Foreign Direct Investment and International Knowledge Diffusion: Evidence Using Patent Citation Data

Stefan Fairbrother

Foreign Direct Investment and International Knowledge Diffusion: Evidence Using Patent Citation Data

Stefan Fairbrother

Supervised by: Dr Chris Hajzler

A thesis submitted in partial fulfilment of requirements for the degree of a Master of Business in Economics

November 2012

Department of Economics

University of Otago

Dunedin

ABSTRACT

This thesis examines the effect of foreign direct investment (FDI) on knowledge diffusion between countries. Standard theory suggests that the extent technological knowledge diffuses across national borders plays an important role in determining long run economic growth and income equality. Because governments of developing countries often compete to attract FDI using tax credits and subsidies, an important question is whether knowledge spillovers are one of the benefits of such policies. To measure knowledge flows, we use all patent citations from the United States Patent and Trademark Office (USPTO), yielding an unbalanced panel of patent citations from 90 countries between the years 1985 to 2010. We model knowledge flows using a gravity framework, and test whether bilateral FDI increases the likelihood that patent applicants in the host country will cite patents originating from the investing country. For developing countries, we find that aggregate FDI inflows have a negative effect on knowledge flows. However, stronger intellectual property rights in the host country stimulate relatively more knowledge flows via FDI and may lead to positive knowledge spillovers in the medium term.

ACKNOWLEDGEMENTS

First and foremost I would like thank my supervisor Dr Chris Hajzler, I truly appreciate all the time and effort you spent to guide me through this thesis. Also to Dr Murat Genç, thank you very much for your invaluable insights and comments throughout this year. To the rest of the economics department, I thank everyone for your inputs both directly and indirectly throughout my studies, I am very proud to have been part of such an awesome department. To my fellow Mbus lads, you're a great bunch which has made the long road that much easier.

To my parents John and Rita, thank you so much for giving me every opportunity to do what I want to do and supporting me every inch of way. Words will never thank you enough. Natalie, you are awesome and I am so grateful to have you in my life. You make me smile every single day, what more can I ask for!

I have been extremely privileged to be a recipient of the Macandrew-Stout Postgraduate Scholarship in Economics, which undoubtedly helped me throughout my postgraduate studies. Lastly, I would like to acknowledge that it is a great honour to be a recipient of the Sophie Kate Elliott Memorial Prize.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The extent that technology diffuses is a key factor determining the long run rate of growth and wealth distribution of an economy. "New" growth theory has emphasised the significance of research and development (R&D) and the subsequent technology for sustained economic growth (see Romer, 1986 and Grossman and Helpman, 1991). Technology is non-rival as well as non-excludable; therefore the benefits from R&D are partly private and partly public. It is the contribution that R&D makes to the urn of knowledge available in the economy that provides the "engines" of economic growth. This is because technology is not only the product of innovation, but it is also an input (Arrow, 1962). "Thus innovation conceivably can be a self-perpetuating process" (Grossman and Helpman, 1991, p.17). The diffusion of technology facilitates innovative activity within regions, states, or countries, hence determines productivity and the long run rate of economic growth.

Technological knowledge spillovers favour income convergence. Therefore, in an international context, the extent to which the diffusion of knowledge is international or intra-national will have important implications for whether or not global inequality is mitigated. In particular, whether developing countries catch up largely depends on the extent of adoption and implementation of new technologies that are already used in developed countries (Borensztein et al., 1998).

Empirical studies have shown that externalities from R&D activity, or knowledge spillovers, are geographically localised (Jaffe et al., 1993; Eaton and Kortum, 1999; Keller, 2002b). Although knowledge is intangible, it is not weightless in the sense it does not flow effortlessly. This is largely because a portion of knowledge is tacit (cannot be codified). The embeddedness of tacit knowledge means it must be passed through personal communication, often face-to-face connections. Consequently, an extensive literature has focused on the determinants of knowledge spillovers and what factors may facilitate the diffusion of technological knowledge; cultural similarities, language barriers, international trade, and foreign direct investment (FDI) are among the most commonly analysed factors.

International diffusion of technological knowledge is the focus of this thesis. In particular, we are interested in whether inward FDI facilitates technological knowledge to spill over into the host country. In recent times governments regularly compete to attract FDI by offering incentives such as tax breaks and import duty exceptions to multinational corporations. These policies are typically advocated under the presumption that the costs of these subsidies are small relative to spillover benefits and resulting income growth. Although economic theories suggest there are mechanisms for this to occur,[1] empirical evidence does little to justify these expensive policies designed to attract FDI. In fact, in a review of the literature, Smeets (2008) states that "empirical inconclusiveness has become so infamous that virtually every study reviewed here begins with this observation as its main motivation" (p.107). Given the disagreement in the literature, further empirical research is needed to investigate whether domestic firms benefit from the presence of larger, more productive multinational firms.

Following the work of Branstetter (2006) and Singh (2004) we test whether bilateral FDI inflows facilitate bilateral knowledge spillovers, as measured by international patent citations. Seminal work by Jaffe and co-authors has shown that patent citations provide a paper trail of knowledge flows. Citations therefore provide the researcher a valuable data source to analyse knowledge flows and potential spillovers.

We use a gravity model framework with an unbalanced panel of 58 developing countries and 32 developed OECD countries to test whether bilateral FDI has an effect on the number of bilateral patent citations. Our analysis differs from existing literature in a number of ways. Firstly, our unit of analysis is at the aggregate national level which allows us to investigate the effect of national level policies on knowledge flows through FDI. In particular, we test whether stronger intellectual property rights (IPR) play an important role in facilitating bilateral knowledge diffusion through FDI. We consider alternative FDI lag structures in our model in order to adequately model the dynamic relationship between FDI and knowledge flows. In a review of the empirical literature where productivity measures are used to study the effects of FDI on spillovers, Görg and Greenaway (2004) state that "most studies use either the contemporaneous level of foreign penetration or relatively short lags (most often one year) [...] If anything, therefore, these studies usually measure short-run effects of foreign presence."[2] Our analysis is novel in the sense it measures the medium term effect of FDI on patent citations. We achieve this by including additional FDI lagged terms, and by doing so we hope to better model the dynamic relationship between FDI and knowledge spillovers.

[1] See Smeets (2008) for a survey of the theoretical and empirical FDI spillover literature.
[2] Note: Aitken and Harrison (1999) includes 8 years of FDI lags to test 'long-term' effects on productivity.

Our findings suggest that developing countries may indeed receive positive knowledge spillovers from foreign subsidiaries; however, this effect is positive and significant only if there are relatively strong intellectual property rights, and tends to only occur with a lag. For countries with relatively weak IPRs, FDI has a negative effect on knowledge flows, even over the medium term. We interpret this finding as evidence that strong IPR regimes attract FDI that relies on more advanced technology, which is more likely to provide knowledge spillovers, whereas the types of FDI (perhaps resource seeking) received by countries with weak IPR regimes crowds out foreign investment in other sectors.

For comparison, we also analyse the effects of FDI on knowledge flows between OECD countries. In contrast to developing countries, we find that aggregate FDI has a positive impact on bilateral knowledge flows, and increasing the strength of IPR has a detrimental effect. This suggests that there may be a threshold for IPR strength in order to induce knowledge spillovers; developed countries already have good IPRs, and it may be the case that strengthening them further increases monopolistic power and dampened spillovers.

The remainder of this thesis is organised as follows. Section 2.1 briefly reviews the literature on growth theory and the importance of knowledge spillovers. In Section 2.2 we discuss measuring knowledge spillovers and review two alternative measures used in the literature. Section 2.3 reviews relevant empirical literature on estimating knowledge spillovers. Section 3 provides an overview of the patent data and discusses patent citations as indicators of knowledge flows. We also discuss why the USPTO data is suitable for our analysis and present data in more detail. In Section 4 we consider some general issues related to the estimation of gravity models. Section 5 presents our model specification, followed by a description of the complete data we use in Section 6. In Section 7 we carry out specification testing for our model. Section 8 presents our empirical results. Section 9 concludes the thesis.

2. LITERATURE REVIEW

2.1. Economic Growth Theory

Within the last few decades, the economic growth literature has evolved beyond exogenous growth theory, which assumes exogenously determined factors contributing to a country's comparative advantage and long run growth. According to the standard Heckscher-Ohlin (neo-classical) theory, all countries are assumed to have identical technology levels and production functions for any given commodity. Within this traditional framework, trade arises due to differences in relative costs of production reflecting different factor prices; and with identical production technologies, price differentials are driven by factor endowment levels of each country. Therefore the Heckscher-Ohlin theory states that countries rich in labour will export labour intensive goods and countries rich in capital will export capital intensive goods. Differences in long run growth across countries, in turn, reflect exogenous differences in productivity growth across the different sectors of specialisation.

The recent development to growth theory is that "technical change is the result of conscious economic investments and explicit decisions by many different economic units" (Griliches, 1992, p.1). Although notions of endogenous growth date back to Schumpeter (1934), the emergence of "new" growth theory in the 1980's was the major development in formalising many of the ideas of increasing returns and R&D externalities in economic models (for influential contributions see Romer, 1986; and Grossman and Helpman, 1991). In "new" growth theory models, the technological progress of economic agents, regions, or countries are no longer taken as an exogenous process of time as in the early traditional growth theory.

> *These models endogenized technological innovation, allowing the rate and direction of inventive activity to be affected by the global pattern of specialization and trade, and also allowing the pattern of trade to be, in turn, affected by the rate and direction of inventive activity, both in the global economy in aggregate and in individual countries.* (Branstetter, 1998, p.518)

The endogenous growth models brought technological differences into the core of the model; as a consequence, the rate of global economic growth depends on the rate of innovative activity.

A related question that needs to be addressed is: how do innovative efforts of individual agents determine the technological capabilities of the economy? One might suppose that when an economic agent introduces a more productive good, service, or process it will gain a comparative advantage, from which they can appropriate the economic benefits. However, the endogenous growth theories (such as Romer, 1986; and Grossman and Helpman, 1991) emphasize two aspects of technology that suggest the returns to R&D investment are not confined to the inventor. Keller (2004) describes these as:

1) Non-rival, in the sense that the marginal costs for an additional agent to use the technology are negligible, and
2) Non-excludable, at least not fully, in the sense that the return to technological investments is partly private and partly public.

The first point above notes that unlike other rival productive inputs, such as labour and physical capital, the marginal costs of an additional agent using the same technology elsewhere are negligible. The second point highlights that the non-rival nature of technology makes it very difficult for the innovator to expropriate even a fraction of its total social benefits even in the presence of intellectual property protection (Griliches, 1992).

In light of this, innovation is the process of using a combination of individual R&D input and existing general "state-of-the-art" knowledge (that is assumed to be accessible by everyone at negligible cost) to create further new designs; these innovations, in turn, contribute to the urn of general "state of the art" knowledge. It is this phenomenon that lays the foundations for endogenous growth theory; R&D provides positive externalities, and these externalities serve as the "engines" of endogenous growth that allow sustained global economic growth without diminishing returns setting in (Branstetter, 1998).

Although "new" growth theory lays great emphasis on the significance of knowledge spillovers, until recently there has been very little empirical research on them. Subsequently, the literature has set about answering various important questions:

How can knowledge spillovers be measured? Do nearby firms benefit more from local R&D? Or does knowledge flow across economic regions, or country borders? What facilitates knowledge to spillover to other economic agents? Do international trade, labour mobility, and foreign direct investment play influential roles in international knowledge diffusion?

In discussing the literature, we begin by discussing the two most widely used methods to measure knowledge spillovers in Section 2.2, and then in Section 2.3 we provide an overview of some of these questions and findings supplied.

2.2. Measuring Knowledge Spillovers

> "*Knowledge spillover occurs when firm A is able to derive economic benefit from R&D activity undertaken by firm B without sharing in the cost firm B incurred in undertaking its R&D.*" (Branstetter, 1998, p.521)

Knowledge is an intangible; therefore it is difficult to trace spillovers of knowledge from one firm to another. One reason is that only a fraction of knowledge is codifiable because it is impossible or very expensive to document such a large and complicated mass of information. Therefore a large portion of knowledge is known as tacit, or more easily understood as knowledge that is learnt through experience and relationships. As noted by Keller (2004), this means that complete contracts for valuable knowledge cannot be written. Therefore it is not surprising that there is no direct measure of the positive externalities created from R&D activity. To address the void of data the researcher must utilise indirect measures to capture the extent of these knowledge spillovers. Two widely used approaches are to measure a) the impact of knowledge spillovers by the effect on productivity, or b) the resulting output of knowledge spillover (patents).

2.2.1. Productivity Measures

Many papers, most notably Coe and Helpman (1995) and Keller (2002a, 2002b) have studied the externalities from R&D by assessing the impact of R&D on productivity. It is commonly known in economics that after subtracting the contribution of inputs such as labour and capital from output, the remaining variation in output is due to the unobservable factor usually labelled the 'level of technology'. Therefore, if the productivity of a firm increases as the R&D efforts of another firm increase, we can implicitly say there has been knowledge spillover. Using this production function framework, Coe and Helpman (1995) derive total factor productivity (TFP) measures for a sample of 21 OECD countries. Using these TFP measures, they estimate the effects of domestic and foreign R&D capital stocks on TFP. They find evidence that there is indeed a link between productivity and R&D. Both domestic R&D, and R&D of trading partners contribute to productivity increases. This is consistent with the notion of knowledge spillovers, in the sense that if the R&D of country j is correlated with TFP of country i, all else equal, country i must be utilising the contribution to the effective stock of knowledge from country j.

