ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

FDI, ownership shares and domestic productivity

A Meta-analysis

Sanne Blankestijn

July 2012

Bachelor thesis for BSc Economics and Business Economics

Supervisor: dr. M.J. Burger
Academic year: 2011-2012
Student number: 333058

KEYWORDS: FOREIGN DIRECT INVESTMENT, META-ANALYSIS,

PRODUCTIVITY SPILLOVERS, OWNERSHIP SHARES

ABSTRACT

In this thesis the results are presented of a meta-analysis of the literature on Foreign Direct Investment and productivity spillovers. Special attention is given to the distinction between minority and majority ownership. First, this thesis provides an overview of the relevant theory on the existence and development of these spillovers. It distinguishes between three effects: competition, demonstration, and the labor mobility effect. Then, a dataset is established with ten empirical studies that take ownership shares into account when researching this topic. On this dataset meta-regression techniques are applied to find which study characteristics influence study outcomes. It was found that several study characteristics have a significant effect on the coefficient size. First, using different definitions for the dependent variable in the original study yields different results. Also, the proxy used for foreign presence has a significant influence on the original coefficient. Next, using cross-section data instead of panel data gives smaller results. This is related to the finding that the number of years included in the original study gives different results. No significant influence is found of the statistical method used, if spillovers are explicitly considered in the original regression or if the country is developed/developing. No evidence was found that study characteristics significantly influence the sign and significance of the coefficient. Finally, some evidence was found that results do differ when minority or majority observations are considered.

Contents

1. INTRODUCTION

In 1997 the Economist published an article about multinationals in China. It described some examples of foreign companies who were trying to become successful in the Chinese market. One of the stories was about Alcatel, a French telecom company that entered the Chinese market by means of a joint venture. They cooperated with a Chinese company that had complete control over China's fixed-line network. The joint venture, called Shanghai Bell, became quite successful, and provided Alcatel with a competitive advantage over competitors such as Siemens and AT&T.

The above story describes an example of Foreign Direct Investment (FDI), which I will further define as the direct investment of a company (by acquisition, extension of current business, or establishment of a new company) in a foreign country. FDI has been one of the major forces behind globalization in the last years. Governments have worked hard to attract foreign firms with attractive subsidies and other (fiscal) policies. The main justification for these policies is the common belief that FDI creates (technological) spillovers, which will increase domestic productivity and stimulate economic growth. Productivity spillovers will occur if subsidiaries of multinational firms "*lead to productivity or efficiency benefits in the host country's local firms*" according to Blomström and Kokko (1998 - p.249).

In 2011, global inflows of foreign direct investment rose by 17%, according to the Economist (2012). This added up to a total amount of an astonishing 1.5 trillion dollars. Interestingly only a little over half a trillion dollars was spend in developed countries. This means that a majority of the investments is done in developing or transitioning countries. For example, the investment in Latin-America was up 35% (The Economist, 2012). As a result a large amount of research has been directed to this field, trying to answer questions such as: What are the consequences of foreign direct investment? What do companies and countries expect to gain from it?

The main question is to prove that a positive influence of FDI can be substantiated by empirical evidence. Second, do foreign entrants cause these spillover effects and if so, to what extent? Are these results similar for all countries? Are these spillovers present in every industry? Finally, does the set-up of the research by the scholar have a (significant) influence?

The empirical evidence on this subject is mixed. Görg and Greenaway (2004) reviewed the literature on export spillovers, wages, and productivity effects of FDI and found a gap between theory and empirical evidence. On a theoretical basis, scholars agree on the likelihood of spillovers but empirical evidence has been inconclusive. Studies have shown different results for different countries but also different results for different industries. Globerman and Chen (2010) also evaluated the literature on spillovers and distinguish between the main variables used to determine the extent of spillovers in the literature. They find six main categories which have an influence on the existence of FDI, namely: Absorptive capacity, openness of the economy, nature of FDI and linkages, regulations, infrastructure and finally industrial structure. They believe that spillovers occur when these variables point to a favorable investment climate. Taking absorptive quality as an example, the authors believe that with a strong absorptive quality, the likelihood of spillovers increases. Absorptive quality expresses how well a domestic company can internalize knowledge. Later, in the theoretical framework this paper distinguishes between three main channels thru which spillovers may occur: Competition (only efficient domestic companies survive), labor mobility (workers take the knowledge form foreign firms with them and apply it in a domestic firm), and demonstration effects (domestic firms learn from the techniques used in foreign firms). For an in-depth explanation of how spillovers occur and the categories of Globerman and Chen, see section 2.

There is also an ongoing debate on the direction of the correlation between FDI inflows and domestic productivity. Options are that the domestic productivity is increased by spillovers or maybe the competition effect is more important. This effect states that by the entrance of foreign firms, the less productive domestic firms are forced out of the market and hence the average productivity is increased. The third option is that multinationals only enter in high-productivity industries, thus not increasing the average productivity level but maintaining it. This option makes it interesting to distinguish not only between countries but also industries.

Notably absent from the discussion and most scientific research is the influence of ownership shares. This variable has not been part of the mainstream FDI literature, but it has been researched by a few scholars. Some studies focus on the distinction between majority/minority and other studies focus on wholly owned subsidiaries versus partially owned subsidiaries. This absence is remarkable given the general believe that local participation increases the chance of spillovers. If firms work together it is easier for the domestic firm to access the knowledge of the foreign firm, learning from it and using the

gained knowledge in their own processes. Governments are so convinced that this is true, that they base some of their FDI policies on it (Blomström & Sjöholm, 1999).

A second reason why the size of ownership shares is fascinating is that theoretical reasoning leads to different results for different degrees of ownership. For instance, one can reason that if a foreign firm has a larger ownership share, it has the right to a larger part of the profits. It will then have a stronger incentive to transfer the newest technology to its subsidiary.

Also, the few studies that do take ownership in to account, find different results. Some studies find that the degree of ownership does not influence the size of spillovers, others do find significant effect. This raises the question: What causes these different results?

This question has not been answered in the literature yet. The literature has only started focusing on this subject over the last ten years, with very few scholars publishing in this subfield. Given this period and the increasing attention for this field, the time is here to evaluate the different results thus far. An answer to this question might be found in the methodology, data type or research design used in the different studies. In order to quantify the influence of these factors, this thesis will provide a meta-analysis of the different studies which do take degree of ownership into account. A meta-analysis is a quantitative research method used to summarize results from different studies and explain why they vary. The studies have to be on a similar topic (Card & Krueger, 1995). Given that in this time and age there is a pressing need to critically review published articles and do this in an objective way, meta-analysis became a necessity. It is a good measure to compare economic empirical research, as the research which is done in the field of FDI (Stanley, 2001)

As a result, the focus of this thesis will be on the meta-analysis of ten studies which take ownership shares into account. I will perform two different meta-regression analyses. First, the coefficient size in the original study is used as a dependent variable – which will be regressed on several study characteristics. This will result in several study characteristics which play an important role when performing a study on FDI, ownership shares and domestic productivity. Next, these properties will be tested on a logistic regression to see if the significance or positivity of a coefficient can be explained by these study characteristics as well.

I find several study characteristics to have a significant result on the coefficient size. Using different definitions for the dependent variable in the original study yields different results.

Also, the proxy used for foreign presence has a significant influence on the original coefficient. Next, using cross-section data instead of panel data gives smaller results. This is related to the finding that the number of years included in the original study gives different results. No significant influence is found of the statistical method used, if spillovers are explicitly considered in the original regression or if the country is developed/developing on effect size. No evidence was found that study characteristics significantly influence the sign and significance of the coefficient.
Finally, some evidence was found that results do differ when minority or majority observations are considered.

This thesis has the following structure. First it will give an overview of the existing literature and different theories of FDI. There will be some attention on the reason why ownership shares are so important. Also, some information is given on the technique of meta-analysis. Section 3 provides an overview of the methodology and data used whereas section 4 will describe the meta-analysis and the resulting analysis. Section 5 will conclude.

2. THEORETICAL FRAMEWORK

This section will provide an overview of the various theories published about this subject over the last 20 years. The first section provides some basic arguments on why a firm would decide to become a multinational and thus engage in FDI by definition. The next section contains a brief review of FDI spillover theory, followed by a brief overview of additional reasons why degree of ownership can have an influence on productivity spillovers[1]. The section will conclude with some short remarks on the theory for meta-regression.

2.1 REASONS WHY A FIRM BECOMES A MULTINATIONAL

Before researching the effects on FDI, it is important to establish the reasons why firms would choose to establish a subsidiary in another country or in other words: become a multinational. It seems logical that the reasons why a firm would internationalize are probably also the basis for their decision on mode of entry (e.g. ownership shares). Dunning and Lundan (2008; p. 3) describe a multinational as follows: *"an enterprise that engages in foreign direct investment (FDI) and owns or, in some way, controls value-added activities in more than one country".* This is a literature wide excepted definition.

The first question is why firms decide to start operating in a foreign market. They probably have less information than local firms on the market situation and they do not have the supplier network in the foreign country. From, a business perspective, Dunning (1981) answers this question. He states that in order for a firm to become a multinational, several conditions have to be satisfied. These conditions are theorized in Dunnings (1981) well-known OLI-framework.[2] Dunning distinguishes between three types of advantages that a firm should have: ownership, location, and internalization. Ownership advantages can result from the possession of superior knowledge assets, but also from entrepreneurial skills of the employees. The empirical literature confirms that multinationals have knowledge assets, as they spend significant resources on R&D (Griffith, 1999).

Another condition which has to be satisfied is the location advantage: the country of destination should offer some advantages in the area of raw materials or infrastructure. Internalization advantages exist if it is more attractive for the firm to produce its product itself in a foreign country instead of using partnership arrangements (e.g. joint ventures, licensing) or trade. The framework adds the three types of advantages to get the *net ownership*

[1] For a more complete overview, see Smeets (2008): Collecting the Pieces of the FDI Knowledge Spillovers Puzzle
[2] Ownership – Location – Internalization (OLI)

advantages. This term expresses the total of advantages that the multinational has in foreign countries (Hagen, 1997).

Markusen (1995) also provides an answer to this question, using a different approach. Markusens reasoning is based on the theory of trade, from a macro-economic perspective. His model takes firm-specific advantages into account. He refers to the knowledge-based assets of firms, which are similar to Dunnings ownership advantages. He states that these assets give rise to economies of scale which in turn make expanding in other countries worthwhile. He also states in his theory of the multinational firm three main reasons for expanding in another country: Legal, cost advantages or market opportunities. The legal system is of importance for the possibilities of the multinational firm to protect its assets. It would, for instance, be important to look at the intellectual property laws. Market opportunities consist when a firm is certain it can establish a certain market share for itself. Cost advantages are for instance present if the average wages are lower in the foreign country.

