A REVIEW OF ECONOMIC LITERATURE
ON FOREIGN DIRECT INVESTMENT

David Riker
Heather Wickramarachi

ECONOMICS WORKING PAPER SERIES

U.S. INTERNATIONAL TRADE COMMISSION
500 E Street SW
Washington, DC 20436

April 2020

Office of Economics working papers are the result of ongoing professional research of USITC Staff and are solely meant to represent the opinions and professional research of individual authors. These papers are not meant to represent in any way the views of the U.S. International Trade Commission or any of its individual Commissioners. Working papers are circulated to promote the active exchange of ideas between USITC Staff and recognized experts outside the USITC and to promote professional development of Office Staff by encouraging outside professional critique of staff research. Thanks to John Fry for helpful suggestions. Please address correspondence to david.riker@usitc.gov or heather.wickramarachi@usitc.gov.
A Review of Economic Literature on Foreign Direct Investment  
David Riker and Heather Wickramarachi  
April 2020

Abstract

We review the theoretical and empirical literature on the economics of foreign direct investment (FDI). We summarize the significant evolution of economic research on this topic over the last five decades, from descriptive industrial organization studies to detailed econometric models with firm heterogeneity. We highlight the key economic concepts in the literature.

David Riker, Office of Economics  
david.riker@usitc.gov

Heather Wickramarachi, Office of Industries  
heather.wickramarachi@usitc.gov
1 Introduction

The theoretical and empirical literature on the economics of multinationals and foreign direct investment (FDI) has evolved significantly over the last five decades. Up through the 1970s, research on multinational firms consisted mostly of descriptive industrial organization studies. From the late 1970s to the middle of the 1990s, there were significant developments as the literature started to apply economic models of international trade with product differentiation and imperfect competition to the study of multinationals. Starting in the middle of the 1990s, there was significant growth in empirical analysis of FDI, as researchers gained access to firm-level data on the operations of multinationals. In the 2000s, there was another surge of developments in the theoretical modeling of FDI, with the addition of firm heterogeneity and firm-level data. The focus of the models in the 2000s shifted to organizational choices – whether multinational production is internalized within a firm or coordinated through arms'-length outsourcing – and continued the emphasis on firm-level data in empirical work. This work on the boundaries of firms builds on early work on transaction costs and the theory of the firm by incorporating insights from the modern literature on incomplete contracting, with its emphasis on agency and incomplete information.

Along the way, the literature has identified several key economic concepts that are important for understanding FDI. We outline these concepts briefly in the Introduction and then in the rest of the paper we trace their development in the economic literature.

• Firm-specific assets, like unique technologies, brands, and reputations for quality, are leveraged globally, and this allows a multinational firm to achieve firm-specific economies of scale.

---

2 Examples include Helpman and Krugman (1985) and Markusen (1995).
3 Yeaple (2003a), Branstetter, Fisman and Foley (2006), Keller and Yeaple (2009), and Harrison and McMillan (2011) are examples of econometric analyses of firm-level data.
4 Helpman, Melitz and Yeaple (2004) is a leading example of this branch of the literature.
5 Helpman (2006) discusses this organizational choice branch of the literature.
• Strategic pricing above marginal costs is a common feature of the concentrated industries that account for most FDI.

• There are several different types of FDI, including horizontal FDI that jumps tariff or other barriers to cross-border trade, vertical FDI that establishes global supply chains to take advantage of international factor cost differences, and foreign production that serves as a platform for exporting to the domestic market or other foreign markets.

• Incomplete contracting and technology transfer can explain why internalization within a multinational firm is more profitable than outsourcing.

• Detailed firm-level data now dominate the econometric literature on FDI.

There are several excellent surveys of the economics literature on FDI and multinationals, including a recent chapter by Pol Antràs and Stephen Yeaple in the fourth volume of the Handbook of International Economics. This paper covers most of the same ground in a less technical way, placing greater emphasis on the usefulness of the literature as a framework for analyzing the impact of international trade and investment policies.

The rest of this paper is organized into six parts. Section 2 focuses on the importance of firm-specific assets and market power. Section 3 discusses different types of FDI. Section 4 focuses on incomplete contracting and the factors that determine the boundaries of firms. Section 5 reviews empirical analysis of FDI. Section 6 addresses the link between FDI and specific trade and investment policies.