In a similar study carried out at the industry level, Keller (2002a) finds that productivity of industry i is correlated with the R&D efforts of industry j, where industry j is either domestic or foreign. Keller estimates that the own-industry R&D accounts for about 50% of productivity, while domestic inter-industry and foreign knowledge spillovers account for 30% and 20%, respectively. Again, using a productivity measure to determine the extent of knowledge spillover, Keller (2002b) studies the international diffusion of knowledge. This study focuses on the impact of distance on the magnitude of productivity gains. The author finds that distance has a negative effect on the gains from foreign R&D. (The role of geographic distance on knowledge flows will be discussed in more detail in Section 2.3.1.)

One problem with using productivity gains as a measure of knowledge spillovers is that TFP is a derived measure which is subject to measurement error and possible bias. If we consider the simple Cobb-Douglas production function:

$$Y = AL^{\alpha}K^{1-\alpha}$$

TFP, as denoted by A, is derived using data on factor inputs labour and capital (L and K), as well as outputs (Y); all of which come with considerable measurement error because the required data rarely comes in the desired form (Keller, 2009). Griliches (1992) argues that often the literature confuses R&D spillovers with results of measurement error. The example he provides is when "R&D intensive inputs are purchased from other industries at less than their full 'quality' price" (p.36). The resulting increase in productivity is due to the measurement of capital and not pure knowledge spillovers. Establishing a causal effect is particularly difficult under these circumstances. Furthermore, changes in productivity levels are subject to both market conditions and technological progress and it is often impossible to disentangle the two effects. We discuss this issue further in Section 2.3.2.

2.2.2. Patent Data Measures

Until recently, generating derived productivity measures of knowledge diffusion was necessary because the nature of knowledge flows means they are very difficult to trace. Or as Krugman (1992) put it, in relation to labour and capital "knowledge flows, by contrast, are invisible" (p.53). So although knowledge flows were known to play an integral part in economic growth it was impossible to measure them directly; therefore the magnitude of the effect of knowledge spillovers was largely unknown.

However, Jaffe et al. (1993) were the first to suggest using patent data for this purpose. As is often quoted in the literature, they state that "knowledge flows do sometimes leave a paper trail, in the form of patent citations" (p.578). That is, when a patent is applied for, the applicant is obligated by law to reference all prior art of which their invention builds upon (henceforth referred to as citations). These citations, at least in principle, capture the stock of knowledge the inventor utilised in the process of developing a new product. Consequently, subject to caveats (which we discuss in detail in Section 3), the researcher is able to use patent citations as a measure of knowledge flows from the cited patent to the citing patent. We elaborate on how patent data has been used in applied research in the following section. In this thesis we also make use of this rich source of information to study bilateral knowledge spillovers.

2.3. Empirical Studies of Knowledge Flows

The importance of knowledge accumulation, spillovers and increasing returns has been stressed in the growth literature for some time. However, it was not until more recently that this literature has focused on estimating knowledge spillovers explicitly. What follows is a review of this literature.

2.3.1. Geographic Distance

A question that researchers have paid particular interest to is: where do spillovers go? The seminal paper by Jaffe et al. (1993) tests whether knowledge spillovers in the U.S. are localised within U.S. states:

> *"Is there any advantage to nearby firms, or even firms in the same country, or do spillovers waft into the ether, available for anyone around the globe to grab?"* (p.577)

As previously stated, the authors pioneered the use of patent citations to study knowledge flows. They use citations received by universities and firms in the U.S. to test whether inventors from citing patents are disproportionally geographically matched to the cited inventor, compared to a control group of non-citing patents that have similar temporal and technological distributions. They find that inventors in the U.S. are more likely to cite innovations from within their own state than they are from other states. Jaffe et al. (1993) conclude that this is evidence that the knowledge trails left by patent citations are geographically localised.

Peri (2002) also uses patent data to analyse knowledge flows at the sub-national region level in Europe and North America. The author estimates a gravity type equation of the diffusion knowledge across regions; the model includes a measure of geographical distance between regions. He finds that regions farther away from one another exhibit much lower flows of knowledge than relatively close regions. Another study, Thompson (2006) also uses patent citation data to assess the localisation of knowledge spillovers. The study provides further evidence that knowledge is geographically localised intra-nationally; He also shows international borders have a negative effect on the diffusion of knowledge.

Specifically looking at international knowledge diffusion, Eaton and Kortum (1999) use country-level patenting activity to estimate models of technology diffusion. They exploit the international patenting system to infer "who is getting what from whom." Their assumption is that an inventor will choose to seek patent protection in countries where their idea is likely to be used, hence where the knowledge will flow to. Eaton and Kortum (1999) find evidence that there are geographic barriers to the spread of technological knowledge; "research performed abroad is about two-thirds as potent as domestic research" (p.537). This is in line with Jaffe and Trajtenberg (1999), who find that patents whose inventors reside in the same country are typically 30% to 80% more likely to cite each other than inventors from other countries. For further evidence that knowledge is geographically localised using patent data see Branstetter (2001); Maurseth and Verspagen (2002); Bottazzi and Peri (2003); Peri (2005); MacGarvie (2005); and Li (2009).

Using an alternative approach, Keller (2002b) examines whether the distance between countries affect the magnitude of productivity gains from each others' R&D efforts. Their hypothesis is this: if knowledge is international in scope (i.e. there exists a global pool of knowledge), distance should not matter for international diffusion of the knowledge generated from R&D spending. The author relates the R&D spending of the G-5 countries[3] to the productivity levels of nine OECD countries. He tests whether geographic distance affects the productivity gains from G-5 countries' R&D. The paper finds that productivity effects decline with geographic distance; for every additional 1,200km there is a 50% drop in technological diffusion.

The consistency in these findings suggests that knowledge is not weightless and does not diffuse frictionlessly. This is at first compelling, given that knowledge is an intangible that seemingly should be unaffected by geographic distance. The distance effect on trade levels found in the literature is often suggested to be partly explained by transport costs; the same intuition does not seem to directly relate in the same way to knowledge diffusion. As explained earlier, knowledge is, however, made up of a large proportion of non-codifiable (tacit) information. The transferability of such knowledge is limited by it embeddedness in individuals, teams and organisations, and therefore detailed communication (likely to be face-to-face contact) is often required in order to pass on the knowledge (Montobbio and Sterzi, 2012). From this standpoint, the ability of distance to inhibit knowledge diffusion is more apparent. The amount of knowledge diffusion declines with distance because in equilibrium it is relatively more costly to transfer knowledge to remote places, so there is less of it (Keller, 2009). Therefore one reason why technically advanced networks form (eg. Silicon Valley) is because firms intentionally exploit knowledge flows that are geographically localised (Döring and Schnellenbach, 2006). However, it is also intuitive to assume that the decrease in communication cost through better technologies (video conferencing, etc.) would mean

[3] France, Germany, Japan, United Kingdom, and the United States.

the distance effect may have diminished over time. The empirical evidence is mixed on this question.

Knowledge is not strictly localised, as we can see from day-to-day life. All kinds of consumer goods have a vast amount of technology embedded in them that originate from all around the globe. So what enables technological knowledge to travel abroad? There has been a great deal of interest in the literature studying the determinants of international knowledge diffusion. Amongst the most frequently mentioned channels are international trade and FDI; I discuss some of the key contributions to the literature on these two channels.

2.3.2. Trade and International Knowledge Diffusion

The most influential attempt to estimate the effect of international knowledge spillovers from R&D through trade is Coe and Helpman (1995). They estimate the TFP residuals for a set of countries using the aggregate measures of factor inputs; capital, and labour. As discussed in Section 2.2.1, TFP is a measure of productivity due to a country's technological level. Coe and Helpman (1995) regress the TFP residuals on aggregate R&D and weighted foreign R&D, weighted by bilateral trade between the countries. They find evidence that there is a close link between productivity and both domestic and foreign R&D capital stocks. In other words, a country's TFP depends not only on its own research efforts, but on the 'imported' R&D that it receives through its trading partners. Moreover, the authors state that the benefit from foreign R&D is increasing in "openness" to international trade with R&D rich countries. This seminal paper brought the first evidence of international trade facilitating knowledge spillovers at the aggregate level.

However, others have questioned whether this approach is actually capturing a knowledge spillover. Branstetter (2001) suggests we may instead be observing common demand or input price shocks, or a common time trend. This is because at an aggregate level the model fails to control for considerable technological heterogeneity. It seems likely that R&D in the pharmaceutical industry will not produce positive externalities to the mechanical sector. Therefore it is difficult to interpret a casual relationship when one finds positive and statistically significant results as did Coe and Helpman (1995). Keller (1998) sheds further doubt on using aggregate R&D measures to infer knowledge spillovers through international trade. Analysing the results of Coe and Helpman (1995), Keller (1998) uses simulation methods to show that weighting foreign R&D by randomly assigned import shares yields similarly high and positive effects of foreign R&D on domestic productivity. In other words, randomly assigned trade patterns lead to similar "trade-related" R&D spillovers as found using "true" trade patterns. This paper does not argue that trade does not facilitate knowledge flows. Instead it casts doubt on whether trade-related knowledge spillovers can be analysed in this way, using aggregate R&D measures. In response, Coe et al. (2009) revisit their approach with modern econometric techniques and an updated data set. They confirm that there is robust evidence that domestic R&D capital, and foreign R&D are significant determinants of TFP.[4] Zhu and Jeon (2007) follow a similar approach and find that the elasticity of TFP with respect to bilateral trade is around 2%.

Sjöholm (1996) provides an alternative method for identifying whether trade encourages knowledge to flow across national borders. Using the share of patent citations in Swedish patent applications to other countries as a proxy for the share of knowledge to flow into Sweden, Sjöholm finds that total trade between Sweden and country *i* has a positive effect on the number of patent citations to country *i*. MacGarvie (2006) also find evidence that inventors from firms that import are more likely to cite foreign patents, all else equal. This is interpreted as importers are more likely to be influenced by foreign technology than non importing firms.

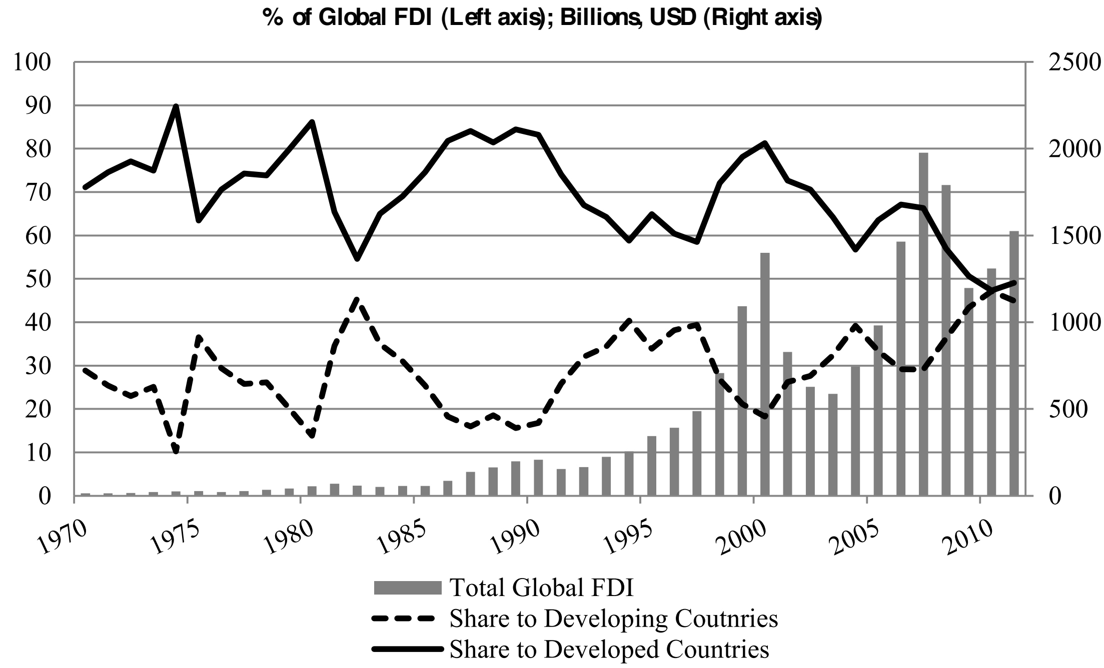
The theory on how trade may facilitate knowledge to spill over national borders has also been widely considered (see Keller (2009), for example). The main idea stemming from this theory is that technological knowledge is embodied in tradable goods and in response to economic opportunities, firms imitate, reverse-engineer, or use these goods in the production process which in turn become the new "state of the art." If a country is open to international trade, not only will they have access to a wider variety of goods but also better quality goods of which domestic producers can use to increase productivity through best practice and further innovation.

[4] Also, Coe and Hoffmaister (1999) show that Keller's weights were not random and if they were the results would be extremely small. See Coe et al. (2009), footnote 19 for a precise explanation.

2.3.3. Foreign Direct Investment and International Knowledge Diffusion

Another channel that may mediate knowledge to flow between countries is through FDI. In recent times countries have created policies that open their shores to multinational corporations. Figure 2.1 shows how rapidly FDI has increased, as well as the increasing relevance of FDI flows into developing countries. Görg and Greenaway (2004) states that FDI is probably the most visible driver of globalisation, and it is growing at least twice as fast as international trade.