So the main difference between the approach of Markusen and Dunning is that Dunning takes the perspective of the firm: what is better for the firm and which advantages can the firm get from internationalizing? Markusen on the other hand, considers this question on country level and investigates what countries have to offer. He does try to incorporate some firm-specific elements in his theory, but this is not his main focus. Both theories are relevant given that they establish different reasons and viewpoints for the firm to become a multinational.

In conclusion: the common theory is based on the *firm-specific advantage hypothesis*. This hypothesis states that a company has advantages which are the basis for the international expansion and as a result become a multinational. These advantages are presumably transferable within the firm and not available to competitors (Bellak, 2004). In this theory the knowledge-based assets, ownership and internalization advantages can easily be found. Location advantages are not firm-specific but to what extent the multinational can internalize these advantages is based on firm-specific characteristics.

2.2 FDI SPILLOVERS

2.2.1 GENERAL COMMENTS

The literature on FDI spillovers is very broad. A common term used in the literature is the ‘black box’. Most studies assume that FDI spillovers are occurring instantly. They just test for the existence of spillovers but do not investigate the causes. The term “spillovers” it self

usually refers to a couple of effects. The spillovers where this thesis especially focuses on are called *productivity spillovers*. This is because the studies which are part of the meta-analysis have as a goal to establish if productivity spillovers occur in their dataset. These spillovers can be defined as the "*changes in productivity of the domestic firm due to the presence of foreign firms*" (Iyer, 2008; p. 5).

In order to research this topic, I will distinguish between the three main causes of spillovers – which are the main causes described in the literature: the competition effect, the labor mobility, and the demonstration effects (Suyanto & Bloch, 2009). In this chapter, I will refer to both vertical and horizontal spillovers. Vertical spillovers are inter-industry spillovers related to forward and backward linkages of firms. Backward linkages exist between the multinational and its local suppliers, forward linkages between the multinational and its buyers. Horizontal linkages are within the industry (intra-industry) on the same level in the supply-chain. This is the effect the multinational has on its competitors.

Two models focusing on vertical linkages are Rodriquez-Clare (1996) and Markusen & Venables (1999). Rodriquez-Clare assumes in his theoretical model a 'love-of-variety'. His assumption is that foreign firms make more complex products, needing more complex and more diverse inputs. This leads to a backward linkage with intermediate suppliers through the increased and/or specialized demand. The model of Markusen and Venables will be discussed in the next section

2.2.2 COMPETITION EFFECT

The competition effect can be described and interpreted in several ways. The result is most of the time the same: the multinational enters the market which results in the exit of domestic firms.

Markusen and Venables (1999) describe the competition effect as followed: through the entrance of the new firm the total output rises. This will result in a drop in the prices which causes the exit of domestic firms and an increase in overall productivity. Markusen and Venables combine this effect with the so-called *positive development effect*. The entrance of the new firm also increases the demand for inputs, forcing the intermediate good suppliers to increase production. This will result in a decrease of average costs and hence increased productivity. This positive development effect is also shown by Rodriquez-Clare (1996), see above.

Another example of the competition effect is the interpretation as a *market stealing effect* in an imperfect competitive market. Now market output is not increased but the market capacity is redistributed with domestic firms losing out. Their production will decrease resulting in rising average costs (Aitken & Harrison, 1999).

The foreign competition effect is not always negative; it might have a positive impact on domestic firms. Kugler (2006) states: competition will provide managers with an incentive to work on their efficiency and make appropriate investments. Next, the exit of non-efficient firms will increase average productivity.

Empirically evidence is found for the existence of the competition effect. Aitken and Harrison (1999) find evidence for their market stealing effect in Venezuela. Also they state that the competition effect is larger in developing countries, which is confirmed by Konings (2000) for Bulgaria, Romania and Poland. In both Bulgaria and Romania the competition effect more than offsets the spillover effect resulting in a negative effect of FDI.

Not only can a distinction be based on development level but also between inter- and intra-industry. Evidence was found by Kugler (2006) that the competition effect is larger for inter-industry than for intra-industry. This is intuitive because of the fact that most direct competitors are in the same industry and will feel the influence of the multinational very quickly. He also finds that productivity advances are in complementary or non-competing sectors. He states that the amount of competition within an industry influences the extent of spillover effects. Javorick and Spatereanu (2008) find that the competition effect in Romania is smaller for partnership companies. This is probably because when working together multinationals and domestic firms have a smaller incentive to compete with each other.

It might also be possible that the multinational itself is influenced by the competition. The multinational might be forced to transfer more knowledge assets to their subsidiary in order to compete with domestic firms (Blomström, Globerman, & Kokko, 2001). This can result in a larger incentive to innovate and thereby affecting the host economy.

An important note for this section is that this is all based on theory. To the best of my knowledge, no empirical research has explicitly studied the effect of competition in a foreign market on the domestic economy.

2.2.3 LABOR MOBILITY

Labor mobility points to the movement of workers between different firms (worker turnover). Workers can learn from their foreign counterparts and spread the knowledge. It is likely that the level of training in foreign firms is higher than in local firms (Smeets, 2008). When workers leave the firm they will take the knowledge and skills with them, spreading it across the industry. Workers trained by foreign firms will use this knowledge when changing jobs.

When an owner of a firm has previous experience in a multinational firm, who operates in the same sector, the productivity of his firm will rise. In contrast, if the owner was trained in a local firm, the productivity would not rise (Görg & Strobl, 2005). These were the results from a panel study of 228 Ghanaian manufacturing firms, but they give rise to the thought that these effects can be more global, supporting the labor mobility hypothesis.

While this thesis focuses on productivity effects, I would like to shift the attention to research on wage levels. This because wage levels and productivity are closely connected: mostly a rise in wages can only be justified by a rise in productivity. This follows from reasoning among the following lines. First, if the firm does not raise productivity it would be very hard to pay for the higher wages of the workers. Second, a pay raise is usually used as an incentive to work harder, increasing the productivity of the firm. Lastly, wages can be used to attract different personnel. Firms can try to attract better trained personnel by offering them higher wages. Empirically, Markusen and Trofimenko (2007) showed an instantaneous and lasting effect on real wages in a firm. These wages went up when the firm hired a foreign expert. Also, for Brazil it was shown that if the amount of foreign workers in the sector increased, consequently the wages went up as well (Poole, 2007). In China, the rise of foreign workers caused an increase in firm productivity, according to the research of Hale and Long (2006).

2.2.4 DEMONSTRATION EFFECTS

Demonstration effects are the effects on the domestic firm by observing the foreign firm. These effects point to the lessons learned by the domestic firm. Demonstration effects can occur in three different ways: imitation, new innovation, and reverse engineering (Suyanto & Bloch, 2009). We speak about imitation if a local partner decides to copy the technology of the foreign firm and use it in their own products. Innovation is the process in which the local firm uses the technology of the foreign firm as a basis for their own new products. Finally, reverse engineering means that the local firm studies the multinationals' products. It does this

in order to find the technology behind the product and the demands which the product has to fulfill.

Alfaro and Rodriquez-Clare (2004) felt that the intended demonstration effects will be stronger in the vertical zone than horizontally. This is quite intuitive: it is in the interest of the foreign firm to have suppliers who can deliver high quality inputs. For this purpose, the sharing of some of the technology through backward linkages may be required or beneficial. Horizontally the foreign firm will have an incentive to protect its knowledge to prevent competitors from taking advantage of it.

Das (1987) developed a theoretical model to see who benefits from technology import by the subsidiary. Her conclusion was that, even though the multinational will face some loss by the learning of domestic firms, they still come out ahead.

Another critical factor for the existence of the demonstration effect results from the firm-specific characteristics of domestic companies. It is important to involve the absorptive quality of a firm in the analysis because this determines the capability of the domestic firm to use the spillovers and internalize them (Wang & Blomström, 1992). Also, the knowledge gap might be a factor in the ownership share decision (see section 2.3). Three points are determining the possibility of the demonstration effects: human capital, the R&D expenditures, and the technology gap. For human capital, see section 2.2.3.

R&D expenditures are used in the literature as a proxy for the absorptive quality of domestic firms. Absorptive quality is *"the ability of those firms to adopt and exploit new technologies and management processes that are brought into the host country through FDI."* (Globerman & Chen, 2010; p. 5). It is hypothesized that the larger the R&D expenditures, the better the firm can adopt its spillovers. Empirically, it was confirmed that for twelve OECD countries R&D played a substantial role in the knowledge transfers (Griffith, Redding, & van Reenen, 2004).

The technology gap refers to the difference in state of technology at the multinational and at the domestic firm. Perez (1997) found that the ability of a firm to actually absorb spillovers does depend on two things: the existing level of technology within the firm and its ability to learn. The latter is consistent with the finding described above that the ability to survive the knowledge gap is related to the amount of R&D expenditures. The size of this gap determines to a large extent the possibility of knowledge spills. If the gap is too small there is not any

space or incentive for the domestic firm to learn, but if the gap is too wide, it will not be possible to cross it (Barbosa & Eiriz, 2009). This conclusion is shared by Kokko *et al.* (1995) based on their research in Uruguay.

2.3 IMPORTANCE OF FOREIGN OWNERSHIP SHARE

This thesis adds to the existing literature with its attention for the influence of the foreign ownership share. By doing this, it takes heterogeneity of firms into account while studying the effect of FDI on firm productivity. It will later distinguish between minority and majority foreign ownership, as is done in the existing literature. Governments believe that partnerships encourage technology spillovers, but is this really the case? Blomström and Sjöholm (1999) were the first scholars to take this effect into account. They argue that spillovers tend to occur when two companies work together but they also say that a partnership can limit the willingness of the multinational to share their technology.

First, Blomström and Sjöholm (1999) argue that the risk of losing assets is important. If multinationals feel that they could lose their assets when working together with domestic firms, they might only share older technology. Based on theoretical reasoning, Coughlin (1983) proved that in countries that discourage fully owned foreign subsidiaries indeed have older technology transferred to them. This does not have to be a negative phenomenon, its actual effect depends on the width of the above described knowledge gap. Javoricik and Spatareanu (2008) agree with this vision. They feel that combined with the easier access which comes from *minority* ownership to technology, large horizontal spillovers can occur. They identify an advantage for local suppliers: less complex inputs can be more likely made by local suppliers. When combined with the knowledge of the local partner it is likely to stimulate the development of inter-industry spillovers.

Second, a *majority* ownership share means that the foreign firm is in control and has a direct interest in the profits. It will be their goal to maximize these profits and make the subsidiary as efficient as possible. In order to achieve this goal, they might be willing to share more technology. With their finest technology, it will be easier for the subsidiary to achieve their fullest potential. With a higher level of ownership (thus sharing), the chance of spillovers will increase. Furthermore, a *minority* ownership would force domestic firms to do more of the work. They will get better acquainted with the technology in order to do this and will be able to learn more. This will result in local people having more knowledge, which can help raise

domestic productivity and GDP (Blomström & Sjöholm, Technology Transfer and Spillovers: does Local Participation with Multinatinals matter?, 1999).