2 Firm-Specific Assets and Market Power

Most of the early literature, dating back to the 1970s, was concerned with why firms shift production across borders and what types of firms will choose to do so. The first research in this area took an industrial organization (IO) approach that is based on an analysis of

See Antrás and Yeaple (2014). Other excellent reviews of the literature include Markusen (2002) and Helpman (2006).
firm-specific assets that give multinational firms a strategic advantage over domestic firms in foreign markets.

In one of the seminal contributions to the IO literature, Dunning (1981) developed the Ownership, Location, and Internalization framework that identifies factors that can influence a firm’s decision to become a multinational. Ownership advantages focus on firm-specific assets that make a firm competitive in a foreign country, including technology and patents, processes, and cost advantages. Location refers to location-specific advantages in the production process, for example to take advantage of international differences in costs and natural resource availability. Internalization relates to a firm’s decision to control multiple stages of production within the same firm, or to contract for some stages at arms-length. There are costs associated with both, including organization costs of the firm and transaction costs of operating at arms-length.

Markusen (1984) supplements the analysis in Dunning (1981) with formal economic modeling. He reconciles the activities of multinationals with trade theory by focusing on technological advantages. He develops a technology-based general equilibrium model of multinationals. These firms have firm-specific activities that allow them to innovate, such as R&D and marketing, and this innovation can then be deployed to other production facilities without having to redevelop the technology for each market.

In Markusen (1997), he expands the theoretical framework further to show that trade and investment liberalization can be complements. He introduces his well-known knowledge-capital model, which he supports with empirical evidence in Carr, Markusen and Maskus (2001). The knowledge capital model essentially explains the volume of production of MNC affiliates in a host country as a function of host and source country characteristics (i.e.,

---

7Markusen (1984) is primarily a model of horizontal FDI (discussed further in section 3, below), but because there is some form of specialization in each location it can be described as ‘hierarchical.’

8This model combines horizontal motivations for FDI with vertical motivations and was further synthesized in Markusen (2002).
market size, transport costs, trade and investment barriers). There are three assumptions:
1) knowledge-intensive activities (R&D) can be geographically separated from production
facilities and can be supplied efficiently from the source that has the global low cost; 2) these
activities require relatively skilled labor; and 3) the knowledge-intensive services can
be simultaneously utilized at multiple production facilities. These three assumptions create
the motive for FDI and multinational production: the first two lend themselves to vertical
investment, the third to horizontal investment, as these terms are defined below.

Bergstrand and Egger (2007) add physical capital and third country effects to Markusen’s
knowledge-capital model to generate a knowledge-and-physical capital model of trade, FDI,
and foreign affiliate sales (FAS). With the introduction of an additional factor of production
to the model, physical capital, intra-industry trade and intra-industry FDI can coexist for two
identical countries and that it is possible to explain trade, foreign affiliate sales, and FDI
simultaneously. In addition, the introduction of third country effects into the theoretical
framework helps explain the complementarity of FDI, foreign affiliate sales, and trade to
changes in a country pair’s characteristics (size, other gravity variables). It also provides a
rationale for estimating gravity-like equations of FDI and FAS.

Finally, another important, and very distinct, contribution to this early branch of the
literature is Neary (2007) with its focus on how FDI can facilitate market power. Neary
focuses on firms’ decisions to undertake cross-border mergers, which account for the majority
of FDI. He uses a Cournot oligopoly model embedded in a general equilibrium setting.
Ricardian international differences in productivity lead to mergers, as low-cost firms purchase
their higher-cost foreign competitors.
3 Different Types of FDI

There is wide variation in the way that FDI is modeled in the literature, reflecting the many different forms that FDI takes in practice. Firms can decide to export a product it produces domestically, or they can instead decide to produce the same product within the foreign market it wants to serve. Producing in a foreign market at the same value chain stage as the home market is commonly referred to as horizontal FDI. Additionally, firms can decide to disperse different stages of the value chain to multiple countries, including the home market, for eventual export. This type of investment is referred to as vertical FDI. Many firms engage in a combination of the two strategies Helpman and Krugman (1985) focuses on vertical FDI. In their model, production is fragmented into stages and dispersed across borders. Capital-intensive headquarters is located in a capital-abundant country, labor-intensive assembly is located in a labor-abundant country, and there is intra-firm, international trade in services typically related to administrative functions such as accounting and human resources. These are often referred to as “headquarters services”. Helpman et al. (2004), on the other hand, focuses on horizontal FDI. Firms decide whether to serve a foreign market through FDI or through cross-border exports. The most profitable mode of supply depends on the productivity of each firm wherever it locates), as well as fixed costs of affiliate production and exporting and variable costs of exporting.