Figure 2.1 Global FDI inflows 1970 - 2011



Source: UNCTAD

Often policies are deliberately designed to attract multinational firms in technologically advanced industries.[5] There seems to be a widely held assumption among policy makers that FDI will not only increase capital stock, but will also provide positive spillovers to domestic firms.

The early influential empirical studies on the spillovers from FDI largely concentrated on the effect FDI had on the productivity of domestic manufacturing firms. Haddad and Harrison (1993) use a production function approach to test whether the

[5] For example, in 1994 the U.S. state of Alabama spent $ 230 million, or $ 150,000 per newly created job, to attract a new plant of Mercedes-Benz.

presence of foreign multinationals increases the productivity of domestic firms in Morocco. They find no significant evidence of such a relationship; however, they note that there is lower dispersion of productivity levels among sectors with more FDI. A study of the Venezuelan manufacturing sector by Aitken and Harrison (1999) finds two opposing effects of FDI on domestically owned plants. Plants with higher FDI receive positive productivity gains, and plants with no FDI exhibit productivity declines as FDI increases in Venezuela. They conclude that the net effect on the economy is small.

Using productivity measures to infer knowledge spillovers from FDI is questionable for a number of reasons. It is possible that any observed positive effect on productivity is subject to self selection bias; this endogeneity of FDI is often not accounted for, as many studies use cross sectional data. When choosing where to locate foreign subsidiaries, multinationals will choose countries with the best opportunity for economic gain; hence it is likely that FDI is attracted to more productive countries. In this case, the causal relationship is reversed. Furthermore, multinational firms are often larger and more productive than their domestically owned counterparts. If studies concentrate on aggregate country or industry productivity, the presence of FDI will increase local productivity regardless of any technology gains (hence knowledge spillovers) to domestic firms.

On the other hand, if the effect of FDI on domestic firms' productivity is negative (for example, Aitken and Harrison, 1999), it may be a result of higher competition introduced by the FDI. So although there may be knowledge spillovers from multinationals to domestic firms, productivity measures will be negative if multinationals attract demand away from the domestic firms, forcing them to reduce production and move up their average cost curve.[6] Because this line of literature fails to separate out these alternative sources of productivity gains (losses) and do not explicitly model knowledge spillovers, other authors have focused on more direct measures of knowledge flows using information from patents.

Almeida (1996) studies the knowledge flows of multinational firms in the semi-conductor industry in the U.S. They ask whether foreign subsidiaries in the U.S. contribute to local knowledge to a greater extent that would be expected of similar domestic firms. Using patent citations, they find that foreign subsidiaries' patents indeed are cited more frequently from local domestic firms than would be otherwise expected.

[6] See Aitken and Harrison (1999) for a more detailed explanation.

Branstetter (2006) also uses patent data to directly measure the impact of FDI on knowledge spillovers. The author is able to link firm-level data on Japanese manufacturing firms with foreign investment in the U.S. to citation patterns in their U.S. patents. The panel data set allows the author to study the effects of changes in "FDI presence" on facilitating knowledge spillovers between the two countries, as indicated by patent citations. He finds FDI does provide a channel for knowledge to flow across national borders. Furthermore, the spillovers from investing Japanese firms to domestic U.S. firms is strongest amongst greenfield affiliates, which usually possess superior technology.

Singh (2004) analyses data from patent applications to identify whether the assignee firm of the patent is a foreign subsidiary. The author then measures knowledge flows between foreign subsidiaries and domestic firms using patent citations. Using a regression analysis to model the probability of a citation between two patents, Singh finds that domestic firms are more likely to cite patents of foreign subsidiaries than a reference group representing cross-border inter-organisation citations. Furthermore, the paper shows that multinational firms are particularly good at relaying knowledge between their subsidiaries and home base. This provides evidence that multinational firms facilitate knowledge flows across national borders.

Without access to micro-level data similar to what is used in these papers, this thesis concentrates on estimating knowledge flows at the aggregate level. We ask a similar question: does FDI provide knowledge spillovers to the domestic firms? However, by looking at an aggregate national level we are able to concentrate on national policies related to the transmission of knowledge through FDI, such as the strength of a country's IPR regime.

A natural choice of empirical framework for investigating factor flows between countries is the gravity model, which has been extensively used in the trade literature with a great deal of success. (We introduce the basic gravity model and discuss some estimation issues in Section 4.) Maurseth and Verspagen (2002); MacGarvie (2005); Li (2009); Picci (2010); and Montobbio and Sterzi (2012) are examples of studies that extend the gravity model to bilateral knowledge flows, in particular using patent data.

Picci (2010) uses information on the country of residence of inventor(s) and applicant in patent application data from more than 80 patent offices. From this information he defines a patent as "international" if at least one inventor or applicant resides in a different country to the others. A gravity model is applied to assess various determinants of the intensity of collaboration between countries. The idea here is that international collaboration is a direct measure of sharing knowledge. This paper confirms that cultural and technological similarities positively affect bilateral collaboration, whereas the study fails to find an unambiguous effect of FDI. The author uses aggregate FDI flows in the sense that they do not distinguish the source country (bilateral FDI), which the author suggests is an area for further research.

Similarly, Montobbio and Sterzi (2012) use the idea that the collaboration among inventors represents technological knowledge flow between the inventors. They consider patents with inventors residing in different countries to analyse the determinants of international technological collaborations. Using a gravity model, they estimate the impact of geographic and technological distance between the two countries and the strength of IPR (among other control variables) on international knowledge flows. Geographic distance is found to have an insignificant effect once technological and cultural distances are controlled for. Technological proximity, which is a measure of how similar two countries' technological activities are, is a particularly important factor in explaining international collaboration between inventors. This study provides evidence that the effect of increasing IPRs on knowledge flows depends on the initial strength of IPR. There seems to be an inverse "U" shape effect, where improving IPR increases international knowledge flows from countries affording weak IPR and the opposite effect for countries with strong IPR. The reasons could be many; the authors suggest that stronger IPRs should increase economic openness via FDI, imports and joint ventures. Montobbio and Sterzi (2012) indeed show that an increase in IPR does facilitate international technological collaborations (hence knowledge flows) if the two countries also have an increasing trade relationship. The similar effect in relation to FDI is not explored; this is a question of particular interest that we ask in this thesis.

The previous two papers we have discussed use the collaboration between inventors as indicators of knowledge flow; however this does not allow the researcher to directly determine the direction of knowledge flow. As we have discussed, patent citations do provide such information. The following papers exploit patent citation data and assess international knowledge flows at an aggregate level with a gravity type model.

MacGarvie (2005) uses a panel data set of all patent citations of USPTO patents from ten developed countries to study the determinants of bilateral patent citations. The paper provides evidence that the usual "gravity terms" have the expected effect on the number of patent citations between countries. The inventive "mass" of each country, as measured by aggregate patent counts, has a positive effect while the distance between them decreases the number of patent citations. The paper also provides some evidence that bilateral FDI flows between the two countries has a positive effect on number of cross-country citations.[7] Li (2009) also employs a gravity model framework to study the factors affecting knowledge diffusion between countries, as measured by patent citations. The paper specifically concentrates on the effects of distance and national borders, and finds that distance has a negative effect on the number of patent citations between two regions.[8] Maurseth and Verspagen (2002) applies a similar framework to study knowledge flows between European regions. They test the effects of geographic distance, national borders, technological specialisation, and language on the number of patent citations between regions.

Although MacGarvie (2005) included FDI as a control variable, none of these aggregate-level studies using a gravity framework focus on the effect bilateral FDI has on knowledge spillovers (as indicated by patent citations). That is the primary focus of this thesis.

[7] Although given the journal it was published in, the paper is very brief. It gives little explanation on the type of FDI data used, and also deals with missing FDI data in a questionable manner.
[8] The paper also provides a thorough derivation of the gravity equation of knowledge spillovers.

3. PATENT DATA OVERVIEW

As we have mentioned, we utilise the unique wealth of information that is captured in patents.

> *A patent confers, by law, a set of exclusive rights to applicants for inventions that meet the standards of novelty, non-obviousness and industrial applicability. It is valid for a limited period of time (generally 20 years), during which patent holders can commercially exploit their inventions on an exclusive basis. In return, applicants are obliged to disclose their inventions to the public so that others, skilled in the art, may replicate them. The patent system is designed to encourage innovation by providing innovators with time-limited exclusive legal rights, thus enabling them to appropriate the returns of their innovative activity.* (WIPO, 2011, p.35)

Patent applications provide detailed information on the invention, inventor(s), and the assignee who owns the rights to the invention. Furthermore each patent applicant is legally bound to make reference (citations) to all previous patents as well as other forms of intellectual property that represent "prior art" on which the applicant's invention builds on.

3.1. Patent Citations

Patent citations provide researchers with some indication of the timing and the direction of knowledge flows. The legal purpose of patent citations is to indicate which parts of the knowledge are claimed in the application and which parts have been claimed by previous patents or non-patent form of intellectual property (Criscuolo and Verspagen, 2008). To familiarise the concept, these citations somewhat resemble references in academic papers; however, there are distinct differences. In academic research, references can be added to build motivation even though the work does not necessarily build upon the cited paper (as well as other strategic reasons)[9]; by contrast, inventors

[9] Criscuolo and Verspagen (2008) give an example where the author of the cited paper may be a potential reviewer.

have the incentive to cite less prior art as it directly limits the scope of their claim.[10] The extent to which the inventor can ‘get away’ with withholding citations is ultimately determined by the patent examiner. Patent examiners are considered to be experts in the field and are responsible for judging the degree of novelty of the patent. After the applicants disclose their citations, patent examiners conduct their own prior art searches. They may challenge the applicant’s claims, and can add additional citations to the patent application. Examiner added citations are widely expressed as a shortcoming of using patent citation data to analyse knowledge flows and spillovers. The common point made is that citation counts may not accurately reflect the unobservable knowledge flow because citations added by the examiner may introduce noise if the inventor(s) was unaware of the technology underlying the cited patent(s).[11] This concern is especially relevant in the case of USPTO because examiners add approximately two thirds of all citations.[12] To justify the use of citations as indicators of knowledge spillovers, most papers in the literature make reference to Jaffe et al. (2000). This paper provides survey evidence that cited patents are more likely to be sourced by the inventor than similar non-cited patents. Interestingly, the authors find that one-third of citing inventors were unaware of the cited patent. Jaffe et al (2000) conclude that patent citations can be used as a signal of the presence of knowledge spillovers, albeit a noisy one. Duguet and MacGarvie (2005) also provide evidence from the European Patent Office (EPO) and survey data on French firms that citation counts contain relevant information of technology flows.

Contrary to the view that citations added by the examiner do not represent a knowledge flow (see Jaffe et al., 1993). Breschi and Lissoni (2005) suggest “there is no reason to exclude that examiner’s citations (i.e. ‘unaware citations’) may signal knowledge flow [...] At most we can presume that citing and cited inventors do not know one another” (p.623). However the authors suggest that knowledge may well flow through a common acquaintance, or a social chain of personal relationships. Therefore they conclude examiner added citations need not cause noise “as long as one recognises that

---

[10] Keeping in mind inventors are legally bound to disclose known prior art. Failure to do so can result in fraud charges. However they are not required to search prior art. For an excellent explanation of the incentives and disincentives for inventors to include citations see Alcácer et al. (2009)

[11] See Breschi and Lissoni (2005) for a well constructed discussion on three relationship scenarios that could exist between citing and cited patent.

[12] The USPTO average for 2001 – 2003. The distinction between inventor and examiner citations at USPTO was not possible until 2001 (see Alcácer et al., 2009). In contrast, Maurseth and Verspagen (2002) states that in the U.S., inventors add the majority of citations even though it is the examiners who finally determine which citations to include.

knowledge of technical contents of a patent may travel independently from information about the existence of that patent, or from exact references to the relevant documents." (Breschi and Lissoni, 2005, p.641).

Another issue with the data is that citations to patents owned by the same assignee as the citing patent (so called self-citations) do not, in general, represent knowledge spillover. Presumably, these citations represent knowledge that is internalised; hence it is common practice to eliminate them when studying externalities. The common approach would exclude citations with the same assignee code. However, this is far from satisfactory given the data available. Even if one is able to exclude citations between inventors belonging to the same firm, it is almost impossible to control completely for self-citations between subsidiaries, affiliates, and firms resulting from mergers and acquisitions throughout history. It would take an incredibly laborious manual task of checking whether the assignee names are sufficiently similar or co-owned, etc.[13] Judgement is then needed to decide on the degree of interaction between the related firms. As Thompson (2006) points out: "This does not seem to be a criterion that leads itself to measurement" (p.388). Self-citations have been measured to be around 11% of total citations (Hall et al., 2001). However, given that we are only considering bilateral citations in this study we expect self-citations to make up much less of total bilateral citations in our sample. Admittedly, self-citations do add noise to the analysis of knowledge spillovers. Nevertheless, we follow the prevalent understanding that studying bilateral citations including self-citations will still represent benefits to the receiving country even if they are partly explained by the transfer of knowledge within firms.

In line with the literature, we acknowledge the limitations of patent data and proceed with the assumption that citations allow us to identify knowledge flows between inventors and citation counts provide a proxy for the extent of knowledge spillovers (especially in large samples).