Next, Dimelis and Louri (2002) add to the discussion by arguing that the decision around ownership shares is based on net returns. They state that an increase in foreign share brings some extra profits for multinationals, but this effect is decreased by additional monitoring costs. The multinational wants to protect its property rights add the costs to do this should be a factor in the decision making process.

Combining the arguments above, we find a tradeoff, modeled by Müller and Schnitzer in 2006. They explicitly focused their research on joint ventures, but gave some interesting insights in the relationship between the size of ownership share and the size of the knowledge transfer. They argue that a larger ownership share comes with a larger knowledge transfer from the parent company but lowers the access of the local firm to the technology. They believe that minority companies are more likely to create spillovers. Firms with majority foreign ownership actually have productivity increases as a result of the spillovers.

The question to which side the balance will point, is only to be answered empirically. Unfortunately, the empirical studies find contrasting results, also due to different approaches. Blomström and Sjöholm (1999) find that the ownership share (minority/majority) does not influence domestic productivity in Indonesia. Dimelis and Louri (2002) find the opposite for Greece, only majority or full owned subsidiaries have an influence on domestic productivity. Javorick (2004a) found only positive spillovers for partially foreign owned subsidiaries. Abraham et al. (2007) find that the spillover effect for majority ownership shares is smaller than in case of a minority share for Chinese plants. Finally, Javorick and Spatereanu (2008) find results which indicate a different influence of ownership shares on productivity spillovers. Only joint projects deliver vertical spillovers through backward linkages. Ramachandran (1993) finds that both in his theoretical model and empirical analysis, multinationals subsidize knowledge transfer in wholly owned subsidiaries more heavily.

From all the theory and empirical evidence stated above, we can conclude that a definite answer is not possible. In general the distinction is made between fully owned and partially owned. The different effects can offset each other making the outcome (and size of the influence) on domestic productivity uncertain. Both a majority and a minority ownership share have different influences on the host economies and channels of spillovers, making it an interesting field of research. The different results of this research give rise to the idea that the

research approach plays a role as well. This is a very important observation, given that, as Smeets (2008; p. 124), says: "The actual relation between multinational enterprise ownership and knowledge spillovers may turn out to be an empirical matter."

2.4 META-REGRESSION

The standard approach used in science when comparing studies on one topic, is to write a literature review. An overview is provided combined with some ordering and the scholar tries to provide a systematic summary. This is a subjective process where the author has the freedom to make his own choices.

In order to quantify this procedure, Stanley and Jarrell (1989) introduce the technique of meta-regression. This is a branch of meta-analysis, according to them, especially designed to explore empirical research done in the field of economics. A definition of meta-regression is the following: *"In a meta-regression analysis, the dependent variable is a summary statistic, perhaps a regression parameter, drawn from each study, while the independent variables may include characteristics of the method, design and data used in these studies. Thus, meta-regression analysis can identify the extent to which the particular choice of methods, design and data affect reported results."* (Stanley, 2001; p. 131-132).

Meta-regression analysis is used to objectively examine empirical evidence. From the conclusions of the meta-regression, recommendations can be built for further research. An important variable in a meta-analysis should be included in further research.

Meta-regression uses mostly dummy variables to express the different study characteristics. For example, a dummy introduced to describe the type of data (cross-section/panel). A regression is than conducted in order to find significant results. If there are significant results, it can be concluded that that study characteristic plays an important role in the way a study should be set-up.

From this, a standard approach for meta-regression can be derived (Stanley, 2001). These steps are followed in this thesis. Step 1 is to find all relevant studies for the thesis. Next, a choice should be made for the summary statistic (dependent variable in the meta-regression). Here it's important to look at the different methods used by the original studies and find a common denominator. For example, the regression coefficients can fulfill this role. Third, a decision should be made on which independent variables to include. Following, it's time to conduct the meta-regression. In this step the meta-regression is estimated, and the results are

discussed and analyzed. Last, the meta-regression should be tested in order to establish if it fulfills the requirements from the statistical method used. [3]

Of course, this method has some shortcomings. Stanley (2001) mentions five, four of which are relevant here[4]. One of the weaknesses is the selection of study characteristics that are included in the regression. This decision is still up to the author, which adds a subjective aspect. Second, the weights of the articles included. For instance, some articles give multiple observations while others only add one. Here, a risk for overweighing some studies exists. Furthermore, if all studies included in the meta-analysis have the same problem, the meta-analysis will not notice it. Last, the selection of studies itself is a sore point. Again, the author has some freedom here. Also, the studies are mostly just selected based on their method of research and relationship to the topic of interest. Quality of the original study is not considered.

[3] For instance, tests for autocorrelation, normality and heteroskedasticity
[4] The fifth limitation concerns research for publication bias.

3. DATA AND METHODOLOGY

In order to evaluate the different studies included in the meta-analysis, multiple steps were taken, as suggested by Stanley's (2001) approach. Here the first three steps are explained, as described in the theoretical framework, section 2.4. First, the selection of studies for the meta-analysis is described. Next, an overview is given of the main results of these studies. Furthermore, the methodology for the meta-analysis is described and clarified.

3.1 DESCRIPTION OF THE SAMPLE

The sample of papers consists of ten studies, considering eight countries. Of these studies eight are published in academic journals and two papers are from a working paper series. An overview of the papers is provided in table 1. The selection of papers was primarily based on the list of articles selected by Smeets (2008). He made a compilation of five studies which take ownership shares into account. Second, I used online search engines (such as Google Scholar and Scopus) to search for articles with references to these five studies and with the key words 'ownership' or 'foreign direct investment'. Using only English key words, the papers included in this meta-analysis are solely written in English. I also used the references within the selected papers to search for other articles with the same topic. There may be more studies concerning this topic, which were not found using these research methods. However, I am confident that the studies which are included are the leading literature on this topic.

INSERT TABLE 1 HERE

3.1.1 RESEARCH DESIGN

Three studies consider cross-section data in their studies, others use panel data. Most studies cover multiple years (on average five years) in their study, while the cross-section studies only consider one year. These studies use different definitions for the explanatory variable; some studies even use multiple definitions. For these studies all outcomes are treated as one observation. This will increase the number of observations in the study, which will increase the reliability of the analysis. The total number of observations is 65 (before corrections). Also the studies differ in the type and number of explanatory variables used. Most studies do take firm size and the capital-labor ratio into account.

All studies use plant (firm) level data. The number of observations differs greatly among the studies, with only 658 observations for Spain and 17.645 observations for China. As dependent variables different definitions are used. Blomström and Sjöholm (1999) use the

labor productivity, the same as Schoors and Van der Tol (2002) and Khalifah and Adam (2009). The logarithm of the labor-output ratio is used by Dimelis and Louri in both their 2002 and 2004 studies. Barrios, Dimelis, Louri and Strobl (2004) use labor productivity growth. Javorcik (2004b) uses growth of output. Takii (2005) uses the logarithm of the value added. Abraham, Konings and Slootmaekers (2007) use the logarithm of total factor productivity, which is the same measure Javorcik and Spatareanu (2008) use. Thus, in total six different measures are used to measure the influence of foreign presence on domestic productivity.

Also, different measures are used to express the share of foreign presence. Mainly there are three ways: share of employment, share of equity or share of sales within the industry. Blomström and Sjöholm (1999) use total gross output. Schoors and van der Tol (2002) and Barrios et al. (2002) use sales. Dimelis and Louri (2002) use all three measures for their OLS regression, for the quantile regression they use equity share as a proxy. Javorcik (2004b) uses equity, weighted by output. Dimelis and Louri (2004) and Takii (2005) use employment. Abraham et al. (2007), Javorcik and Spatareanu (2008) and Khalifah and Adam (2009) use equity share.

3.1.2 EMPIRICAL MODEL

Smeets (2008) states in his review of the literature a standard model used in the empirical research. This is the following standard model:

$$y_{ijt} = \beta_0 + \beta_1 Min\_FDI_{jt} + \beta_2 Maj\_FDI_{jt} + \beta_3 X_{it} + \beta_4 Z_{jt} + \varepsilon_{ijt} \quad (1)$$

Min_FDI and *Maj_FDI* measure the share of foreign presence (minority or majority) in industry *j* at time *t.* *X* is a vector of firm-level variables which are included in the analysis. *Z* is a vector of industry-level variables which are taken into account as well. Certain studies also consider the effect of linkages (see chapter 2), which will expand this equation. For instance, Javorcik (2004b) explicitly uses a variable to describe backward linkages.

The factors which are mostly considered for *X* are firm size, capital-labor ratio, and the skill level of employees. For instance, Blomström and Sjöholm (1999) look at the factors above and also take capital utilization into account, distinguishing them from other studies.

The factors included in the *Z* vector are mostly dummies. These include dummies for sector specific differences or dummies to distinguish between multiple industries. Most studies use dummies based on the ISIC division of industries.

3.2 METHODOLOGY

For this meta-analysis it is of interest to use the above described model as a starting point. In order to explain the results, two types of analysis are conducted. The first, a linear regression, is with the coefficient for foreign ownership share from the original study as the dependent variable (named *coefficient_size*). The second is a logistic regression, conducted to predict the influence of the study characteristics on a positive and significant outcome. This means that multiple observations are based on one regression. For instance, if a regression holds both the *Min_FDI* and *Maj_FDI* variable, this is treated as two observations. The coefficient is used because in the different studies different information is given about the regression. In order to be able to compare the different studies, a unit of measurement has to be used which is available in all studies; a role the coefficient has to fulfill. Unfortunately, the coefficient does not give information on the significance of the outcomes. In order to take this into account, the second regression uses a dummy variable as dependent. This is *coefficient_sigpos* which takes the value 1 if the coefficient is both significant (on 5% level or better) and positive. It takes value 0 if the coefficient is either negative or insignificant. In the data, 33 studies have a significant and positive coefficient for foreign presence. If this is the case, it is interpreted as evidence for spillovers (Görg & Strobl, 2001).

For the linear regression itself, the equation used by Görg and Strobl (2001) is followed. They establish the following equation:

$$Y_j = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{jk} + \varepsilon_j \qquad j = 1,2,\ldots.N \qquad (2)$$

Applied to this specific meta-analysis, here Y_j is the coefficient for foreign ownership share reported in observation j. Z_{jk} is a vector of independent variables used to explain the differences between the studies in this thesis. ε_j is the error term.

For the logistic regression, Y_j is the dummy for significance and positivity reported in observation j. Z_{jk} is a vector of independent variables used to explain the differences between the studies in this thesis. The independent variables are the same for both the linear and logistic regression. ε_j is the error term.