Yeaple (2003a) analyzes the possibility that firms may combine vertical and horizontal FDI via complex integration strategies that reflect the importance of both transport costs and factor price differentials. The type and complexity of a firm’s integration strategy depends largely on transport costs. Low levels of these costs lead to the dominance of vertical FDI, and high levels lead to the dominance of horizontal FDI. Two locations can be either substitutes or complements in the production activities of the multinational firm. Grossman, Helpman and Szeidl (2006) expands on the complex integration strategies in Yeaple (2003a). Their
model includes multiple stages of production (intermediate goods and then assembly) as well as headquarters services. The model examines the equilibrium choices of firms that vary in their productivity levels. These choices depend on the relative magnitudes of fixed and variable costs and factor cost savings. There is strategic complementary when locating one production activity in a low-wage country makes it more likely to shift other activities to that country. In equilibrium, there is a mix of strategies in place, but the analysis always takes the ownership structure of the firm as given.

Grossman and Rossi-Hansberg (2008) develops a model of vertical FDI that specifically focuses on offshoring. They represent production as a combination of many tasks and offshoring as trade in tasks after factor specialization between countries. The extent of offshoring depends on the efficiency of remote management and on wage differentials.

Ramondo and Rodríguez-Clare (2013) builds a calibrated general equilibrium model that extends the Eaton and Kortum (2002) framework to include multinational production and international technology transfer and combines horizontal and vertical FDI. Multinationals use imported inputs from their home country in foreign affiliate production. There are no fixed costs in the model, and perfect competition is assumed. Trade is a complement to multinational production and international technology transfer due to intra-firm trade in intermediate goods. At the same time, trade can also be a substitute for multinational production because the two modes are alternatives for supplying a foreign market. Ramondo and Rodríguez-Clare calibrate the model to data for 19 OECD countries for the period 1996-2000. The model quantifies the real wage gains from openness when countries interact through cross-border trade and multinational production. Including multinational production and international technology transfer in the model approximately doubles the real wage gains from trade.
4 Internalization and Incomplete Information

A branch of the FDI literature examines what defines the boundaries of firms that engage in international transactions: why do they combine and internalize geographically dispersed production in one integrated multinational firm rather than building a supply chain through licensing technologies to other parties and conducting business through arms'-length contracts? This is an extensive and technically complex literature, but there are some excellent surveys.

For example, Helpman (2006) demonstrates that the insights and modeling approaches available from incomplete contracts theory can be applied to decisions about location and the boundaries of the firm. Internalization within a multinational firm is generally more efficient than international outsourcing when there are hold-up problems associated with specialized production inputs and investments, imperfect monitoring, and a risk of technological diffusion.

Antràs and Yeaple (2014) offers another great overview of the FDI literature on organizational choice. They explain that internalization is often due to market imperfections, and in particular contract incompleteness around intellectual property. There is a possibility of too little investment due to hold-up inefficiencies. Contracting incompleteness means that equilibrium outcomes are determined by bargaining and participation constraints.

5 Econometric Analysis of FDI

Many of the surveys of the FDI literature – including Markusen (2002), Helpman (2006), and Antràs and Yeaple (2014) – start with an overview of stylized facts about FDI gleaned from aggregate and firm-level data. These stylized facts help to establish targets and motivation for theoretical analysis. The most commonly cited stylized facts are that most FDI is between
developed countries and in R&D and capital-intensive industries; firms that engage in FDI are usually the largest and most productive; and most FDI occurs through mergers and acquisitions rather than greenfield investment.

The econometric literature has greatly elaborated on these stylized facts, relying for the most part on detailed, firm-level data. As Antràs and Yeaple (2014) point out, "models of the internalization decision [of a multinational firm] are essentially models of firm behavior, [so] firm-level data are an ideal laboratory to use in testing certain aspects of the models." In this section, we highlight some of the most influential studies in the large econometric literature on FDI.