3.2. Alternative Patent Offices

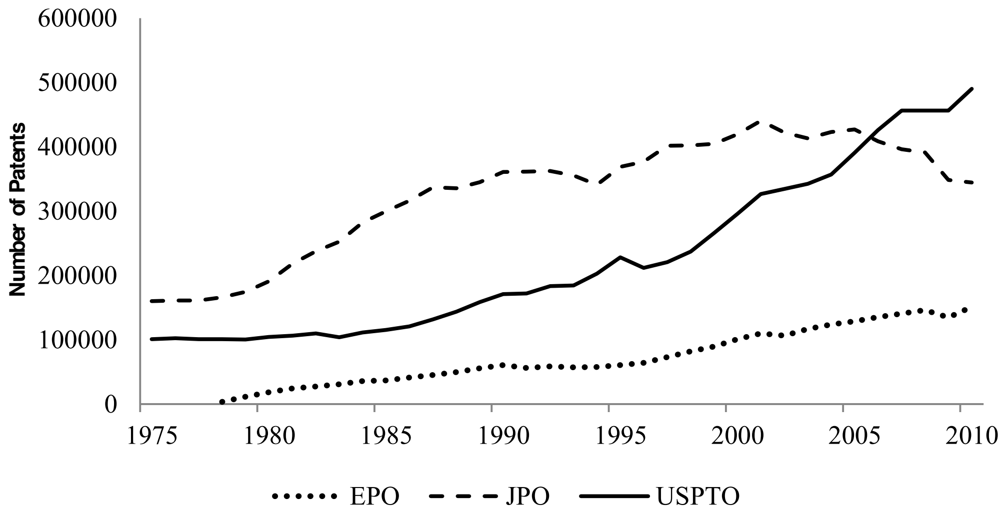
In the literature on knowledge spillovers, data from any of the three major patent offices in the world are generally used; the USPTO, European Patent Office (EPO), and the

[13] Researchers do attempt to do so using databases such as "who owns whom". Given the time available for this thesis, it is beyond the scope of this study.

Japanese Patent Office (JPO). When the study concentrates on regional or state level analysis the national patent office is justifiably used.[14] Similarly, when one is studying the aspects of inventive activity between a chosen pair of countries it makes sense to use the corresponding patent office(s).[15] However, in a study with a worldwide scope, it is not feasible to simply amalgamate the data from individual patent offices. This is because there are overlaps of the same invention, and due to differences in the processing and publishing of patent applications.[16] For example the EPO data lists all patent applications whereas the available USPTO data only contains patents that have been granted. The two patent offices also have different processes for documenting patent citations.

For a global study we would ideally have information on every patent granted worldwide for each unique invention.[17] In the absence of such data we aim to obtain the most comprehensive coverage of all patent applications as possible using any individual patent office. Figure 3.1 below shows the number of patent applications filed through each of the three major patent offices.

Figure 3.1 Patent applications at the top three patent offices



Source: WIPO

---

[14] For example Jaffe et al. (1993)

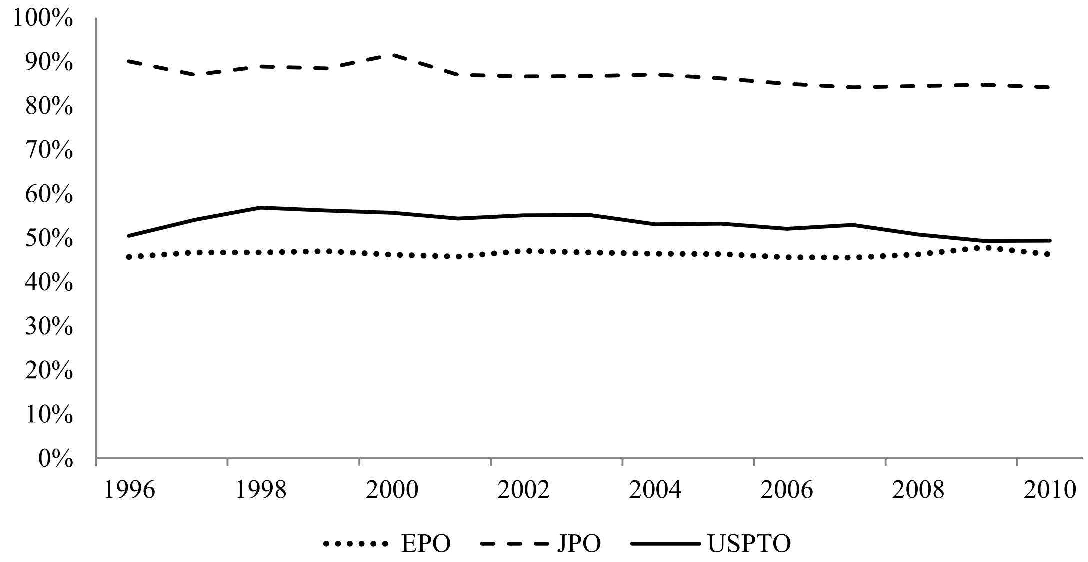
[15] See Branstetter (2006) for a study between U.S. and Japan.

[16] A "Worldwide Patent Statistical Database" known as Patstat has been created by a Patent Statistics Task Force but is in initial stages. It requires a subscription fee which prevented us from using it in this thesis.

[17] The Patstat database is a step in this direction; see Picci (2010). Similarly, the OECD triadic patent family dataset also accounts for patent families; see Dernis and Khan (2004).

It is clear from Figure 3.1 that JPO and USPTO have been the most heavily used patent offices in the world. In fact the two offices have received between 30% and 40% of total world patent applications every year for the last two decades. This suggests either the U.S. or Japanese office will provide the best coverage of worldwide patent data over longer time horizons. However, when we consider where the inventions come from within each patent office, some interesting observations arise. Figure 3.2 shows the proportion of patents applied at each office that are residents of the country the office is in.[18]

Figure 3.2 Proportion of applications from residents



Source: WIPO

Figure 3.2 reveals that the vast majority (over 80%) of patent applications to the JPO are from residents of Japan. In contrast, less than 50% of USPTO applications are from U.S. residents. This suggests that the JPO data do not offer a good representation of patents applied for from around the world. The fact that firms tend to apply for patent protection in their home country is known in the literature as the 'home advantage' effect. This has the potential to bias results if the data are used to compare the inventive activities of countries or regions (Criscuolo (2006).[19]

[18] Because the EPO is a regional office it does not strictly speaking have a home country. For the purpose of illustration I assign any resident of an EPO member country as an EPO 'resident'.

[19] In the present analysis, however, we are able to mitigate this potential bias using country or country-pair fixed effects.

To further understand where the inventions of patent applications come from Figure 3.3 shows the share of applicants that come from four of the highest patenting counties: USA, Japan, Germany, and France.

Figure 3.3 Share of patent applications in 2005

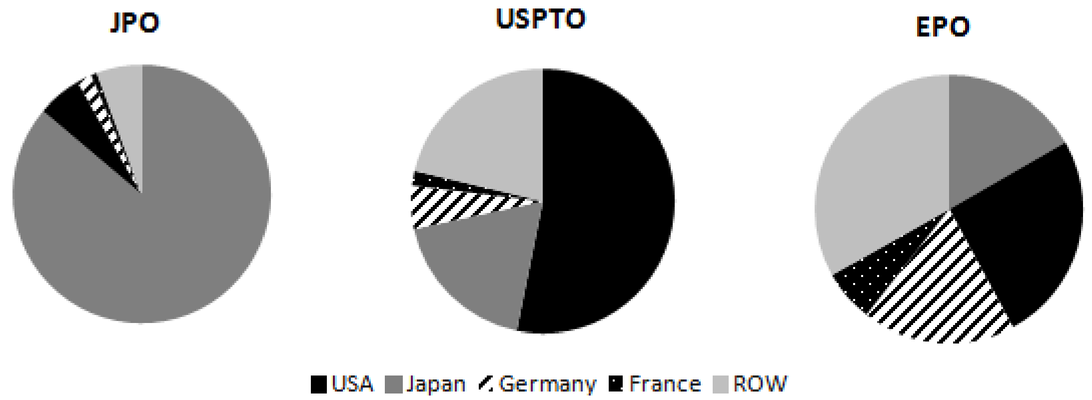


Figure 3.3 confirms that a very small proportion of JPO applications come from outside of Japan, with the remainder largely taken up by the U.S. and the rest of the world (ROW). From the centre diagram we can see that over a quarter of USPTO applications do not come from the U.S. or Japan. This provides further evidence that patent applications to the USPTO come from a mixture of countries and therefore is likely to provide the most suitable data to analyse worldwide knowledge flows. The right-hand-side diagram in Figure 3.3 shows that the EPO is much more evenly spread between the major countries and that the ROW has the largest proportion. This may indicate that the EPO has the most diverse mix of applications; hence one might suggest that the EPO could provide a better source of information for the purpose of this study. Though this may be true, we would like to highlight a few points that could suggest otherwise.

Firstly, EPO contains only 8 percent of total world patents, compared to 25 percent for the USPTO.[20] And although EPO applications are more evenly spread between the major country applicants, there are far fewer EPO applications from countries outside the OECD compared to the USPTO (6,000 and 28,000 respectively).[21] As we are specifically interested in analysing knowledge flows from OECD to

[20] As at 2010, according to WIPO data.
[21] In 2010, according to WIPO data.

developing countries, we think it is important we capture as much of the non-OECD patent applications as possible.

Secondly, it is important to note that patent applications around the world are not necessarily for unique inventions. Often the owner of an invention will seek patent rights through multiple patent offices (this is referred to as a patent family). In essence a firm will seek patents rights in an existing or potential market for their invention, and obviously there can be more than one. So although the EPO applications are more evenly distributed among nations, it is entirely possible that a large portion of EPO applicants also apply at the USPTO, even though their overall share of total USPTO patents is swamped by the large number of applications from the U.S. and Japan. Furthermore, Criscuolo (2006) point out that non-U.S. firms have an interest in protecting their most significant innovations in the USPTO because the U.S. is the largest market for technologically advanced products.[22]

Lastly, a general weakness of patent count data is that they do not accurately proxy for the value of innovative output, as patents vary considerably in their technological and economic value (see Trajtenberg, 1990). By using USPTO data we are more likely to capture the majority of valuable inventions from around the world. This is important when considering knowledge spillovers because we are most interested in economically valuable knowledge spillovers that are the driving force of endogenous growth.

3.3. USPTO Patent Data

A long term research and data creation effort from various NBER researchers, mainly B. Hall, A. Jaffe, M. Trajtenberg, M. Fogarty, and R. Henderson, has made a very detailed data set of USPTO patent application information available to the public. The data covers all patents granted (emphasis) from 1975.[23] To the best of our knowledge all literature using USPTO data uses the NBER patent data set, either the first edition that covers 1975 – 1999, or the version updated to 2006 (for a detailed description see Hall et al., 2001).[24] In this thesis we make use of a newly available data set containing the same

[22] Branstetter (2006) also note that Japanese firms seek to patent all their valuable ideas in both the U.S. and Japan, so that trends in their U.S. patents should be reflective of their total innovative activity.
[23] This is the earliest USPTO application information with citations available in digital form.
[24] The first edition was made available in Jaffe and Trajtenberg (2002) or the NBER website. http://data.nber.org/patents/ The updated version is on the official project website. https://sites.google.com/site/patentdataproject/Home

raw USPTO information and updated to 2010. This data was constructed by Lai et al. (2011) and made publically available online.[25] This updated data set has a total of 3.9 million granted utility patents, over 1 million more than the commonly used first version of the data.[26] Since 2006 there have been 297,000 patents granted at the USPTO. Figure 3.4 shows annual number of patents granted by application year and Figure 3.5 shows the average number of patents granted by application year per country within each group.

Figure 3.4 Number of patents granted by application year

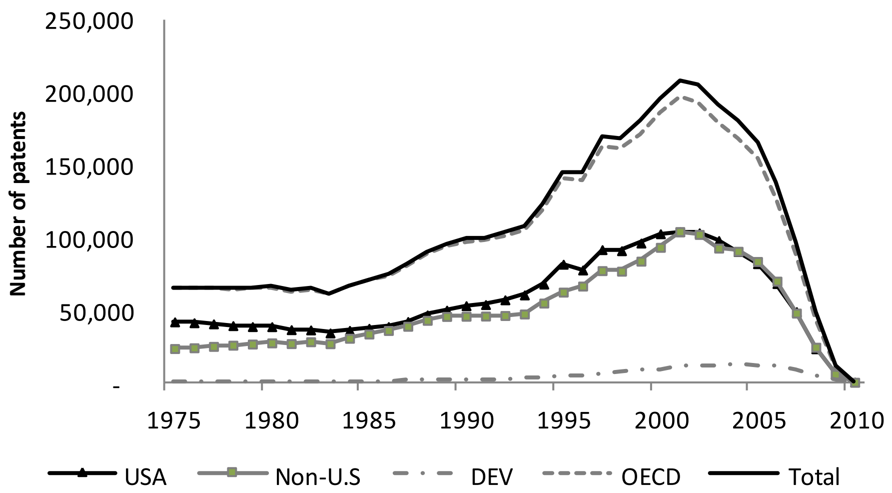
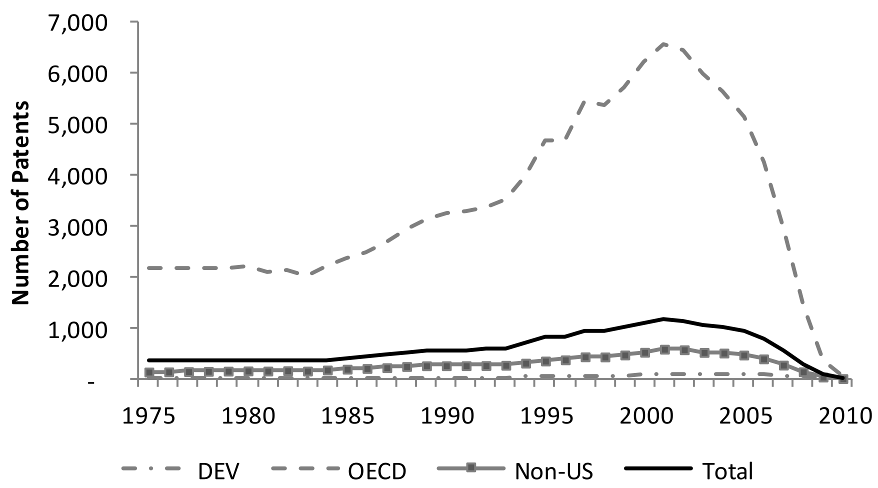


Figure 3.5 Average number of patents granted per country by application year



---

[25] On the Harvard dataverse website. http://dvn.iq.harvard.edu/dvn/dv/patent

[26] Following Hall et al. (2001) we only consider utility patents. There are three other major types of patent categories: design, reissue, and plant patents. The vast majority patents are utility patents.