3.2.1 SELECTION OF VARIABLES

The selection of variables took place thru two steps. First, a list of variables used in other meta-analyses about FDI was compared (Görg and Strobl, 2001; Havranek & Irsova; 2011).

From this list, the variables used in both studies were selected to use in this meta-analysis. Second, these variables were researched for the studies included in this meta-analysis. For example, for each study it was determined what definition was used for the dependent variable. For each different definition a dummy variable was created. A list of variables is provided in table 2. Lastly, every observation was coded for the regression and study of source.

INSERT TABLE 2 HERE

Combined, in the analysis twenty-five (dummy) variables were included. Given the limited number of observations available in this study, the degrees of freedom are limited as well. This limits the possible amount of dummies to include in the model (Görg & Strobl, 2001). In order to correct for this, not all dummies were included in every model (see results).

The variable *cross_section* was included because Görg and Strobl (2001) found evidence that the type of data has an influence on the magnitude of spillovers. *Developing* was included because of the theoretical believe that investors in developing countries use less inputs from local firms. This stems from the notion that local firms are not able to provide these high quality inputs (Rodriquez-Clare, 1996). Next, *time_span* is included because of the large variety in number of observations per study. Mostly, this is because studies with larger amounts of observations consider more years in their study (Havranek & Irsova, 2011). Unfortunately, number of observations is not available for every study (see table 1), but *time_span* is frequently used in the literature in order to control for that. Also, the average number of Furthermore, the *dep_#* and *for_sh#* variables are taken into account as well, following Görg and Strobl (2001). Lastly, the method of estimation and spillovers variables are again based on Havranek and Irsova (2011).

3.2.2 DATA ANALYSIS

For this meta-analysis, the technique of meta-regression is applied. First, the data is analyzed to see if there are certain properties of the data which have to be taken into account. Descriptive statistics for the variables can be found in table 3.

INSERT TABLE 3 HERE

First, the data is examined for outliers. A 95% confidence interval was established for the *coefficient* which plays the role of dependent variable. This resulted in the exclusion of 4 observations from different studies which had extremely large coefficient sizes. This leaves

61 observations in the regression. The first observation left out was from a study by Dimelis and Louri (2004). They give the explanation that in their data, large positive externalities stem from small firms. Large firms only show negative effects of FDI. Combined, there is a large significant effect, also because their data has almost three times as many observations for small firms. The other three observations left out come from Takki (2004). He does not give an indication for a reason for the large coefficient sizes. For him, the results are not too large in the light of his entire study. Given that most of his coefficients are quite high, these findings fit in his research and will not raise any flags for him.

Next, the data is tested for correlations. From the correlation table it can be seen shown that most variables are correlated to each other. An overview of the correlations is given in table 4.

INSERT TABLE 4 HERE

The high correlations are caused by the limited amount of observations. For instance, only two studies use cross-section data. In the majority of the observations from these studies, *dep_log_laboutput* is used. Automatically, a high correlation will appear. Given that certain variables have a very high correlation with each other, a problem with multicollinearity will probably appear. When these problems are encountered, the variable with the highest average correlation is excluded from the analysis. When this is the case, it will be pointed out in the results chapter of this thesis.

Next, different regressions are estimated. For the regression on coefficient sizes, the method of Ordinary Least Squares regression is applied. For the results, I refer to chapter 4 (especially table 5). Stanley and Jarrell (1989) point out that the fact that *coefficient_size* was taken from so many different studies, can cause some problems with the regression. They fear that the different characteristics of the original studies cause the error terms in this meta-regression to be heteroskedastic. In order to prevent this from interfering with the results, I use the White estimator in order to construct White heteroskedasticity-consistent standard errors.

For the logistic regression, other methodological issues were encountered. Given that the amount of variance in the dummies is limited and the amount of correlation between the variables is high, problems were encountered with both multicollinearity and the regular estimation of logistic models. In the results section, the different problems are mentioned and some solutions are offered.

4. RESULTS AND ANALYSIS

In this chapter the results of the meta-regression will be presented. Next, the results of the regressions will be analyzed and linked with the theoretical framework. Special attention will be drawn to the distinction between minority and majority ownership.

Again, with reference to Stanley (2001), this section explains lasts step of his approach.

4.1 RESULTS

The results from the linear regression will be discussed in section 4.1.1. In section 4.1.2 the results of the logistic regression will be explained.

4.1.1 OLS RESULTS

The first step here was to establish four separate models with as dependent variable the coefficient size from the original studies. Each model focused on one or multiple types of dummies. The first model includes all the different dummies for definitions used in the original studies for the dependent variable (*dep_#*). In order to avoid the dummy variable trap, *dep_labpro* is left out of the model. *Dep_labpro* is the reference category. This technique is repeated with one model for all *for_sh#* variables (model 2 – *for_equity* is used as reference) and one for all *stat_#* variables (model 3 – *stat_OLS* is used as reference). Next, a model is established similar to the second model used by Görg and Strobl (2001). They establish a model which only includes the variables which concern the data. In this case, the model includes *cross_section, time_span, developing,* and *spillovers*. From these models, all significant variables are gathered. These variables are included in models 5-7. Model 5 is estimated for all data, 6 for all data referring to majority ownership and 7 for all data considering minority data. Observations which refer to general foreign direct investment are considered in all models. This method was used in order to empirically assess which variables play a role in explaining the differences in the included studies. Table 5 gives an overview of the seven models estimated.

INSERT TABLE 5 HERE

Table 5 shows seven different regressions. Each regression includes some dummies from table 2. The table shows the coefficients for these variables, as well as the White heteroskedasticity-consistent standard errors. The table shows several significant results for *dep_#* and *for_sh#*. Interestingly, *stat_#* has no significant variable. When only looking at the

variables which describe the data used in the original studies, both *cross_section* and *time_span* show a significant result. Also, it's interesting to see that there is a noteworthy difference between the models in R^2. The range is from 4.6% for model 3 (statistical methods) to 63.5% for model 5 (significant variables – only major observations). Eventually, the following variables are found to be significant: *dep_log_laboutput, dep_outgro, dep_log_tfp, dep_log_tva, for_sales, for_output, cross_section,* and *time_span*. These variables were used to establish models 5-7. Unfortunately, including all these variables meant issues with multicollinearity. To avoid this, *dep_log_laboutput* was excluded from model 5. [5] This was done, based on the correlation. *Dep_log_laboutput* is highly correlated with *cross_section* (0.899653). For model 6 and 7, the same issue occurred leading to the exclusion of both *cross_section* and *dep_log_laboutput*.

The regressions show the importance of the definition of the dependent variable, the definition of the foreign share as well as the timespan used in the original study. This provides some starting points for the analysis of coefficient size.

4.1.2 LOGISTIC REGRESSION RESULTS

Taking the models estimated in section 4.1.1 as a starting point, a new type of regression was estimated. This regression was used to predict the influence of study characteristics on the significance and positivity of the coefficient for the 61 observations. Each model from the OLS regression was re-estimated, using logistic regression techniques. *Coefficient_sigpos* is used as dependent variable. The independent variables are the same as in the OLS regression models.

The idea to look at both sign and significance was also used by Wooster and Diebel (2006). They use a logistic regression for two separate types of models. They use this method to test if the study characteristics influence the significance or the sign of the coefficient in the original study. Following them, these regressions were also estimated in this thesis, see table 7 and 8 in the appendix. Unfortunately, all the models estimated then, show errors with quasi-complete separation. Also, some models show issues with singular covariance. This means that the interpretation of the coefficients and their significance in these models is useless (Field, 2009). As mentioned before, this is probably caused by the lack of variance in the data.

[5] The model was also estimated including *dep_2* and leaving *cross_section* out. This did not lead to different conclusions. Both took the same coefficient sizes, standard errors and probabilities.

The next step taken was to combine the two dependent variables into one variable. This is variable *coefficient_sigpos*. The results from these regressions are mentioned in table 6.

INSERT TABLE 6 HERE

Again, some problems surfaced when establishing these models. The first model established included all *dep_#*, leaving *dep_labpro* as the reference category. Unfortunately, the total model could not be estimated. *Dep_log_tva* has no variance: each observation which uses this definition for the dependent variable produces a positive and significant outcome. After closer examination, it turns out that *dep_log_tva* is only used in the study by Takii (2005). This is the same study from which multiple observations were excluded because they were outliers. Excluding *dep_log_tva* leads to a logistic regression model with no significant results. This is contrary to the OLS regression which shows a significant influence for almost all definitions.

The next two regressions do not show any significant signs for the definition of the foreign share or statistical method used either. For the foreign share another problem is encountered due to high correlations (*for_equity* with *for_empl*: -0.589571, *for_equity* with *for_sales*: -0.553912). In order to see if this causes any problems, the model is also established stepwise. In multiple steps, each time special attention is given to the shifts in coefficient size and standard deviations. Fortunately, the coefficients stay more or less stable. Still, the interpretations of the coefficients should be done with caution. For the third model, the estimation did not cause problems as such. It is notable however that the coefficient of both *stat_olleydiff* and the constant is very small, with a very large standard deviation. As a consequence the probability (p-value) of these two is 1. *Stat_olleydiff* is used in the study by Javorcik (2004). In this study she finds both very small significant and positive results for partial ownership and even smaller insignificant results for full ownership. The results that she finds are very close together which might explain the high p-value and small coefficient. Given these high probabilities is the explanatory value of this model limited.

Model 4 is established to look at all data characteristics. This model can be established but again without any significant results. The issue which is particularly important here, are the high correlations between the four variables in the model. They are all correlated to each other on high levels (>|0.626581|). This leads to difficulties for the reliability of the model.

Furthermore, the total model, model 5, is estimated. The model differs from OLS model 5 because *dep_log_tva* is not present. This is because of the lack of variance for that variable.

This model yields no significant results. The factors included are correlated to each other, but not to extreme levels. This is true, except for the correlation between *cross_section* and *time_span* (-0.824561). In this model all observations are included. The probabilities are around 0.25, except for the *time_span* (0.9310). This means that *coefficient_sigpos* and *time_span* move with each other.

Next, an attempt was made to also establish models 6 and 7, distinguishing between majority and minority ownership. This attempt was however unsuccessful due to quasi-complete separation. It was not possible to establish a model which would be reliable. The variation within the dataset is so small that it does not meet the standards which are necessary to establish such a model.

In order to be able to say something about the influence of minority or majority ownership, model 8 was established. Model 8 includes the dummy for major and the dummy for minor observations. Here, it turns out that *major* does have a significant influence. *Minor* however does not.