Carr et al. (2001) reports econometric tests of the knowledge-capital model of multinationals in Markusen (1997), using BEA foreign affiliate sales data from 1986–1999. In their model, foreign affiliate sales in the host country are dependent on the difference in GDP between the host and source country, the difference in skill level between the host and source country, the investment costs of the host country, the trade costs of the host country, the trade cost of the source country, distance, and the interaction of host trade cost and skill difference. Their results confirm the predictions of the model: increasing trade costs in the host country lead to an increase in affiliate production; an increase in trade costs in both the host and source countries lead to a decrease in affiliate production (trade and investment are complements); affiliate sales increase when income converges between the two countries; and when the U.S. is the source country, an increase in skilled labor in the host country increases affiliate sales.

Yeaple (2003b) is an econometric analysis of the factors that determine the location and magnitude of U.S. outbound FDI. He finds that patterns of FDI at the level of host country and industry are explained by transportation costs, plant scale economies, market size, and the relative factor intensity of the industry interacted with the relative factor abundance of the country. He finds significant evidence of both horizontal FDI driven by
market access issues like trade costs and vertical FDI driven by factor proportions-based comparative advantage. Helpman et al. (2004) test their model of horizontal FDI using industry-level BEA foreign affiliate sales data. In their regressions, the log of the ratio of exports to FDI sales (i.e., foreign affiliate sales) is the dependent variable. The independent variables include variable costs such as trade costs (freight and tariff rates), a measure of plant-level fixed costs, several alternative measures of the dispersion of productivity across firms within the industry, and the industry’s capital-to-labor ratio and R&D intensity. Their econometric estimates confirm the predictions of the theoretical model: variable trade costs have a negative impact on exports relative to FDI sales favoring proximity; scale economies have a positive impact leading to concentration of production and exporting; the dispersion of productivity levels has a negative effect; and greater capital-to-labor ratios and R&D intensities both favor FDI over exports. The econometric analysis also confirms the model’s prediction that sorting into FDI will reflect within-industry dispersion in firm productivity.

Baltagi, Egger and Pfaffermayr (2007) uses bilateral U.S. FDI stock of manufacturing and non-manufacturing sectors from the BEA, covering 1989–1999, to estimate the bilateral determinants of FDI as well as the spatially-weighted third-country determinants of FDI. They estimate a ‘complex FDI’ version of the knowledge-capital model previously developed by Markusen (1997) and tested in Carr et al. (2001) (and follows the literature on ‘complex’ MNEs, for example Yeaple (2003a)). They find that ‘bilateral and third-country effects of changes in skilled and unskilled labor endowments tend to be substitutes for vertical and complex vertical FDI.

Head and Reis (2008) develop a model of FDI as a market for corporate control. They use the model to derive a gravity model for FDI. In their model, parent companies have incomplete information on potential acquisitions in host countries and monitoring costs increase as function of distance. The model takes multilateral effects into consideration in that the likelihood of a firm successfully acquiring a subsidiary (through bidding) depends
on the parent’s location relative to firms in other countries, in addition to its own distance from the acquisition. The authors use bilateral FDI data for 30 OECD countries and 32 partner countries to estimate the parameters of the model and test its predictions. They find that bidders enter as predicted and that the level of development of the source country is a significant determinant of outward FDI. They use their model to calculate predicted shares of FDI from the model estimates and find that the model fits the data well.

Davies, Desbordes and Ray (2018) is an econometric analysis of transaction-level FDI data for many sectors and countries for the period 2003—2010. The model identifies factors that are important determinants of greenfield FDI and international mergers and acquisitions. Greenfield FDI usually involves a transfer of an asset developed in the origin country, so it reflects the financial development and comparative advantages of the origin country. On the other hand, FDI through mergers and acquisitions usually reflect the institutions, financial development, and cross-border barriers of the destination country.

Several studies use firm-level data to examine technology transfer within multinational firms. Branstetter et al. (2006) investigates how international technology transfer between the parent and affiliates of multinational firms responds to reforms in the protection of intellectual property rights in host countries. The econometric analysis utilizes affiliate-level data from individual U.S.-parented multinational firms in 1982-1999. The authors find that there was greater transfer of technology, measured by royalty payments back to the U.S. parents when patent regimes in the host countries were strengthened, and the impact was large, with royalties increasing by more than 30 percent for the most patent-intensive firms. Keller and Yeaple (2009) estimates international technology spillovers to U.S. manufacturing firms from U.S. inbound FDI between 1987 and 1996. The study finds that "FDI leads to substantial productivity gains for domestic firms," with a larger impact on smaller firms and in high-tech sectors.

