Journal of Economic Surveys*, *23* (1), 165–196.

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *The Journal of Political Economy*, *92* (3), 451–471.

Helpman, E. (1985). Multinational corporations and trade structure. *Review of Economic Studies*, *52*, 443–458.

Helpman, E. & Krugman, P. (1985). *Market Structure and Foreign Trade.* Cambridge, MA: Cambridge University Press.

Horstmann, I. J. & Markusen, J. R. (1987). Strategic investments and the development of multinationals. *International Economic Review*, *28* (1), 109–121.

Hymer, S. H. (1976). *The international operations of national firms: A study of direct investment.* Cambridge, MA: MIT Press.

Kemp, M. C. (1964). *The pure theory of international trade.* Englewood Cliffs, NJ: Prentice Hall.

Kindleberger, C. P. (1969). *American business abroad: Six lectures on foreign direct investment.* New Haven, CT: Yale University Press.

Krugman, P. R. (1983). The 'new theories' of international trade and the multinational enterprise. In C. P. Kindleberger & D. B. Audretsch (Eds.), *The multinational coporation in the 1980s* (pp. 57–73). Cambridge, MA: MIT Press.

MacDougall, G. D. (1960). The benefits and costs of private investment from abroad: A theoretical approach. *Economic Record*, *36*, 13–35.

Markusen, J. R. (2002). *Multinational firms and the theory of international trade.* Cambridge, MA: MIT Press.

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

Markusen, J. R. (1997). Trade versus investment liberalisation. NBER working paper no. 6231. Cambridge, MA: National Bureau of Economic Research.

Markusen, J. R. & Venables, A. J. (1998). Multinational firms and the new trade theory. *Journal of International Economics*, *46*, 183–203.

Markusen, J. R. & Venables, A. J. (1997). The role of multinational firms in the wage-gap debate. *Review of International Economics*, *5*, 435–451.

Markusen, J. R. & Venables, A. J. (1996). The theory of endowment, intra-industry, and multinational trade. NBER working paper no. 5529.

Markusen, J. R. & Venables, A. J. (2000). The theory of endowment, intra-industry, and multi-national trade. *Journal of International Economics*, *52*, 209–234.

Perry, M. (1978). Vertical integration: The monopsony case. *American Economic Review*, *68*, 561–570.

Zhang, K. H. & Markusen, J. R. (1999). Vertical multinationals and host-country characteristics. *Journal of Development Economics*, *59*, 233–252.