The dramatic decline in recent years highlights the extent of truncation in the data. Because the data only includes patents that have been granted, there will be an increasingly large portion of patents that have been applied for but not yet granted. Therefore in the years close to 2010 we are only observing the patents that were granted relatively quickly, and not the ones that are yet to be granted (but were applied for in 2009, for example).[27]

Figure 3.5 shows that the average number of patents granted per year by an OECD country is much greater than the average of all countries. It also shows that when the U.S. is excluded the average annual applications significantly drops, indicating that the U.S. patents a large number of innovations relative to all other countries.[28] Also the very small, often zero, average number of patents granted to developing countries is an important observation shown in Figure 3.5. This reflects the large number of developing countries that do not file any patents through the USPTO each year. In other words, the distribution of the number of patents granted to a developing country is heavily skewed to the right (see Table A 1.1 in Appendix A1 for further descriptive statistics on the number of patents granted). The large number of zero values is an important attribute of the data and will be discussed in detail in Section 4.

For each of the 3.9 million patents in our data set there is detailed information of the invention itself and its inventors (e.g. their geographical location). The data also includes all citations to previous patents. "These citations open up the possibility of tracing multiple linkages between inventions, inventors, (and) locations" (Hall et al., 2001, p.4). An example of a patent document is included in Appendix A1. The information we use from the patent data set is described in Table 3.1 below.

Table 3.1 Patent data set information

Variable	Description
$Patent_i$	8 character alphanumeric identification assigned by the USPTO
$AppYear_i$	Year of application by defined patent
$Country_i$	Country of residence of inventor
$Class_i$	Primary Patent classification (Up to 10)
$InvSeq_i$	Patent Inventor sequence (0 = primary Inventor)
$Citation_j$	Patent number of the patent cited by defined patent

[27] See Hall et al. (2001) for more descriptive analysis on the data.
[28] At least at the USPTO.

We follow common practice in the literature to take the country of residence of the inventor(s) as the place where the innovation took place. Furthermore, for patents with multiple inventors the common approach is to use the residence of the "primary" inventor.[29] The residence information for the citing and the cited patent is used to identify the recipient and source (respectively) country of technological knowledge flow as indicated by the patent citation. The dependent variable (C_{ijt}) for our analysis is the total number of bilateral citations to the source country. That is the number of times a patent from the recipient country i cites a patent from the source country j in time t (where $i \neq j$). It is important to note that the distinction between the source and recipient country is crucial in indicating the direction of knowledge flow ($C_{ijt} \neq C_{jit}$). Recall that if $patent_i$ cites $patent_j$, this indicates that knowledge has travelled from $patent_j$ to $patent_i$ and not vice versa. The year of application ($AppYear_i$) is used to indicate the timing of knowledge flow. As highlighted by Hall et al. (2001), the inventors have a strong incentive to apply for a patent as soon as possible once the invention has taken place. Therefore using the application date should as closely as possible represent when the inventor received the knowledge spillover. However, when a patent is granted depends on the application process, which on average takes two years. Furthermore the USPTO has changed its process over the years, adding further reason not to use the date when a patent is granted (for further details see Hall et al., 2001).

Figure 3.6 below shows the annual number of bilateral citations by application year. The distribution resembles that of Figure 3.4. As we would expect, the more patents granted the more citations are made. However, as shown in Figure 3.7, the average number of bilateral citations per patent has increased over time causing the distribution of total citations to increase faster than the number of patents applied. An increasing trend in the average number of citations per patent could simply reflect that there are more patents (hence more technology) to cite. However, Hall et al. (2001) suggest this is also partly due to the system becoming computerised during the 1980's, meaning examiners could find potential citations more easily. Because we cannot distinguish between real changes and those due changes in citations practices at the USPTO it is important we control for the time variable in our regression analysis.

[29] The primary or lead inventor is identified in the data.

Figure 3.6 Total number of bilateral citations by application year

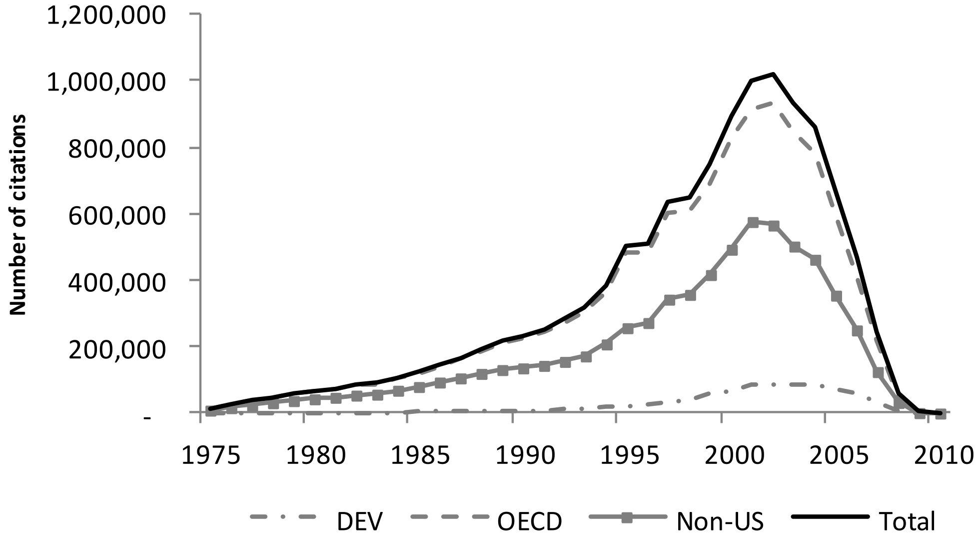
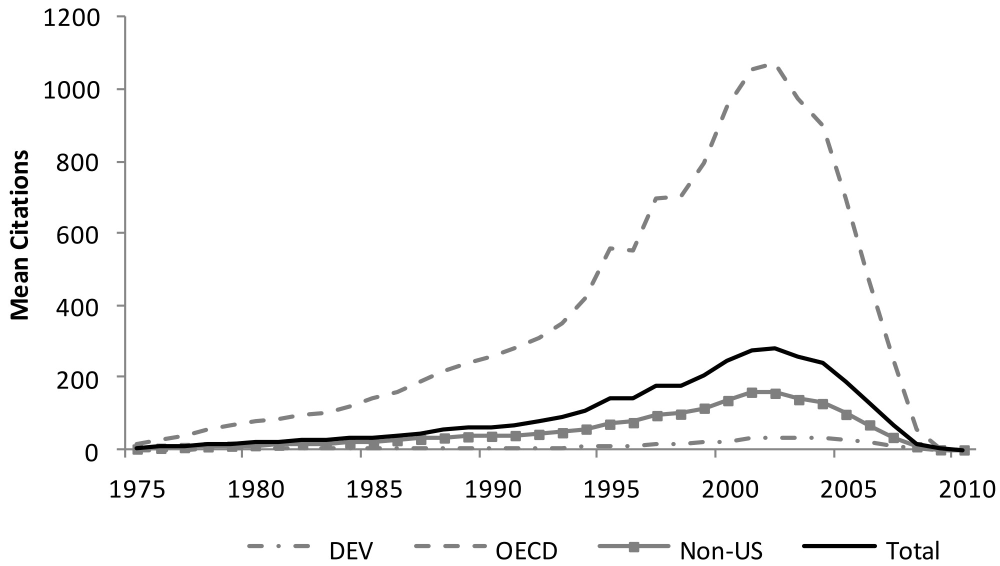


Figure 3.7 Average annual bilateral citations per country



Note: This is the average number of citations excluding the countries with no patents

Table 3.2 below provides an insight to how the total and average number of all patent citations compare to the total and average number of bilateral citations (to OECD countries). As we would expect, we can see that bilateral citations make up a large portion of developing countries' total citations.

Table 3.2 Total and bilateral patent citations

year	1980	1985	1990	1995	2000	2005
Developing Countries						
Total	1,441	3,038	8,051	25,254	74,967	84,080
Avg	7	11	20	47	86	112
Bilateral	1,218	2,756	7,140	21,573	64,381	70,883
Avg Bilateral	8	13	26	70	141	180
OECD Countries						
Total	160,288	276,422	546,034	1,288,308	2,313,130	1,825,367
Avg	293	421	693	388	1,977	1,709
Bilateral	66,043	121,124	224,586	483,197	827,732	605,523
Avg Bilateral	185	286	497	980	1,442	1,142
All Countries						
Total	161,729	279,460	554,085	1,313,562	2,388,097	1,909,447
Avg	217	298	465	895	1,168	1,050
Bilateral	67,261	123,880	231,726	504,770	892,113	676,406
Avg Bilateral	133	196	319	629	865	733

Note: The average represents the average number of citations between country pairs (eg. On average a developing country will cite a developed country 141 times in year 2000). It excludes country pairs with zero citations.

4. GRAVITY MODEL ESTIMATION

Following the work of Maurseth and Verspagen (2002); MacGarvie (2005); Li (2009); Picci (2010); and Montobbio and Sterzi (2012) we model the determinants of bilateral knowledge flows using an augmented gravity model specification. The gravity model has long been considered one of the most successful empirical models in economics. It has been a particularly popular model to explain international trade flows, first suggested by Tinbergen (1962). Gravity models, despite their success in empirical literature, were initially criticised for lacking theoretical foundations. However, Anderson (1979) explains how the gravity equation can be derived from a simple trade model with identical, constant income elasticity demand for tradable goods across countries. Over the last few decades gravity models have been used in a range of empirical studies of trade and factor movements between economic regions; such as investigating migration (see Karemera et al., 2000), tourism (see Genç, 2013), and FDI flows (see Bergstrand and Egger, 2007).

At its core, the basic gravity framework is perhaps best summarised by Anderson (2011, p.2):

> *The traditional gravity model drew on analogy with Newton's Law of Gravitation. A mass of goods or labor or other factors of production supplied at origin i, is attracted to a mass of demand for goods or labor at destination j, but the potential flow is reduced by distance between them.*

In other words, the usual interpretation is that the flow of economic goods, services or other factors of production between two countries is proportional to the product of the two economies' 'masses', and inversely proportional to the square of the distance between them. A simple representation of the model is:

$$flow_{ij} = G\left(\frac{E_i E_j}{D_{ij}^2}\right), \qquad (4.1)$$

where *flow*$_{ij}$ depends on the focus of the study. The most common application is modelling the determinants of international trade in commodities. E_i is the mass of economy *i*, D_{ij} is the distance between *i* and *j*, and G is a gravitational constant. In studies of international trade, the common interpretation of (4.1) is that larger, more productive economies will both supply and demand more goods to (from) the international goods market. However the amount of trade between countries decreases the further apart they are. The "traditional" interpretation of the negative effect of distance on trade is that the cost of transporting goods between countries increases with distance; therefore trade decreases as transport costs increase. Although somewhat intuitive, transport costs almost certainly do not explain the magnitude of the negative effect. Disdier and Head (2008) do a meta-analysis on the negative impact of distance on international trade. From the 103 papers analysed they find that, on average, a 10% increase in distance decreases bilateral trade by 9%. This is obviously a very large impact, which cannot be explained by the transport costs alone. In most industries transportation costs only represent a small fraction of the total value.[30] Given that transport costs cannot explain the large negative impact of trade, the literature has focussed on other factors that may inhibit bilateral trade. Cultural differences and language barriers are probably the most widely accepted variables to be included in gravity models. In alternative applications of the gravity model the variables added depend on the focus of the study; time zone differences, the level of telephone minutes, whether countries share a border, or if the country is part of a trade union (such as the European Union) are examples of explanatory variables added in various studies using a gravity model.