The econometric literature also examines the impact of FDI on labor markets. Most
notably, Harrison and McMillan (2011) uses firm-level data on U.S.-parented multinationals from the BEA benchmark years between 1982 and 1999 to estimate the impact of competition from workers in foreign affiliates (specifically changes in affiliate wages) on employment in their U.S. parents. They find that there was a mix of complementary effects (with low-wage affiliates and intra-firm trade) and substitution effects that partly offset each other, and they conclude that, on net, offshoring is responsible for only a small part of the decline in U.S. manufacturing employment between 1982 and 1999. The effects depend on the type and location of the FDI. Their model distinguishes between affiliates in high-income countries and affiliates in low-income countries.

Finally, the empirical literature also considers the impact of institutions on FDI. Benassy-Quere, Coupet and Mayer (2007) estimates a gravity equation for bilateral FDI stocks that includes novel institutional indicators to examine the impact of institutional quality and the difference between national institutions on FDI. They find that institutions do matter for foreign investment, independent of GDP per capita. The most significant determinants of FDI were level of bureaucracy and corruption. Like physical distance in traditional gravity estimations of trade, institutional distance between two countries reduces bilateral FDI.

6 Policy Analysis Focused on FDI

Most of the economic literature of FDI focuses on model predictions about whether multinationals form and why international trade occurs within firms rather than at arms’-length. In contrast, there is much less analysis of how trade and investment policy affect FDI. However, many of the modeling frameworks within the literature can be adapted for policy analysis, either by extending the structural models to incorporate policy variables or by simply reinterpreting the parameters already in the models as policy variables. For example, investment policies and non-tariff measures could be represented as shocks to the fixed costs in the mod-
els, and tariffs could be represented as shocks to the variable trade costs in the models. In this section, we highlight policy studies that develop and apply FDI models.

One group of studies examines the impact of bilateral investment treaties (BITs) on FDI. Egger and Pfaffermayr (2018) estimates several variants of the knowledge-capital model. Their econometric analysis utilizes bilateral outward FDI stock data from OECD countries, covering the years 1982-1997. BITs are theorized to increase FDI because they increase transparency, thereby reducing the risk of investing in a country. Egger and Pfaffermayr find that ratified BITs have a positive and significant impact on outward FDI stocks. They also find that signed BITs also have a positive, albeit smaller, impact on FDI stocks, but the effect is insignificant in most specifications. Egger and Merlo (2007) uses a dynamic adjustment model to estimate the impact of BITs on outward FDI stock in OECD countries and the transitioning economies of Central and Eastern Europe. The dynamic model uses lagged FDI values to account for the sluggish adjustment of FDI over time. Bilateral investment treaties between a developed source country and a developing host country stimulate FDI by communicating the costs of expropriation. They find that BITs trigger a larger impact on FDI in the long-run (a 8.9 percent adjustment) than in the short-run (a 4.8 percent adjustment).

A second group of studies uses industry-specific partial equilibrium models of FDI to analyze the impact of changes in trade and investment policy. Riker (2015) is an econometric analysis of the impact of trade restrictions on U.S. foreign affiliate sales of services in 46 countries for the period 2009—2012. The model focuses on the effects of barriers to each mode of supply on substitution between cross-border exports (mode 1) and foreign affiliate sales (mode 3). The estimates indicate that eliminating existing restrictions on mode 3 supply would increase foreign affiliate sales by approximately 74 percent. On the other hand, eliminating restrictions on substituting mode 1 supply would reduce foreign affiliate sales by approximately 24 percent. Khachaturian and Riker (2017) analyzes the impact
of barriers to trade on foreign affiliate sales and cross-border exports of legal services and architectural and engineering services, using a partial equilibrium simulation model based on Helpman et al. (2004) and is calibrated to industry data. The model estimates that reducing the fixed costs of mode 1 and mode 3 trade by half would have large effects on cross-border imports into the U.S. market and on foreign affiliate sales in the U.S. market, but would have only small effects on sales of domestic producers and on overall prices of these services in the U.S. market.

Riker and Schreiber (2019) develops a pair of industry-specific simulation models that can be used to estimate the effects of tariff changes on incentives for tariff-jumping and export platform FDI. Montgomery and Riker (2020) develops a model of the impact of tariff changes when cross-border ownership and control of price decisions is shaped by FDI.
References