In conclusion, attempts were made to establish logistic regression models with different dependent variables. All models have different issues as described above. Eventually for model 1-5 in table 6, reasonable estimates could be made. These estimates did not show any significant results but had several problems. Above, these problems are mentioned. They make it necessary to interpret the results with caution.

4.2 ANALYSIS

In this section, the different models established will be discussed. Both the results from the OLS and logistic regression are compared with other meta-analyses and the theory as described in the theoretical framework. First, the OLS regressions are discussed, in section 4.2.1 the logistic results. A conclusion follows in section 5.

4.2.1 OLS ANALYSIS

In general, for OLS regression, a positive and significant coefficient in a meta-regression means that the study characteristic is associated with a larger value of the dependent variable. In contrast, a negative and significant coefficient will decrease the dependent variable. Insignificant variables do not have these properties and seem to be without consequences. The coefficient shows the shift in the dependent variable when the independent variable changes one unit. Dummy variables show the effect from being part of a certain category in

comparison to the reference category. The constant shows the expected effect when all dummy variables are zero (Hardy, 1993).

For the first model, all dependent variables (except *dep_labprogro*) lead to different results in the OLS regression in comparison with the definition of *dep_labpro*, which is labor productivity, they are all significant. All other definitions yield smaller coefficient sizes in comparison (they are negative). For example, on average for every observation which falls in category 2 the coefficient is 0.14 smaller. This outcome is interesting, given that other meta-analyses do not find a significant influence of this definition. Görg and Strobl (2001) find no significant evidence in their study of FDI influence on productivity spillovers. Their research explicitly focusses on the impact of FDI on productivity, using 22 different studies and only 3 definitions of dependent variables into account. One has as a definition 'other', the other two focus on output (growth of output versus output per worker). Their conclusion is not shared by Wooster and Diebel (2006) who find that definition does influence the size of the coefficients. An explanation might be found in the data used by both studies. Görg and Strobl (2001) use a mix of developed and developing countries while Wooster and Diebel focus purely on developing countries. Also the latter couple establishes three different regressions (sign, size, and significance), while Görg and Strobl focus on a combined result.
A more theoretical reasoning based on the definitions itself, shows that *dep_labpro* and *dep_labprogro* measure very similar concepts. *Dep_labprogro* measures the growth of *dep_labpro*. Given these similarities, it logically follows that both do not significantly differ. The other 4 variables particularly focus on output. Even though output and labor productivity are related concepts, they do measure something different. Output takes, using a standard Cobb-Douglas production function, both labor-productivity and capital-productivity into account[6].

The second model looks at the definitions used as a proxy for foreign share. Four definitions are used in the different studies, namely equity, sales, employment, and gross output. From these definitions, employment and gross output significantly differ from the results when equity is used as a proxy. Both yield on average a larger result. This means that the definition of the foreign presence might be an important explanatory factor for the differences across studies. Again, comparing this to other meta-analyses, this is an interesting conclusion. Görg and Strobl (2001) find that when another definition is used (not employment or output share)

[6] This is mostly used as a basis for the regressions in the original study. See for example Dimelis and Louri (2002)

there is a significant difference in outcome. They however find a negative influence. Wooster and Diebel (2006) also find a negative influence, but not a significant one. Havranek and Irsova (2011) find significant results when foreign presence is measured in employment or output (not assets). The results are thus inconclusive.

Next, a model is established which looks at the different statistical methods used. Not surprisingly, no significant results are found. Authors of the original papers will use a statistical method which best fits their data. For instance, Dimelis and Louri (2002; p. 450) choose for a quantile regression because their data appears to be *"highly skewed with a long right tail"*. The outcome found here is thus the ideal situation: authors are successful in choosing a statistical method which fits their data and does not color it.

Then a model is established which looks at the data characteristics. Here, both the type of data used (cross-section versus panel data) and the time span are significant. Spillovers and the question if the country is developed/developing do not seem to make a difference. Other meta-analyses all find that the type of data influences the coefficient (Görg and Strobl, 2001; Wooster and Diebel, 2006; Havranek and Irsova, 2011). In our meta-analysis, cross-section data yields smaller coefficients. On average cross-section data uses a 0.28 smaller coefficient than if the study uses panel data. Görg and Strobl (2001) find a positive effect, Wooster and Diebel (2006) find both positive and negative effects and last Havranek and Irsova (2011) find positive results. The interpretation of this factor should however be done with caution. Only three studies used in this meta-analysis use cross-section data. They do however yield 24 of the 61 observations included, and they are the older studies. The fact that now more panel data studies are used might be due to the larger availability of data. Given that FDI research only exists for the last fifteen years of so, data gathering has improved greatly. Also, using panel data has a large advantage because one can control for longitudinal factors. Differences which only appear in a couple of years can then be taken into account (Takii, 2005).

The relation between number of years included and type of data can be shown by the correlation. The correlation between both is -0.824561 (significant at 1% level – p-value is 0.000). This means that if cross-section is used, the number of years decreases. Again, by this large correlation, a warning is given for the interpretation of the coefficients. Large correlation might cloud the outcome of the regression.

Theoretically, it makes sense that there will be a difference in data for only one year instead of multiple years. Most mechanisms which are caused by FDI will take a couple of years to show results. For example, labor mobility is not a phenomenon which appears instantly. It will take time for workers to move to other firms and to learn the different techniques. Also using a couple of years of data, will give the opportunity to control for macro-economic factors. Most original studies do include time dummies (which are mostly significant).

Now, model 5-7 include all significant effects from the first four models[7]. In model 5, the foreign share is not significant anymore. The fact that this meta-analysis does not find a significant effect in contrast to others (see above) might be due to the dependent variable used. In this OLS the focus lies explicitly on the size of the effect. Other meta-analyses use the t-statistic which may yield different results. The one study which does look at size only, Wooster and Diebel (2006) also do not find a significant influence for foreign presence when they include this in their final model. They do find however that cross-section data is significant. This is a shared conclusion with this meta-analysis. In both studies the result is that using cross-section data lowers the coefficient. Here, with on average 0.28. The same reasoning to explain this as above applies. Also in model 5, the definitions of the dependent variables used in the original regression stay significant. The total explanatory power of the combined model is a R^2 of 0.316643. This means that the combined variables only explain 31.7% of the variation, which is relatively low.

The next two models are of special interest in this study. The observations are split out into two separate categories: minor and major observations. The results here differ. Not only is the explanatory power of the major model much higher (R^2_{major} = 0.635253 and R^2_{minor} = 0.369242) but also is there a difference in the size of the coefficients. This is too shown in model 6 self: the coefficients have smaller p-values.
The original coefficients for minor observations have a mean of 0.148, for major observations this is a negative -0.02. Based on the two models, it is possible to say that for both types of observations, it is important to look at study characteristics. Both are sensitive for the type of data, measurement of productivity, proxy for foreign presence and the timespan.
Given the difference in coefficient sizes, it can be said that minor observations are more sensitive for changes in study characteristics. This is not surprising given the range of the

[7] Excluding *dep_2* due to multicollinearity issues.

original coefficients. For major observations, the coefficient size ranges between -0.466 and 0.39 (thus 0.856). For minor observations this is between -0.35 and 0.72 (thus 1.07).
No other meta-analyses take this difference into account. Havranek and Irsova (2011) do include a dummy for ownership share (fully owned versus partially owned) in their research for publication bias. They do find that it makes a significant difference if the firm is foreign owned. However given that they have a whole other objective, the results cannot be interpreted one on one with the results here.
This regression cannot be interpreted as proof for the different theories described above on the importance of ownership shares, however it can proof that there is a difference between the two types. When conducting a study, an author should be aware of this and use the information provided when he makes his research design.

4.2.2 LOGISTIC REGRESSION ANALYSIS

In general, in logistic regression, the coefficient describes the change in the logit of the dependent variable (here *coefficient_sigpos*) when the independent variables change one unit. The logit is the natural logarithm (ln) of the odds that the dependent variable will occur. In this case it is the natural logarithm of the odds that the coefficient is both positive and significant (5% level or better). It's important to realize that the coefficients show the influence of a one unit change *all other things constant* (Field, 2009).

As mentioned above in the results sections, some problems occurred when establishing the logistic regression. The main issue was rather simple, namely the lack of a sufficient number of observations. This leads to a serious warning when interpreting these results. Due to these problems only 6 of the 8 models can be established. For these six models, just one model gives a significant outcome. The reason for this is most likely due to the problems described in the results section. Another logistic regression on FDI, done by Wooster and Diebel (2006) does find some factors to have a significant influence. For instance, they find that if output is used to describe the foreign presence, then it is more likely that the coefficient is significant.

The one model (8, table 5) which yields significant results shows that if the observation is from majority ownership, the chance on a positive and significant coefficient decreases. When looking at the raw data, one can see that from the 24 majority observations, only 6 yield a positive and significant outcome. For the minority observations, 19 of the 28 observations yield this outcome. In this an explanation might be found for the significance of the *major* dummy.

5. CONCLUDING REMARKS

Over the last ten years various papers have been published on the influence of ownership shares on productivity spillovers caused by foreign direct investment. These studies show mixed results. This thesis provided an overview of the literature of this field, both theoretical and empirical.

Based on the literature, Aitken and Harrison (1999) found for instance that if a multinational works in a more productive sector, the likelihood of a positive relation between industrial productivity and foreign presence increases. They also found that this does not directly have to be evidence of a positive spillover effect . This is because the positive effect turns into a negative effect when industry dummies are included. This conclusion is already a sign that research design plays an important role in study outcomes.

In this thesis, using methods of meta-regression, an empirical study was done to find out if study characteristics could explain some of the mixed evidence found. Based on a database with 10 different studies, 61 observations, it is here concluded that some study characteristics do have a significant influence on the outcome of studies. Both the influence of these characteristics on the size of the coefficient and the sign/significance of the coefficient are discussed.

On average it is shown that the definition used in the original study for both foreign presence and for the dependent variable has an influence on the coefficient size. For example, using sales or gross output as a proxy for foreign share, gives a larger outcome then when equity share is used. Another significant factor is the number of years included in the study. Increasing the number of years, gives a different result. It was shown that this is related to the use of cross-section data instead of panel data. Cross-section yields on average smaller coefficients. Another important conclusion was the difference in outcomes when only major or minor observations where included. This outcome does point in the direction that ownership shares do produce different results.

With the use of logistic regression techniques, different attempts were made to find the influence of study characteristics on the sign and significance of the original coefficients. Unfortunately, this made the several shortcomings of this study more apparent. Due to the lack of variation, the insufficient number of observations, high correlation between variables and overall low number of studies included in this thesis only some models could be

established. These models did not yield any significant results for the data and study characteristics. It was however established that if the observation was for majority ownership share, the chance on a positive and significant outcome decreases. The minor dummy was however not significant.