Typically, gravity models have been used where the dependent variable is continuous; such as the value of imported goods. The dependent variable we use in this thesis, however, is the count of patent citations. That is, our dependent variable takes on nonnegative integer values with no upper bound. Although it is common in the literature to use a linear transformation and estimate the model using linear regression, such as OLS, we will discuss the shortcomings of this approach later in this section. With count data it is better to model $E(Y_i|\boldsymbol{X}_i)$ directly and to assume data generating processes that ensure positive outcomes. A common example is the exponential function:[31]

$$E(Y_i|\boldsymbol{X}_i) = \exp(\boldsymbol{X_i\beta}). \tag{4.2}$$

[30] Glaeser and Kohlhase (2003) report that 80% of shipments occur in industries where transport costs represent as little as 4% of the total value of the shipment.
[31] See Wooldridge (1997) for an excellent discussion on estimation methods for count data.

The Poisson regression model is the benchmark model for count data analysis. Wooldridge (1997) explains that one reason for its popularity is because the "Poisson distribution is the nominal distribution for count data in much the same way the normal distribution is the nominal distribution for unbounded, continuous distributed data." (p.355). Furthermore, it turns out that the Poisson regression model has some nice robustness properties. In particular, consistent estimates of the conditional mean do not require the data to follow the Poisson distribution. This property has meant that the Poisson model is not only an attractive model for count data but any constant elasticity mean function. The Poisson regression model can be represented as:

$$\Pr\ (Y_i = y_i | \boldsymbol{X}_i\ ) = \frac{e^{-\mu_i}\ \mu_i^{y_i}}{y_i!} \tag{4.3}$$

where the mean parameter is $\mu_i = E(Y_i|\boldsymbol{X}_i\ ) = \exp\{\ \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip}\}$; Y_i is the outcome variable, and $\boldsymbol{X}_i = \{\ 1, X_{i1}, X_{i2}, \dots, X_{ip}\}$ is a vector of independent explanatory variables including a constant. Estimating the parameters of the Poisson model is straightforward and may be done by maximum likelihood. It can be shown that the Poisson distribution imposes a restriction that the conditional variance equals the conditional mean:

$$Var(y_i|\boldsymbol{X}_i) = E(Y_i|\boldsymbol{X}_i\ ) = \mu_i \tag{4.4}$$

This variance-mean equality is known as equidispersion property of the Poisson distribution. As we will see later in Section 6, our data is overdispersed, where the conditional variance is greater than the conditional mean. Overdispersion is common in economic applications and therefore is commonly highlighted that Poisson may not be suitable for analysis of such data.

Nevertheless, it turns out that Poisson model estimates of the conditional mean are consistent in the presence of deviations from the Poisson distribution. In fact, the consistency of Poisson maximum likelihood estimates does not require any distributional assumptions of $y_i|\boldsymbol{X}_i$ as long as the conditional mean is correctly specified.[32] When the Poisson model is estimated by maximum likelihood and the data do not follow the assumed Poisson distribution, the estimation procedure is commonly referred to as Poisson Pseudo Maximum Likelihood (PPML). Therefore, we can say that PPML

[32] See Wooldridge (2002) p. 649.

estimates of the conditional mean are consistent in the presence of overdispersion.[33] Wooldridge (2002) explains that in the case of overdispersion, the standard errors obtained from PPML can greatly underestimate the asymptotic standard deviations. This is similar to the case of heteroskedasticity in linear models. A simple solution is to adjust the standard errors, or to report robust standard errors.

As we have discussed, the Poisson distribution assumption is often too restrictive and almost never holds in economic applications. An alternative way to deal with overdispersion is to use models with less restrictive distributional assumptions, such as the negative binomial (NB). The NB model is an example of a continuous mixture model.[34] It specifically allows the data to be overdispersed by allowing the variance to adjust independently of the conditional mean. It does this by including one more parameter in the condition variance function than the Poisson model. More formally, the first two moments of the NB model are:

$$E(Y_i|\boldsymbol{X}_i) = \mu_i \tag{4.5}$$

$$Var(Y_i|\boldsymbol{X}_i) = \mu_i(1 + \alpha\mu_i). \tag{4.6}$$

The extra α parameter allows the variance to be larger than the conditional mean, unlike the Poisson model. In the context of the Poisson model, there is overdispersion when $\alpha > 0$, and it is easy to see that the NB model reduces to the Poisson when $\alpha = 0$. The added flexibility of the NB model has been found to fit the data well in applied economic studies, hence it is common to use NB whenever the data are overdispersed. Cameron and Trivedi (2005) warn us however, that such a mechanical approach should be avoided because the overdispersion may be due to misspecification of the conditional mean, and NB and Poisson maintain the same conditional mean function. We specifically test for misspecification of the conditional mean, in Section 7.2.

Despite the attractive features of the Poisson model, it is common in most applications of the gravity model in the literature to fit linear models to logarithmic transformations of the data. However, as pointed out by Flowerdew and Aitkin (1982), it turns out that a log linear specification of the gravity model suffers from three major problems: the bias created by the logarithmic transformation, the failure of the assumption that all error terms have equal variance, and the sensitivity of the results to

[33] Santos Silva and Tenreyro (2011) confirm that the PPML estimator is generally well behaved even when equidispersion does not hold.

[34] See Cameron and Trivedi (2005) ch.20 for further details on the NB model.

zero-valued flows (see also Burger et al., 2009). As in Flowerdew and Aitkin (1982) and Burger et al. (2009), we discuss these three problems in more detail below.

The first problem is that the logarithmic transformation affects the nature of the estimates produced. The log linear model produces estimates for the natural log of the dependent variable instead of its level form (i.e. ln C_{ij} and C_{ij} respectively). It is well known from Jensen's inequality that $E(\ln C_{ij}) \neq \ln E(C_{ij})$. Accordingly, Haworth and Vincent (1979) show that the antilog of the estimates produced by the log-linear gravity model are biased and underpredict large trade flows and total trade flows. This fundamental problem was largely ignored until recently (see, for example Santos Silva and Tenreyro, 2006).

The second major problem with a log-linear specification is that the model is based upon a strong assumption of homoskedasticity, or that the error terms have equal variance for all country pairs. As Flowerdew and Aitkin (1982) explains, under homoskedastic error terms it is assumed that a country pair with an observed trade flow of 1 in relation to an expected flow of 2 is equally as probable as an observed flow of 100 in relation to an expected flow of 200. Using the usual log linear specification with trade data Santos Silva and Tenreyro (2006) find significant evidence that the error terms fail to satisfy this assumption and conclude that the error terms are heteroscedastic. Intuitively, one might expect similar heteroskedasticity when studying knowledge flows instead of trade flows. Furthermore Santos Silva and Tenreyro (2006) show that estimates obtained using a log linear model in the presence of heteroskedasticity can be 'highly misleading' where estimates are no longer consistent or efficient.

The third problem arises because a log linear model is incompatible with the existence of zero value of the dependent variable, as often occurs in bilateral trade and patent citation data. This is because the logarithm of zero is undefined. Adhoc solutions to deal with the "zero problem" using a log linear model have been proposed. A common one found in empirical studies is to simply drop all zero valued observations. In studies on cross-country trade, close to half of the observations can be disregarded and in the current study an even larger proportion of bilateral citation counts are zero (as we will see in Section 6). By deleting all zero valued flows, a large amount of important information is lost because, as argued by Frankel et al. (1997), the most obvious reason for zero-valued trade flows is the lack of trade between small distant countries. A similar argument could be made for the flow of technological knowledge. Consequently, the truncation of the data will lead to endogenous sample selection problems and biased estimates, particularly when the zero-valued flows are non-randomly distributed (Burger et al., 2009). The second widely used approach to deal with the “zero problem” is to add a small non-negative constant to all flows, ensuring the logarithm is defined. This value is usually an arbitrary value, often between 0.1 and 1, that lacks theoretical and empirical justification (Linders and de Groot, 2006).What is more striking is that Flowerdew and Aitkin (1982) show that small differences in the chosen value can lead to significantly different results.[35]

In consideration of these problems with log linear gravity models, in addition to having count data, we focus our analysis on recently suggested estimation methods within the Poisson family, namely the PPML and NB models (see Silva and Tenreyro, 2006; and Burger et al., 2009). Such non-linear models allow us to estimate the determinants of the dependent variable directly without taking the logarithm by estimating the model in multiplicative form. Also, zero is a natural outcome in the Poisson distribution therefore we directly account for the “zero problem” faced by the inadequate log linear gravity models. Furthermore, when the independent variables are expressed in logs, Poisson model coefficients are elasticities, which is usually the main appeal in using a log linear specification.

Besides the attractive features of Poisson estimation already mentioned, a major advantage relates to the heteroskedasticity of the error terms. Silva and Tenreyro (2006) showed that estimates are consistent in the presence of heteroskedasticity when using a simple Poisson model estimated with pseudo maximum likelihood. Poisson estimates are also “reasonably efficient, especially in large samples” (King, 1988). Therefore as long as we report heteroskedasticity robust standard errors to account for the deflated standard errors caused by overdispersion, we retain consistent and relatively efficient estimates. It is worth noting that if we were interested in estimating more than the conditional mean (i.e. the probability of observing a given outcome; eg. $Y_i = 0$), satisfying other distributional assumptions is then important. However, the focus of this paper is on estimating the coefficients of the modified gravity model.

In line with the literature we use both the Poisson and negative binomial models and compare the results, while recognising from our discussion above, that the negative binomial model is not required in our case to deal with the overdispersion of the data.

---

[35] Also see King (1988); he shows that by altering the size of the constant, the researcher can generate any parameter estimate of their liking.

5. EMPIRICAL MODEL SPECIFICATION

To assess the determinants of bilateral knowledge flows we adopt an 'augmented' gravity model. This means we extend equation 4.1 with additional variables that may facilitate or inhibit knowledge flows between two countries. Primarily, we are interested in whether FDI plays a significant role in facilitating knowledge flows from the investing country to the host country. We estimate the following model:

$$C_{ijt} = P_{it}^{\alpha_1} P_{jt}^{\alpha_2} Dist_{ij}^{\eta} FDI_{ijt}^{\gamma} \exp\,(\boldsymbol{X}'\boldsymbol{\beta})\, e^{\tau t} \qquad (5.1)$$

where the dependent variable C_{ijt} is the number of patent citations by country i to country j in time period t. This represents the extent of knowledge flows from country j to country i, as explained in Section 3.1. The explanatory variables control for attracting and inhibiting factors of knowledge diffusion for each country pair as well as individual country characteristics.

P_{it} is the "inventive mass" of the recipient country i, that we proxy for with the number of patents applied for by country i in period t.[36] P_{jt} is a measure of "knowledge mass" or "stock of knowledge" of the source country j, that we proxy for using various measures of previous patents applied for by country j (these are explained in detail in Section 7.2). These "mass" terms are expected to have positive effects on knowledge flows analogous to the usual mass terms used in standard gravity models.

$Dist_{ij}$ is the geographic distance between the country pair. It is the distance (computed with the great circle formula) between the most important cities (in terms of population) of each country, measured in kilometres. The estimate of η is the (negative) elasticity of bilateral patent citations with respect to geographic distance. In line with the extensive literature we discussed in Section 2.3.1, we expect distance to have a negative effect on bilateral knowledge diffusion. The further apart two economies are the less likely they will 'learn' from one another.

[36] Recall that our data contain only patents that are granted. So this is the number of patent applications in year t, that were eventually granted.

FDI_{ijt} is the bilateral flow of FDI into the recipient country i from source country j. This is the key variable of interest in this thesis. We want to test whether the amount of FDI flow into i from j has an effect on knowledge flows from country j to country i, as indicated by the number of citations of country i patents to country j patents.

$\boldsymbol{X'}$ is a vector of explanatory variables representing factors that are expected to influence the amount of knowledge flows between a country pair. These variables are:

$$\boldsymbol{X'} = (GDP_{it}, GDP_{jt}, Import_{ijt}, IPR_{it}, IPR_{jt}, Language_{ij}, Colony_{ij}, Prox_{ijt})$$

$GDP_{it,jt}$ control for the size or "mass" of each economy. These terms are included essentially in accordance with gravity model convention, but underlying theoretical interpretation for a market size effect in a model of knowledge flows are not as clear-cut as the trade literature. Although inventive mass will be correlated with country size, one might still expect that, taking as given the inventive activity in a country and the stock of patents that may potentially be cited in the partner country, domestic market size is positively related to the proportion of global products developed for sale in this market, increasing the relevance of foreign technology for domestic innovative activity. This relationship may be more prevalent in the presence of increasing returns to global R&D and product development. Similarly, the size of the source country market may be positively related to citations because more domestic product development is intended for eventual sale in this market, increasing the relative relevance of technologies originally developed in this market.

$Import_{ijt}$ is the amount that country i imports from country j in period t. In line with the literature discussed in Section 2.3.2, tradable goods may be one channel that facilitates bilateral technological knowledge flows. If imports positively affect the number of patent citations, it would suggest that international trade indeed does facilitate technological knowledge to travel across national borders.

$IPR_{it,jt}$ is an index that indicates the strength of a county's intellectual property rights at the beginning of period t. The index, constructed by Ginarte and Park (1997), is made up of five sub-indices: (1) extent of coverage, (2) membership in international patent agreements, (3) provisions for loss of protection, (4) enforcement mechanisms, and (5) duration of protection. Each sub-index is given a score from 0 to 1 and the un-weighted sum of the five sub-indices make up the IPR index which ranges from 0 to 5, with higher values indicating stronger intellectual property rights.