Except for the shortcomings stated above, another problem was with using coefficient size as a dependent variable for the OLS regression. This meant including all observations, even if they weren't significant. It would have been preferred to use the *t-statistic* as a dependent variable, but this information wasn't available for all studies. The *t-statistic* should have been used because in the different studies different measurements are used for the foreign ownership share. In order to be able to compare the different studies, a universal unit of measurement has to be found; a role the t-statistic can fulfill (Stanley & Jarell, 1989). Furthermore, the limitations which were described in section 2.4 apply here as well. The sample was constructed to the best of my knowledge, but it might have been more robust if other studies were included as well.

Concluding, several study characteristics influence the size of the coefficient in a study of FDI spillovers. The influence of these factors differs between major and minor observations.
Thus this thesis provides some findings which clearly indicate that a research design should be built with caution. Then, the main point of this thesis should be a warning for future researchers in this field. Leaving the shortcomings of this thesis as it is, it provides a strong signal that study characteristics play a role in explaining different observations. It also underlines the importance of more research on the role of ownership shares.

6. BIBLIOGRAPHY

Abraham, F., Konings, J., & Slootmaekers, V. (2007). FDI Spillovers, Firm Heterogenity and Degree of Ownership: Evidence form Chinese Manufacturing. *CEPR Discussion Paper 6573.* London: Centre for Economic Policy Research.

Aitken, B., & Harrison, A. (1999). Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela. *American Economic Review*(72), 605-18.

Alfaro, L., & Rodíiquez-Clare, A. (2004). Multinationals and Linkages: An Empirical Investigation. *Economia, 4*(2), 113-69.

Baldwin, J. (1996). Productivity Growth, Plant Turnover and Restructuring in the Canadian Manufacturing Sector. In D. Mayes, *Sources of Productivity Growth.* Cambridge: Cambridge University Press.

Barbosa, N., & Eiriz, V. (2009). Linking Corporate Productivity to Foreign Direct Investment: An Empirical Assessment. *International Business Review, 18*(1), 1-13.

Barrios, S., Dimelis, S., Louri, H., & Strobl, E. (2004). Efficiency Spillovers from Foreign Direct Investment in the EU Periphery: A Comparative Study of Greece, Ireland, and Spain. *Review of World Economics, 140*(4), 688-705.

Bellak, C. (2004). How Domestic and Foreign Firms Differ and Why Does It Matter? *Journal of Economic Surveys*.

Blomström, M., & Kokko, A. (1998). Multinational Corporations and Spillovers. *Journal of Economic Surveys*(12), 247-77.

Blomström, M., & Sjöholm, F. (1999). Technology Transfer and Spillovers: does Local Participation with Multinatinals matter? *European Economic Review*(43), 915-23.

Blomström, M., Globerman, S., & Kokko, A. (2001). The Determinants of Host Country Spillovers from Foreign Direct Investment. In N. P. ed., *Inward Investment, Technological Change, and Growth.* London: Palgrave.

Card, D., & Krueger, A. (1995). Time-Series Minimum-Wage Studies: A Meta-analysis. *American Economic Review, 85*, 238-43.

Coughlin, C. (1983). The Relationship Between Foreign Ownership and Technology Transfer. *Journal of Comparative Economics, 7*(4), 400-14.

Das, S. (1987). Externalities, and Technology Transfer through Multinational Corporations: A Theoretical Analysis. *Journal of International Economics, 22*(1-2), 171-82.

Dimelis, S., & Louri, H. (2002). Foreign Ownership and Production Efficiency. *Oxford Economic Papers*(54), 449-69.

Dimelis, S., & Louri, H. (2004). Foreign Direct Investment and Technology Spillovers: Which Fiirms really Benefit? *Review of World Economics, 140*(2), 230-53.

Dunning, J. (1981). Explaining the International Direct Investment Position of Countries: Towards a Dynamic or Developmental Approach. *Review of World Economics, 117*(1), 30-64.

Dunning, J. (1993). *Multinational Enterprises and the Global Economy.* Workingham, England: Addison Wesley.

Dunning, J., & Lundan, S. (2008). *Multinational Enterprises and the Global Economy.* Northhampton, Massachusetts: Edward Elgar Publishing, Inc.

Field, A. (2009). *Discovering Statistics using SPSS.* London: SAGE Publications Ltd.

Globerman, S., & Chen, V. (2010). *Best Policy Practices for Promoting Inward and Outward Foreign Direct Investment.* The Conference Board of Canada.

Globerman, S., Ries, J., & Vertinsky, I. (1994). The Economic Performance of Foreign Affliates in Canada. *Canadian Journal of Economics*, 143-56.

Görg, H., & Greenaway, D. (2004). Much Ado About Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment? *The World Bank Research Observer, 19*(2), 171-197.

Görg, H., & Strobl, E. (2001). Multinational Companies and Productivity Spillovers: A Meta-analysis with a Test for Publication Bias. *The Economic Journal, 111*(475), F723-39.

Görg, H., & Strobl, E. (2005). Spillovers from Foreign Firms through Workers Mobility: An Empirical Investigation. *Scandinavian Journal of Economics, 107*(4), 693-709.

Griffith, R. (1999, June). Using the ARD Establishment Level Data to Look at Foreign Ownership and Productivity in the UK. *Economic Journal, 109*, F416-F442.

Griffith, R., Redding, S., & van Reenen, J. (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Countries. *Review of Economics and Statistics, 14*(5), 922-40.

Hagen, A. (1997). *Deutschen Direktinvestionen in Grossbritannien.* Jena: Franz Steiner Verlag.

Hale, G., & Long, C. (2006). What Determines Techonological Spillovers from Foreign Direct Investment? Evidence from China. *Yale Economic Growth Center Discussion Paper 934.* New Haven, CT.

Hardy, M. (1993). Regression with Dummy variables. *Sage University Paper series on Quantatitve Applications in the Social Sciences, 07-093.* Newbury Park, CA: Sage.

Harris, R., & Robinson, C. (2002). The Effect of Foreign Acquisitions on Total Factor Productivity: Plant-level Evidence form UK manufacturing 1987-1992. *Review of Economics and Statistics, 84*(3), 562-68.

Haskel, J., Pereira, S., & Slaughter, M. (2002). Does Inward Foreign Direct Investment Boost the Productity of Domestic Firms? *NBER Working Paper Series No. 8724.*

Havranek, T., & Irsova, Z. (2011). Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is. *Journal of Internatonal Economics, 85*(2), 234-44.

Iyer, C. (2008). *Effect of Foreign Firms and Their Linkages on Domestic Firms in the Indian Manufacturing Industry: Evidence from 1989-2004.* Indira Gandhi Institute of Development Research (IGIDR) Working Paper Series. Opgehaald van http://www.cds.edu/globelics/chidambaram.pdf

Javorcik, B., & Spatareanu, M. (2008). To Share or not to Share: Does Local Participation Matter for Spillover from Foreign Direct Investment? *Journal of Development Economics, 85*(1-2), 194-217.

Javorick, B. (2004b). Does Foreign Direct Investment Increase the Productivity of Domestic Fims? In Search of Spillovers through Backward Linkages. *American Economic Review, 94*(3), 605-27.

Khalifah, N., & Adam, R. (2009). Productivity Spillovers from FDI in Malysian Manufacturing: Evidence from Micro-panel Data. *Asian Economic Journal, 23*(2), 143-67.

Kokko, A., Tansini, R., & Zejan, M. (1996). Local Technological Capability and Productivity Spillovers from FDI in the Uruguayan Manufacturing Sector. *Journal of Development Studies, 32*(4), 602-11.

Konings, J. (2000). The Effects of Direct Foreign Investment on Domestic Firms: Evidence from Firm Level Panel Data in Emerging Economies. *CEPR Discussion papers*.

Kugler, M. (2006). Spillovers from foreign direct investment: Within or between industries? *Journal of Development Economics, 80*(2), 447-77.

Markusen, J. (1995). The Boundaries of Multinational Enterprises and the Theory of International Trade. *Journal of Economic Perspectives, 9*(2), 168-89.

Markusen, J., & Trofimenko, N. (2007). Teaching Locals New Tricks: Foreign Experts as a Channel of Knowledge Transfer. *NBER Working Paper 12872.* Cambridge, Mass.: National Bureau of Economic Research.

Markusen, J., & Venables, A. (1999). Foreign Direct Investment as a Catalyst for Industrial Development. *European Economic Review*(43), 335-36.

Müller, T., & Schnitzer, M. (2006). Technology Transfer and Spillovers in International Joint Ventures. *Journal of International Economics, 68*(2), 456-68.

Perez, T. (1997). Multinational Enterprises and Technological Spillovers: An Evolutionary Model. *Evolutionary Economics, 7*(2), 169-92.

Poole, J. (2007). *Multinational Spillovers through Worker Turnover*. Department of Economics, University of California Santa Cruz.

Ramachandaram, V. (1993). Technology Transfer, Firm Ownership, and Investment in Human Capital. *Review of Economics and Statistics, 75*(4), 664-70.

Rodriquez-Clare, A. (1996). Multinationals Linkages and Economic Development. *American Economic Review, 86*(4), 852-73.

Schoors, K., & Tol, B. v. (2002). Foreign Direct Investment Spillovers within and between Sectors: Evidence from Hungarian Data. *Ghent University Working Paper*.

Smeets, R. (2008). Collecting the Pieces of the FDI Knowledge Spillovers Puzzle. *World Bank Research Observer, 23*, 107-38.

Stanley, T. (2001). Wheat from Chaff: Meta-Analysis as Quantative Literature Review. *Journal of Economic Perspectives, 15*(3), 131-50.

Stanley, T., & Jarell, S. (1989). Meta-Regression Analysis: A Quantative Method of Literature Surveys. *Journal of Economic Surveys, 3*, 161-70.

Suyanto, R., & Bloch, H. (2009). Does Foreign Direct Investment Lead to Productivity Spillovers? Firm Level Evidence from Indonesia. *World Development, 37*(12), 1861-76.

Takii, S. (2005). Productivity Spillovers and Characteristics of Foreign Multinational Plants in Indonesian Manufacturing. *Journal of Development Economics, 76*, 521-42.

The Economist. (1997, April 19). Multinationals in China: Going it Alone. *The Economist*.

The Economist. (2012, January 28). Foreign direct investment. *The Economist*.

Wang, J.-Y., & Blomström, M. (1992). Foreign Investment and Technology Transfer: A Simple Model. *European Economic Review, 36*(1), 137-55.

Wooster, R., & Diebel, D. (2006). Productivity Spillovers from Foreign Direct Investment in Developing Countries: A Meta-Regression Analysis. *Working Paper*.