$Language_{ij}$ is a dummy variable equal to one if the country pair shares a common official language, and $Colony_{ij}$ is a dummy variable equal to one if the country pair has ever had a colonial relationship. These variables are included in accordance with the idea that two individuals (countries) are more likely to share common interests, have economic relationships, exchange contracts, or communicate with one another if they have cultural similarities and are free of language barriers.

$Prox_{ijt}$ is a variable that indicates how close the two economies are in technological space. Controlling for such a relationship was first suggested by Jaffe (1986); the author states "I assume the existence of technological spillovers implies that a firm's R&D success is affected by the research activity of its neighbours in technological space" (p.5). Patent application data provide the researcher a unique opportunity to observe a firm's (or in our case, country's) 'technological position' from the technology-based patent classes assigned to each patent. With K classes, a firm's technological position is characterised by a vector $F = (F_1, \ldots, F_k)$ where F_k is the share of patents granted into technological class K. Jaffe's measure of proximity in technological space TP_{ij} is the un-centred correlation of each firm's vector F.

$$TP_{ij} = \frac{F_i F_j'}{[(F_i F_i')(F_j F_j')]^{1/2}} \tag{5.2}$$

This measure ranges from 0 to 1, where it is one for firms with identical position vectors and zero for firms whose vectors are orthogonal (Jaffe, 1986).

The USPTO's classification system comprises of over 400 main (3-digit) patent classes, and over 120,000 nested subclasses. Each patent must be assigned one "original" classification, in addition to any number of subsidiary classes and subclasses. Even if the patent application includes two or more claims that belong in separate classes, for administration and legal purposes a patent can only have one "original" class. This "original" patent class is decided according to a detailed hierarchy of rules.[37] Although most patents have more than one main patent class, Hall et al. (2001) state "for

[37] See the USPTO Examiner's Handbook Chapter Four – "Determination of class for Original Classification or Assignment for Examination." http://www.uspto.gov/patents/resources/classification/handbook/four.jsp

the vast majority of uses one is most likely to resort only to the original 3-digit patent class"; hence the widely used NBER patent data set only includes the "original" patent class. However, Benner and Waldfogel (2008) explains in detail why one patent class does not represent the patent's location in technological space, hence is not suitable for proximity measures. Furthermore, Benner and Waldfogel provide reasons why using coarser patent class partitions is preferred over fine partitions (such as the 400 main classes) for assessing technological distances.

Hall et al. (2001) provide higher-level or coarser divisions where the 400 patent classes are aggregated into 6 main categories, including: Chemical (excluding Drugs); Computers and Communications; Drugs and Medical; Electrical and Electronics; Mechanical; and Others. We group all main classes assigned to each patent into one of these categories, the location in technological space of patent n is then represented by a 1x6 vector (V_n) where the elements are the proportion of total main classes allocated to patent n that belong to the corresponding categories. The technological location of country i at time t is then calculated by the averaging across all individual patent vectors invented in the country in time t.

$$F_{it} = \frac{\sum_{n}^{N_{it}} V_n}{N_{it}} \tag{5.3}$$

where N_t is the number of patents applied for by country i in time t. $Prox_{ijt}$ is then calculated according to equation 5.2 for each time period. To summarise, $Prox_{ijt}$ "measures the extent to which the distributions of patents across technological classes in the two countries overlaps" (MacGarvie, 2005, p.122). This variable controls for the fact that an inventor is more likely to cite a patent within a similar technological field as his/her own invention.

$\boldsymbol{X'}$ also includes dummy variables for the recipient and source country, as well as a time dummy to control for unobserved heterogeneity (when we refer to the Poisson and NB models, these dummy variables are included). Alternatively we estimate a Poisson fixed effects (FE) model that controls for country-pair fixed effects.[38] In the latter model, the coefficients of time-invariant variables cannot be estimated; consequently they are dropped from the model. In this model the analysis then solely

[38] The negative binomial FE model in STATA is not a regular fixed effects model. Therefore, we were recommended against estimating with this technique.

focuses those variables that have significant variation over time. All variables except the dependent variable C_{ijt}, $IPR_{it,jt}$, $Prox_{ijt}$, and dummy variables are measured in logs.

Each time period in our model represents a 5-year period (years t to $t-4$), the variables that vary over time are aggregated over the period (except $Prox_{ijt}$ is the average). Firstly, this is to align our data to the IPR index, which is available in 5-year periods. Secondly, we suspect the relationship between FDI and knowledge flows to be a complicated dynamic process that varies between countries and industries. Therefore, aggregating the data into 5-year periods reduces the precision necessary for modelling a dynamic knowledge flow process and enables us to more accurately estimate the medium term effect of FDI on knowledge spillovers. The trade-off, of course, is that some of the information is discarded. We discuss FDI lags in more detail in Section 7.

6. DATA DESCRIPTION

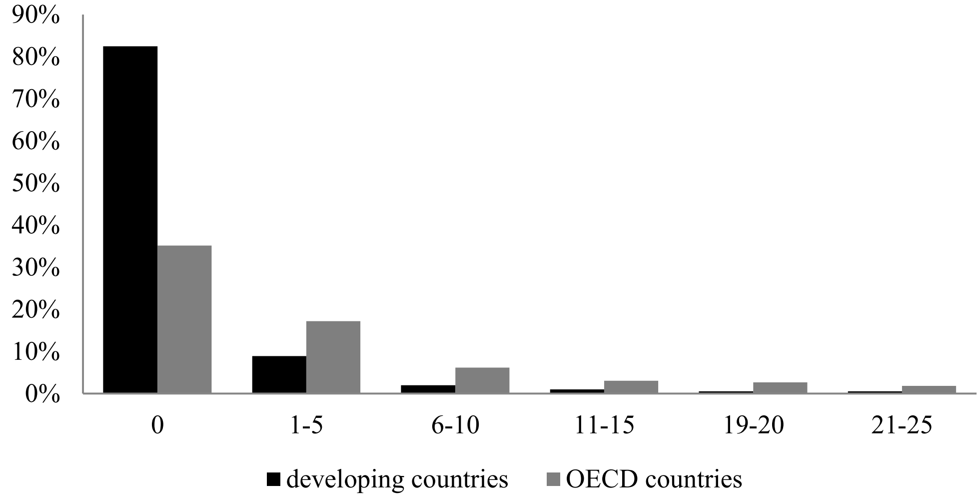
Our data set begins with the recently available Harvard patent database that includes all patent granted to the USPTO from 1975 to 2010. Having data this up to date is particularly important for studying the effects of FDI. As shown in Figure 2.1, the amount of FDI globally has rapidly increased since the late 90's. Therefore recent data provides us the best possible opportunity to analyse the impact an increase in FDI has on bilateral knowledge spillovers. Our sample starts with 122 countries, including 32 OECD countries which we refer to as developed countries, and the remaining countries which are referred to as developing countries. We construct two samples: the "developing sample" has developing countries as the citing country (which in this context we refer to as the recipient country)[39] and OECD countries as the cited country (which we refer to as the source country).[40] This sample allows us to study the knowledge flows from developed to developing countries. The second "OECD sample" includes OECD countries as both the recipient and source country. This sample allows us to study knowledge flows between developed countries. The countries included in our sample are determined by the availability of the IPR index. A complete list of the countries is provided in the Appendix A1. The total number of possible observations is given by the number of country-pair combinations ($N \times n$), where N is the number of recipient countries, n is the number of source countries, multiplied by the number of time periods, T.[41]

As we described in Section 5, our dependent variable is the number of patent citations from the recipient country to the source country in time t. This variable is highly skewed with a large number of countries not citing a given source country at all in period t. A zero observation can either be because the recipient country did not have any patents in the period, or they had patents but did not cite the source country. With developing countries patenting far fewer inventions, zero-valued dependent variable observations make up a larger portion of the developing sample. Figure 6.1 shows that

[39] Referring to the recipient of knowledge flows.
[40] Referring to the source of knowledge flows.
[41] However, this is significantly decreased due to missing data for our explanatory variables, particularly bilateral FDI data.

for over 80% of all possible observations in the developing country sample, the dependent variable is zero. Furthermore, from the summary statistics in Table A1.5 provided in Appendix A1, we can see that this variable is overdispersed because the variance is much larger than the mean (in both samples).[42] As we have discussed in Section 4, by estimating using a Poisson model we directly account for the issue of having a large number of zero-value dependent variable observations.

Figure 6.1 Proportion of country-pair citation counts within each range



Note: This diagram is truncated at 25 bilateral citations for presentation purposes. Also note that this is the distribution of all possible country-pairs, not of those included in the regression analysis (which is a considerably smaller sample due to missing data of the explanatory variables).

The main variable of interest in this thesis is the FDI flow into the recipient country. We use the OECD bilateral FDI flows data available from 1985 to 2010. These data are nominal FDI flows in current US dollars.[43] FDI is defined as "the objective of obtaining a lasting interest by a resident entity in one economy ("direct investor") in an entity resident in an economy other than that of the investor ("direct investment enterprise")" (OECD, 1999, p.7). The data are for OECD member countries as the reporting country and all countries as the partner country for both inward FDI flows and outward FDI flows. We use the OECD to developing country outward FDI flows as the developing country inward FDI flows from the OECD country.[44] It is worth noting here

[42] We also carry out a formal test for overdispersion where the null hypothesis is that $\alpha = 0$ in equation (4.6). The null hypothesis is rejected, which confirms our data are overdispersed.
[43] The gravity model trade literature uses both deflated and non-deflated values for bilateral trade flows.
[44] In theory this should be the same figure. However when analysing the data where the reporting and partner country is an OECD member, Outward FDI from j to $i \neq$ Inward FDI into i from j. This is due to various reporting issues.

that FDI flows can take on negative values, representing a decrease in the FDI presence in the partner country.[45]

We also collect data for GDP (source: World Bank), bilateral imports (UN Comtrade), intellectual property rights (Park and Ginarte, 1997), distance, language and colonial relationships (CEPII), and construct a technological proximity value (Harvard patent data set). Table A 1.4 in Appendix A1 provides precise definitions of each variable and the data sources, and Table A 1.5 provides summary statistics for each sample.

[45] A change in FDI may also be a result of a change in its market value. The OECD Benchmark Definition recommends market value as the conceptual basis for valuation. Market valuation places all assets at current prices rather than when purchased or last revalued, and allows comparability of assets of different vintages. See the OECD website for further information on the valuation on FDI.

7. SPECIFICATION TESTING

7.1. Lagged Effects of FDI

The major research question in this thesis is whether foreign direct investment plays a significant role in facilitating knowledge to flow from the investing country to the host country. From the outset, we acknowledge that aggregate FDI flows follow a complex dynamic process that captures responses to several kinds of economic incentives. For instance, a firm may establish a manufacturing plant in another country with cheaper labour to decrease the cost of production; or a firm may pursue a vertical expansion to secure the supply of productive inputs. Alternatively, another country may provide a more secure business environment which may be the deciding factor of where a foreign firm sets up new operations. The various possible reasons for FDI suggest that all FDI is not homogenous in terms of its impact on knowledge flows. Different modes of FDI are expected to have different effects on the extent and rate of knowledge diffusion and therefore we do not expect the relationship between FDI and knowledge flows (as measured by patent citations) to be, in any sense, straightforward to model.

In line with this, we include the current period FDI and the first lagged FDI term in order to capture the medium term effects of FDI on knowledge flows. We justify this by analysing the FDI lag structure using data in its annual form.[46] We begin with the baseline model specification represented by equation 5.1. To this model we incrementally add additional annual FDI lagged terms until the last term is insignificant. The corresponding estimation output is in Appendix A2, Tables A 2.1-2.6. In both samples and all three estimation techniques,[47] we find that increasing the length of the lag structure improves the model, as indicated by larger log likelihood values, and smaller AIC values. However, which individual lagged terms are statistically significant depends on the sample, and the estimation technique. From this analysis it is unclear what the ‘optimal’ lag structure is for the FDI process. However, because FDI lagged 5, 6, or 7 years is significant in some of the models, it signifies that it is not only the recent FDI that has a significant effect on the number patent citations. This number of

[46] Annual data refers to when the data are not aggregated into 5-year periods and we linearly interpolate the IPR index data to fill the missing years.
[47] Poisson, NB, and Poisson FE

significant annual lags falls within the first lagged 5-year period.[48] As mentioned in Section 5, a benefit of aggregating the data into 5-year periods allows the model to control for the medium term effects of FDI without necessarily selecting a precise lag structure that is bound to vary between countries.

7.2. Measures of Knowledge Stock

When determining the model specification we consider alternative measures of the "stock of knowledge" of the source country. The underlying theory of the gravity equation is that the observed factor flow is proportional to the mass of both country partners and negatively proportional to the distance between them. Patent citations control for the current inventive activity ("inventive mass") of the citing country, and the "stock of knowledge" available for them to cite in the source country. In other words, we would expect that a country that is actively patenting more frequently overall to cite more patents from other countries compared to a country that is patenting few inventions, all else equal. Similarly, for a given level of country *i* domestic inventive activity, the larger country *j*'s stock of knowledge, the more likely country *i* will cite an idea from country *j*, all else equal.