7. APPENDIX

TABLE 1: STUDIES OVERVIEW

Overview of Studies included in the Meta-analysis

	Author(s)	Country	Year	Data	Number of observations	Aggregation	Results
1	Blomström & Sjöholm (1999)	Indonesia	1991	Cross-section	13.663	Firm (manufacturing)	Minority and majority FDI shares had equal spillover effects
2	Dimelis & Louri (2002)	Greece	1997	Cross-section	4.056	Firm (manufacturing)	Minority FDI shares had greater spillover effects than majority FDI shares
3	Schoors & Van der Tol (2002)	Hungaria	1997-1998	Cross-section	1.021	Firm	Although insignificant, the influence of majority ownership is larger than minority ownership. When distinguishing between minority, majority <95% and majority >95% is it shown that majority >95% has an large (insignificant) coefficient.
4	Dimelis & Louri (2004)	Greece	1997	Cross-section	3.742	Firm	Majority-held foreign firms that exhibit higher productivity, spillovers are important for small domestic firms and stem mostly from small joint ventures where the foreign partner owns a minor part of equity.
5	Barrios, Dimelis, Louri and Strobl (2004)	Greece	1992,1997	Cross-section	2.301	Firm	For Greece and Spain it is important if minority ownership firms are included in the analysis. For Greece if also minority firms (10-30 % ownership) are considered, the effect of spillovers turns (insignificant) positive. For Spain the
		Spain	1992,1997	Cross-section	658	Firm	

							inclusion of those firms reduces the influence marginally but remains significant.
6	Javorcik (2004b)	Lithuania	1996-2000	Panel	4.000	Firm (manufacturing)	Shared ownership and domestic ownership had positive spillover effects. Only firms with shared ownership create backward spillovers. No difference between minority and majority ownership.
7	Takii (2005)	Indonesia	1990-1995	Panel	Not mentioned in the article	Firm	Extent of the spillovers decreases when there are additional majority- or wholly owned foreign plants.
8	Abraham, Konings & Slootmaekers (2007)	China	2000-2004	Panel	17.645	Plant	Minority FDI shares had a negative influence on locally owned firms. Majority FDI had no effect on locally owned firms.
9	Javorcik & Spatareanu (2008)	Romania	1999-2003	Panel	13.129	Firm	Shared foreign and domestic ownership had positive vertical spillover effects and negative horizontal spillover effects. Wholly owned subsidiaries do not create these effects.
10	Khalifah & Adam (2009)	Malaysia	2000-2004	Panel	4.504	Establishment	There is no difference in the magnitude of spillovers from minority or majority owned foreign subsidiaries.

TABLE 2: VARIABLE OVERVIEW

Overview of Variables used in the Meta-analysis

Variable	**Description *(variable name)***	**Dummies**
Dependent meta-regression	Coefficient size of foreign direct investment in the original regression *(coefficient_size)*	Coefficient size in the original regression
	Coefficient is positive and significant (≤5%) *(coefficient_sigpos)*	1 = positive and significant 0 = negative or insignificant
	Coefficient is positive *(coefficient_pos)*	1 = positive 0 = negative
	Coefficient is significant (≤5%) *(coefficient_sig)*	1 = significant 0 = insignificant
Data	Type of data used in the analysis *(Cross_section)*	1 = data are cross-section 0 = data are panel
	Developing *(developing)*	1 = developing country 0 = developed country
Time span	Number of years used in the analysis *(time_span)*	Total number of years
Spillovers	Spillovers are explicitly considered in the regression *(spillovers)*	1 = yes 0 = no
Dependent variable	Dependent variable is labor productivity *(dep_labpro)*	1 = yes 0 = no
	Dependent variable is logarithm of labor-output ratio *(dep_log_laboutput)*	1 = yes 0 = no
	Dependent variable is labor productivity growth *(dep_labprogro)*	1 = yes 0 = no
	Dependent variable is growth of output *(dep_outgro)*	1 = yes 0 = no
	Dependent variable is logarithm of total factor productivity *(dep_log_tfp)*	1 = yes 0 = no
	Dependent variable is logarithm of total value added *(dep_log_tva)*	1 = yes 0 = no
Measurement of foreign share	The foreign share is measured as foreign share in equity *(for_equity)*	1 = yes 0 = no
	The foreign share is measured as foreign share in employment *(for_empl)*	1 = yes 0 = no
	The foreign share is measured as foreign share in sales *(for_sales)*	1 = yes 0 = no
	The foreign share is measured as foreign share in gross output *(for_output)*	1 = yes 0 = no
Method of estimation	The method of estimation is OLS *(stat_OLS)*	1 = yes 0 = no
	The method of estimation is quantile regression *(stat_quanregr)*	1 = yes 0 = no
	The method of estimation is Olley-packs *(stat_olley)*	1 = yes 0 = no

	The method of estimation is OLS first differences *(stat_OLSdiff)*	1 = yes 0 = no
	The method of estimation is Olley-Packs first differences *(stat_olleydiff)*	1 = yes 0 = no
Major/minor	Major if wholly owned/majority owned *(major)*	1 = yes 0 = no
	Minor if joint venture/minority owned/partially owned *(minor)*	1 = yes 0 = no

TABLE 3: DESCRIPTIVE STATISTICS

Variable name	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Coefficient_size	61	1.2460	-.4660	.7800	.091636	.2445982	.060
Coefficient_sigpos	61	1.00	.00	1.00	.5410	.50245	.252
Coefficient_pos	61	1.00	.00	1.00	0.8197	.38765	.150
Coefficient_sig	61	1.00	.00	1.00	0.6066	.49257	.243
Cross_section	61	1.00	.00	1.00	.3934	.49257	.243
Developing	61	1.00	.00	1.00	.5902	.49588	.246
Time_span	61	5.00	1.00	6.00	2.9836	1.95356	3.816
Spillovers	61	1.00	.00	1.00	.3770	.48867	.239
Dep_labpro	61	1.00	.00	1.00	.1639	.37329	.139
Dep_log_laboutput	61	1.00	.00	1.00	.3443	.47907	.230
Dep_labprogro	61	1.00	.00	1.00	.0656	.24959	.062
Dep_outgro	61	1.00	.00	1.00	.1475	.35759	.128
Dep_log_tfp	61	1.00	.00	1.00	.2295	.42401	.180
Dep_log_tva	61	1.00	.00	1.00	.0492	.21804	.048
For_equity	61	1.00	.00	1.00	.6393	.48418	.234
For_empl	61	1.00	.00	1.00	.1639	.37329	.139
For_sales	61	1.00	.00	1.00	.1475	.35759	.128
For_output	61	1.00	.00	1.00	.0492	.21804	.048
Stat_OLS	61	1.00	.00	1.00	.5246	.50354	.254
Stat_quanregr	61	1.00	.00	1.00	.1639	.37329	.139
Stat_olley	61	1.00	.00	1.00	.0984	.30027	.090
Stat_OLSdiff	61	1.00	.00	1.00	.1475	.35759	.128
Stat_olleydiff	61	1.00	.00	1.00	.0656	.24959	.062
Major	61	1.00	.00	1.00	.3934	.49257	.243
Minor	61	1.00	.00	1.00	.4590	.50245	.252

TABLE 4: CORRELATIONS

Overview of Correlations between the Independent Variables used in Meta-regression 1 (table 5)

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1.	Coefficient_size	1																						
2.	Coefficient_sigpos	.431**	1																					
3.	Coefficient_sig	.259*	.874**	1																				
4.	Coefficient_pos	.638**	.509**	.233	1																			
5.	Cross_section	.080	.270*	.168	.290*	1																		
6.	Developing	-.165	-.166	-.057	-.304*	-.762**	1																	
7.	Time_span	-.312*	-.229	-.041	-.466**	-.825**	.784**	1																
8.	Spillovers	-.202	-.098	-.135	-.251	-.627**	.648**	.810**	1															
9.	Dep_labpro	.251	.052	-.006	.208	-.085	.369**	-.225	-.344**	1														
10.	Dep_log_laboutput	-.021	.252	.160	.250	.900**	-.869**	-.742**	-.564**	-.321*	1													
11.	Dep_labprogro	.368**	-.155	-.193	.124	-.213	-.318*	-.134	-.206	-.117	-.192	1												
12.	Dep_outgro	-.118	-.081	-.138	.195	-.335**	.347**	.433**	.535**	-.184	-.301*	-.110	1											
13.	Dep_log_tfp	-.133	-.045	-.039	-.454**	-.440**	.455**	.568**	.702**	-.242	-.395**	-.145	-.227	1										
14.	Dep_log_tva	-.352**	-.247	.183	-.485**	-.183	.190	.354**	-.177	-.101	-.165	-.060	-.095	-.124	1									

15. For_equity	-.173	.062	-.046	-.086	-.024	.068	.328**	.584**	-.497**	.113	-.353**	.312*	.410**	-.303*	1								
16. For_empl	.012	-.214	-.006	-.138	-.085	-.261*	-.019	-.344**	-.196	-.041	.598**	-.184	-.242	.514**	-.590**	1							
17. For_sales	.083	.105	.051	.195	-.051	.065	-.283*	-.324*	.565**	-.010	-.110	-.173	-.227	-.095	-.554**	-.184	1						
18. For_output	.227	.057	.028	.107	.282*	.190	-.233	-.177	.514**	-.165	-.060	-.095	-.124	-.052	-.303*	-.101	-.095	1					
19. Stat_OLS	-.058	-.086	-.028	-.276*	.095	.141	-.110	-.275*	.422**	-.001	-.278*	-.437**	.051	.217	-.442**	.067	.396**	.217	1				
20. Stat_quanregr	-.092	.141	.085	.093	.550**	-.531**	-.453**	-.344**	-.196	.611**	-.117	-.184	-.242	-.101	.333**	-.196	-.184	-.101	-.465**	1			
21. Stat_olley	.048	.194	.266*	.012	-.266*	.275*	.344**	.425**	-.146	-.239	-.087	-.137	.605**	-.075	.248	-.146	-.137	-.075	-.347**	-.146	1		
22. Stat_OLSdiff	.189	-.173	-.233	.195	-.335**	-.029	.147	.153	-.184	-.301*	.637**	.479**	-.227	-.095	-.073	.315*	-.173	-.095	-.437**	-.184	-.137	1	
23. Stat_olleydiff	-.072	-.022	-.058	.124	-.213	.221	.276*	.341**	-.117	-.192	-.070	.637**	-.145	-.060	.199	-.117	-.110	-.060	-.278*	-.117	-.087	-.110	1

TABLE 5: STUDY CHARACTERISTICS THAT ANALYZE THE EFFECT SIZE

Results of meta-regression – dependent variable: coefficient size (coefficient_size)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Constant	0.23*** (0.027)	0.06 (0.039)	0.08 (0.054)	0.52*** (0.134)	0.42** (0.163)	-0.13*** (0.036)	-0.21* (0.119)
Cross_section	-	-	-	-0.28** (0.118)	-0.28** (0.123)	□	□
Developing	-	-	-	0.01 (0.103)	-	-	-
Time_span	-	-	-	-0.12*** (0.025)	-0.034 (0.033)	0.16*** (0.032)	0.30** (0.120)
Dep_labpro	•	-	-	-	-	-	-
Dep_log_laboutput	-0.14*** (0.045)	-	-	-	□	□	□
Dep_labprogro	0.19 (0.131)	-	-	-	-	-	-
Dep_outgro	-0.21*** (0.028)	-	-	-	-0.23*** (0.007)	-0.67*** (0.127)	-1.24** (0.483)
log_tva	-0.20* (0.098)	-	-	-	-0.22** (0.097)	-0.86*** (0.167)	-1.08** (0.513)
Dep_log_tva	-0.51*** (0.042)	-	-	-	-0.499*** (0.046)	-1.093*** (0.161)	-1.93*** (0.602)
For_equity	-	•	-	-	-	-	-
For_empl	-	0.04 (0.117)	-	-	-	-	-
For_sales	-	0.08* (0.044)	-	-	-0.13 (0.078)	0.01 (0.018)	-0.182 (0.117)
For_output	-	0.27*** (0.049)	-	-	-0.05 (0.134)	0.36*** (0.018)	0.18*** (0.025)
Spillovers	-	-	-	0.11 (0.089)	-	-	
Stat_OLS	-	-	•	-	-	-	
Stat_quanregr	-	-	-0.04 (0.055)	-	-	-	
Stat_olley	-	-	0.05 (0.068)	-	-	-	
Stat_OLSdiff	-	-	0.12	-	-	-	

			(0.105)				
Stat_olleydiff	-	-	-0.05 (0.055)	-	-	-	
# of obs.	61 (all)	61 (all)	61 (all)	61 (all)	61 (all)	24 obs (major only)	28 obs (minor only)
F	5.233081	1.333567	0.678899	3.766474	3.508322	4.934611	2.048881
R^2	0.322371	0.065584	0.046250	0.211999	0.316643	0.635253	0.369242

Notes:

1. Regressions estimated using White heteroskedasticity-consistent standard errors. These are in parentheses.
2. ***, **, * symbolize statistical significance at 1%, 5% and 10%.
3. • reference dummy
4. □ left out due to multicollinearity

TABLE 6: STUDY CHARACTERISTICS THAT CONTRIBUTE TO SIGNIFICANCE AND POSITIVITY

Results of logistic linear regression – dependent variable: dummy = 1if coefficient is significant on ≤5% and positive (total)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Constant	-0.15 (0.581)	0.26 (0.334)	4.95E-11 (0.369)	0.02 (1.319)	-1.16 (1.832)	-1.88* (1.117)	0.00 (0.632)	2.08* (1.088
Cross_section	-	-	-	1.13 (1.123)	1.97 (1.479)	□	37.543 ($5.023^{E}7$)	-
Developing	-	-	-	0.51 (1.040)	-	-	-	-
Time_span	-	-	-	-0.33 (0.404)	-0.04 (0.457)	□	□	-
Dep_labpro	•	-	-	-	-	-	-	-
Dep_log_labo utput	1.07 (0.769)	-	-	-	□	□	□	-
Dep_labprogr o	-0.94 (1.338)	-	-	-	-	-	-	-
Dep_outgro	-0.07 (0.910)	-	-	-	1.13 (1.295)	-149.32 ($3.406^{E}32$)	37.54 ($7.10^{E}7$)	-
Dep_log_tfp	0.15 (0.805)	-	-	-	1.35 (1.222)	□	□	-
Dep_log_tva	□	-	-	-	□	-149.32 ($4.817^{E}32$)	-151.203 ($6.812^{E}32$)	-
For_equity	-	•	-	-	-	-	-	-
For_empl	-	-1.11 (0.788)	-	-	-	-	-	-
For_sales	-	0.44 (0.804)	-	-	1.44 (1.244)	2.30 (1.523)	0.00 (1.183)	-
For_output	-	0.44 (1.310)	-	-	-0.08 (1.404)	-	-188.75 ($6.812^{E}32$)	-
Spillovers	-	-	-	1.04 (1.141)	-	-	-	-
Stat_OLS	-	-	•	-	-	-	-	-
Stat_quanregr	-	-	0.84 (0.809)	-	-	-	-	-
Stat_olley	-	-	1.61 (1.201)	-	-	-	-	-
Stat_OLSdiff	-	-	-0.69	-	-	-	-	-

			(0.825)					
Stat_olleydiff	-	-	-5.57E-11 (1.107)	-	-	-	-	-
Minor	-	-	-	-	-	-	-	-1.33 (1.164)
Major	-	-	-	-	-	-	-	-3.18*** (1.190)
# of obs.	61 (all)	61 (all)	61 (all)	61 (all)	61 (all)	24 obs (major only)	28 obs (minor only)	

Notes:

1. Regressions estimated using logistic regression estimations. Standard errors are in parentheses.
2. ***, **, * symbolize statistical significance at 1%, 5% and 10%.
3. • Reference dummy's
4. □ left out of the estimation
5. Models 6 and 7 have problems with quasi-complete separation. The results in these models have to be ignored.

TABLE 7: STUDY CHARACTERISTICS THAT CONTRIBUTE TO SIGN – POSITIVE/NEGATIVE

Results of logistic linear regression – dependent variable: dummy = 1 for positive coefficients (coefficient_pos)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Constant	0.53 (0.412)	1.35*** (0.410)	0.94** (0.41)	5.65** (2.231)	3.88 (1.407)	2.73 (1.693)	3.00*** (1.084)
Cross_section	-	-	-	□	□	□	□
Developing	-	-	-	3.96 (5.497)	-	-	-
Time_span	-	-	-	-1.88 (1.305)	-0.75** (0.314)	-0.42 (0.379)	□
Dep_labpro	•	-	-	-	-	-	-
Dep_log_laboutput	2.47** (1.137)	-	-	-	□	□	□
Dep_labprogro	34.27 (18422212)	-	-	-	-	-	-
Dep_outgro	34.27 (12281875)	-	-	-	35.48 (19207057)	□	□
Dep_log_tfp	□	-	-	-	□	-2.24 (1.511)	-2.30 (1.420)
Dep_log_tva	□	-	-	-	-35.06 (33467179)	□	□
For_equity	-	•	-	-	-	-	-
For_empl	-	-0.51 (0.823)	-	-	-	-	-
For_sales	-	34.05 (15487348)	-	-	33.08 (16181925)	33.25 (29819025)	□
For_output	-	34.04 (27158989)	-	-	32.51 (33044358)	33.04 (53472623)	32.09 (46670836)
Spillovers	-	-	-	0.629 (1.511)	-	-	-
Stat_OLS	-	-	•	-	-	-	-
Stat_quanregr	-	-	1.26 (1.174)	-	-	-	-
Stat_olley	-	-	0.67 (1.215)	-	-	-	-
Stat_OLSdiff	-	-	35.57	-	-	-	-

			(42651260)				
Stat_olleydiff	-	-	35.11	-	-	-	-
			(34766152)				
# of obs.	61 (all)	61 (all)	61 (all)	61 (all)	61 (all)	24 obs (major only)	28 obs (minor only)
Pearson statistic	0.842105	0.859649	0.857143	1.207281	0.712236	1.124440	1.080000

Notes:

1. Regressions estimated using logistic regression estimation (logit – binominal count). Standard errors are in parentheses.
2. ***, **, * symbolize statistical significance at 1%, 5% and 10%.
3. • reference dummy's
4. □ left out due to multicollinearity or separation issues
    a. Model 1: dep_log_tfp and dep_log_tva are all captured in dep_var2 and thus left out. Quasi-complete separation
    b. Model 2: quasi-complete separation
    c. Model 3: quasi-complete separation
    d. Model 4: cross-section left out due to singular covariance.
    e. Model 5: cross-section left out due to singular covariance, quasi-complete separation
    f. Model 6: quasi-complete separation
    g. Model 7: quasi-complete separation

TABLE 8: STUDY CHARACTERISTICS THAT CONTRIBUTE TO SIGNIFICANCE

Results of logistic linear regression – dependent variable: dummy = 1if coefficient is significant on ≤5% (coefficient_sig)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Constant	0.74** (0.401)	0.36 (0.337)	0.49 (0.329)	0.20 (0.606)	0.56 (0.621)	-0.79 (0.927)	0.70 (0.975)
Cross_section	-	-	-	□	□	□	□
Developing	-	-	-	-0.25 (0.904)	-	-	-
Time_span	-	-	-	0.32 (0.371)	-0.06 (0.184)	0.07 (0.272)	0.21 (0.383)
Dep_labpro	•	-	-	-	-	-	-
Dep_log_laboutput	□	-	-	-	□	□	□
Dep_labprogro	-1.84 (1.270)	-	-	-	-	-	-
Dep_outgro	-0.97 (0.807)	-	-	-	-0.46 (0.890)	□	□
Dep_log_tfp	-0.45 (0.692)	-	-	-	□	-0.25 (1.26)	-1.04 (1.559)
Dep_log_tva	34.58 (19180098	-	-	-	32.77 (8713658)	□	□
For_equity	-	•	-	-	-	-	-
For_empl	-	0.04 (0.749)	-	-	-	-	-
For_sales	-	0.33 (0.805)	-	-	0.24 (0.849)	□	□
For_output	-	0.3302 (1.311)	-	-	0.20 (1.377)	35.97 (57755413)	-97.90 (1.25E+21)
Spillovers	-	-	-	-1.48 (1.232)	-	-	-
Stat_OLS	-	-	•	-	-	-	-
Stat_quanregr	-	-	□	-	-	-	-
Stat_olley	-	-	34.77 (22113311)	-	-	-	-

Stat_OLSdiff	-	-	-1.18 (0.802)	-	-	-	-
Stat_olleydiff	-	-	-0.486	-	-	-	-
# of obs.	61 (all)	61 (all)	61 (all)	61 (all)	61 (all)	24 obs (major only)	28 obs (minor only)
Pearson statistic	1.035714	1.070175	0.964912	1.053468	1.092450	1.150622	1.096032

Notes:

6. Regressions estimated using logistic regression estimation. Standard errors are in parentheses.
7. ***, **, * symbolize statistical significance at 1%, 5% and 10%.
8. • reference dummy's
9. □ left out due to multicollinearity
    a. Model 1: dep_log_laboutput left out due to singular covariance, quasi-complete separation
    b. Model 3: quasi-complete separation
    c. Model 4: cross_section left out due to singular covariance
    d. Model 5: cross_section left out due to singular covariance, quasi-complete separation
    e. Model 6: quasi-complete separation
    f. Model 7: quasi-complete separation