Following Maurseth and Verspagen (2002); MacGarvie (2005); and Picci (2010) we use the total number of patents applied for by country *i* in the current period as a proxy for the "inventive mass" of the reporting country. However we believe it is unclear how the model should control for the "stock of knowledge" of the partner country. We suggest that including the total number of patent applications (in the current period) for both reporting and partner country does not control for the total number of bilateral citations possible in the period because inventors may also cite earlier patents.[49] In particular, bilateral citations are influenced by the cumulated stock of patents in the source country, albeit perhaps adjusted for "depreciation" of older, outdated knowledge. In our view, the literature fails to address this, or at least in a transparent manner.[50]

We explore the following alternative measures of the source country's stock of knowledge: all existing patents in the sample, all existing patents in the sample adjusted

[48] Including the second 5-year period lag resulted in a large loss in the time-series dimension of the panel.
[49] In fact, Hall et al. (2001) show that around half of citations in their sample are made to patents at least 10 years older than the citing patent.
[50] Maurseth and Verspagen (2002) includes the product of the number of patents from recipient and source regions. However, it is unclear what the measures of patents are, current period or stock.

for depreciation,[51] patents from the last 5 years, patents from the 5 years prior to the current period, and all patents from the last 10 years. We also considered the interaction terms between the number of patents applied for by the reporting country and the various stock measures of the partner country (unrestricted versions of the model), but the model performed significantly worse than the restricted versions; therefore we left them out of the analysis.

Similar to the way we analyse alternative FDI lag structures, we consider all the alternative patent stock measures for both samples, and the three estimation techniques. In addition, we test the specification of the mean function for each model using a regression error specification test (RESET). As discussed in Section 4, the pseudo maximum likelihood estimates of the parameters in our model are consistent as long as the conditional mean function is correctly specified. One way to test this is using a modified version of the Ramsey RESET test (Ramsey, 1969). Sapra (2005) developed a RESET test for generalised linear models (GLM) as an extension of the typical RESET test that is commonly used for linear models. The modified RESET test compares the GLM with no higher order terms to a GLM with the higher order terms. We follow Sapra (2005) and include the squared and cubed predicted values of the dependent variable; the test is formally expressed as:

$$Y_{1i} = \beta' \boldsymbol{x}_i$$

$$Y_{2i} = \beta' \boldsymbol{x}_i + \gamma_1 \hat{Y}_{1i}^2 \tag{7.1}$$

$$Y_{3i} = \beta' \boldsymbol{x}_i + \gamma_1 \hat{Y}_{1i}^2 + \gamma_2 \hat{Y}_{1i}^3 \tag{7.2}$$

where $\hat{Y}_{1i} = \hat{\beta}' \boldsymbol{x}_i$ and $\hat{\beta}$ is the PMLE of β. We test whether the higher order terms have any explanatory power in (7.1) and (7.2).

$$H_0 : \gamma_1 = \gamma_2 = 0$$

$$H_a : At\ least\ one\ is\ not\ equal\ to\ zero$$

When the model is estimated using maximum likelihood, a likelihood ratio (LR) test is generally preferred. However, we cannot assume the conditional distribution is correctly specified. In this case, the likelihood function is unlikely to be correctly

[51] We apply a 50% annual depreciation rate.

specified and the LR test will generally be invalid. Therefore a robust Wald test is the much preferred method for the RESET test for our purposes.[52] If we fail to reject the null hypothesis of no misspecification, we can assume that the mean function follows a log-linear specification; hence Poisson PML and negative binomial PML estimates are consistent.

The regression estimates along with the statistics used for model comparison and p-values from the RESET are reported in the Appendix A2. Tables A 2.7-2.9 are for the developing countries sample, while Tables A 2.10-2.12 are for the OECD sample. Each table uses an alternative estimation method; Poisson, Poisson FE, and NB respectively. However, we concentrate on RESET test, and the p-values are summarised in Table 7.1 below. The columns of Table 7.1 represent an alternative measure of patent stock for the partner country; (1) all existing patents, (2) all existing patents adjusted for depreciation, (3) number of patent applications in the current 5-year period, (4) number of patent applications in the previous period, and (5) number of patent applications in the last two periods (10 years).[53]

Table 7.1 RESET test p-values

Sample	Model	Test	(1) Stock	(2) Dep Stock	(3) 5yr	(4) L.5yr	(5) 10yr
DEV	Poisson	(7.2)	.953	.989	.844	.930	c
		(7.1)	.944	.964	.562	.712	.577
	Poisson FE	(7.2)	.013	.010	.02	.000	.000
		(7.1)	.063	.077	.006	.008	.000
	NB	(7.2)	c	c	0.282	0.209	c
		(7.1)	c	.330	.530	.369	.464
OECD	Poisson	(7.2)	.000	.000	c	c	c
		(7.1)	.074	.017	.007	.000	.001
	Poisson FE	(7.2)	.000	.000	.169	.017	.463
		(7.1)	.268	.201	.107	.084	.375
	NB	(7.2)	c	c	c	c	c
		(7.1)	.273	.495	0.238	0.647	.432

c. The model did not converge

In Table 7.1, p-values greater than 0.05 fail to reject the null hypothesis of the RESET test at the 5% level, suggesting that the evidence for an incorrectly specified conditional mean is not especially strong for many of the model specifications. In some

[52] We confirm this by personal communication with J. Santos Silva.
[53] The columns in the full tables in the appendix follow the same format.

instances the model does not converge when the cubed fitted value is added to the model,[54] in which case the outcome of the more powerful version of the RESET test cannot be determined.

At this point it is worth noting that the coefficient estimates of all three models are very robust to alternative patent stock measures. Across the different measures, almost all of the estimates are both quantitatively very similar and of equal statistical significance (except the variable we change). However, the patent stock measure does appear to matter for the specification of the conditional mean.

Table 7.1 shows, for the developing country sample, that all specifications for Poisson models do not reject the null hypothesis in the RESET test.[55] For the negative binomial model, in all specifications where we obtain convergence to a solution we can conclude that the mean function is correctly specified. However, in the case of the Poisson FE models, we can reject the hypothesis that the conditional mean is correctly specified at the 5% level of confidence. Therefore, in further analysis we primarily focus on the Poisson and negative binomial models and refer to the Poisson FE where there are significant differences. In doing so we must be mindful that the fixed effects models estimates may not be reliable.

For the OECD sample, we reject the null hypothesis for all five specifications for the Poisson model without fixed effects. For Poisson FE models (3) and (5) we do not reject the null hypothesis at the 5% level. (However the p-values are generally quite low, indicating that misspecification my still be of some concern.) Similarly, we do not reject the null hypothesis for all five specifications of the negative binomial models.[56]

On balance, we can conclude that specifications (3) and (5) are correctly specified across the alternative estimation techniques. The fact that the coefficient estimates on the patent stock variable in specifications (1) and (2) are either negative or statistically insignificant, further indicates that these models are likely to be misspecified in some sense. In considering among (3) and (5) which is our preferred specification, we analyse the model comparison statistics reported at the bottom of Tables A 2.7-2.12. In

[54] Non convergence is particularly common with non linear models such as the Poisson and negative binomial. In particular, the Poisson command in STATA is not very good at dealing with numerical problems. We tried a range of possible solutions suggested by J. Santos Silva & S. Tenreyro on their “Log of Gravity” website. However, non convergence persisted.

[55] Specification (5) does not converge when the cubed term is included. However the null hypothesis is not rejected when only the squared term is included.

[56] For the RESET test with only the squared predicted values included in the test. The cubed models do not converge.

particular, the Akaike Information Criterion (AIC) can be used to choose between models, with lower values indicating a relatively better fit. Specification (5) has a lower AIC than (3) in five of the six models; therefore we use the number of patent applications in the last ten years as our baseline measure for the patent stock of the partner country. For robustness we compare our results from specification (3).

8. EMPIRICAL RESULTS

In this section, we present the results from each of the three estimation models we use, for six alternative specifications to our baseline model represented by equation 5.1. We focus primarily on estimating international patent citations for the developing country sample, and briefly compare our findings to those obtained for patent citations between OECD countries. We begin interpreting the results from the Poisson model without country-pair fixed effects, and then compare these to our negative binomial and Poisson FE estimates.

Table 8.1 reports the Poisson model estimates for the developing country sample. The first column is the baseline model represented by equation 5.1. Column (2) includes an additional interaction term, and columns (3) to (6) are extensions of the model that account for zero and negative FDI_{ijt} values and zero-valued P_{it} observations.

Table 8.1 Poisson estimates for the developing sample

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.803**	0.746**	0.789**	0.720**		
	(0.067)	(0.064)	(0.074)	(0.065)		
ln (P_{it} + 1)					0.793**	0.725**
					(0.074)	(0.065)
ln P_{jt}	0.640*	0.561*	0.490*	0.506*	0.490*	0.506*
	(0.259)	(0.250)	(0.233)	(0.235)	(0.233)	(0.235)
ln GDP_{it}	-0.540†	-0.256	-0.686*	-0.447	-0.692*	-0.452†
	(0.302)	(0.289)	(0.294)	(0.272)	(0.294)	(0.272)
ln GDP_{jt}	0.347	0.440	0.692	0.533	0.695	0.539
	(0.769)	(0.704)	(0.674)	(0.639)	(0.674)	(0.638)
ln $Dist_{ij}$	-0.133**	-0.090*	-0.111*	-0.071†	-0.111*	-0.071†
	(0.043)	(0.042)	(0.044)	(0.043)	(0.044)	(0.043)
$Lang_{ij}$	0.144*	0.088	0.160*	0.096†	0.160*	0.095†
	(0.063)	(0.055)	(0.071)	(0.056)	(0.071)	(0.056)
$Colony_{ij}$	-0.168**	-0.137*	-0.217**	-0.189**	-0.217**	-0.189**
	(0.064)	(0.067)	(0.064)	(0.068)	(0.064)	(0.068)
$Prox_{ijt}$	1.919**	1.989**	2.028**	2.055**	2.040**	2.068**
	(0.253)	(0.215)	(0.260)	(0.224)	(0.258)	(0.223)
IPR_{it}	-0.187**	-0.484**	-0.171**	-0.415**	-0.171**	-0.426**
	(0.034)	(0.138)	(0.037)	(0.123)	(0.037)	(0.122)
IPR_{jt}	-0.402	-0.316	-0.359†	-0.308†	-0.359†	-0.306†
	(0.251)	(0.219)	(0.193)	(0.166)	(0.193)	(0.166)

ln $Import_{ijt}$	0.064†	0.140**	0.114**	0.166**	0.113**	0.166**
	(0.039)	(0.036)	(0.040)	(0.038)	(0.040)	(0.038)
ln FDI_{ijt}	-0.043**	-0.131**				
	(0.015)	(0.039)				
ln FDI_{ijt-1}	0.025	-0.016				
	(0.019)	(0.019)				
ln FDI_{ijt}			-0.020	-0.092**	-0.020	-0.095**
(positive)			(0.015)	(0.034)	(0.015)	(0.034)
ln FDI_{ijt-1}			0.000	-0.022	0.000	-0.022
(positive)			(0.012)	(0.016)	(0.012)	(0.016)
ln FDI_{ijt}			-0.005	-0.116*	-0.005	-0.119*
(negative)			(0.014)	(0.054)	(0.014)	(0.054)
ln FDI_{ijt-1}			-0.015	0.038	-0.016	0.037
(negative)			(0.013)	(0.031)	(0.013)	(0.031)
ln FDI_{ijt}			0.155	0.093	0.150	0.084
(zero)			(0.205)	(0.233)	(0.205)	(0.234)
ln FDI_{ijt}* IPR_{it}		0.023*				
		(0.010)				
ln FDI_{ijt-1}* IPR_{it}		0.009**				
		(0.002)				
ln FDI_{ijt}* IPR_{it}				0.018*		0.019*
(positive)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				0.008**		0.008**
(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				0.026*		0.027*
(negative)				(0.013)		(0.013)
ln FDI_{ijt-1}* IPR_{it}				-0.011†		-0.011†
(negative)				(0.006)		(0.006)
Observations	1,040	1,001	1,951	1,885	2,766	2,700
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Chi2 (model)	72504	99543	69735	95532	97321	153923
Log likelihood	-5052	-4536	-7293	-6575	-7316	-6602
AIC	10308	9280	14800	13372	14877	13457
BIC	10813	9790	15397	13987	15599	14200

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

We must be careful when interpreting the magnitude of the coefficients. They should not be interpreted as quantitative estimates of ‘total’ knowledge spillovers. As discussed in Section 3.1, patent citations can be viewed as indicators of knowledge spillovers, keeping in mind patent citations represent knowledge flows between similar and patentable technologies. Patent citations may differ substantially across product or industry categories in terms of the extent and industrial value of the knowledge spillover. Because the underlying focus of this thesis is on the impact of FDI (among other factors) on total bilateral knowledge spillovers, we emphasise the sign and significance of the model estimates without drawing any inference from the size of the estimated elasticities. Nonetheless, all estimated coefficients (of the variables measured in logs) are the elasticity with respect to the number of bilateral patent citations.

8.1. Developing Country Sample Poisson Model Estimates

8.1.1. The Impact of FDI

For the baseline model in column (1), FDI inflows have a negative effect on the number of patent citations, while lagged FDI flows do not have any significant effect. In other words, the more FDI a developing country receives from a developed OECD country, the less knowledge spillovers there appears to be. Taken at face value, this effect is contrary to what we would expect (It is worth noting, however, the findings in the literature are highly mixed). However, as previously discussed, the FDI flows are a complex and dynamic process which is difficult to model precisely using available data. In particular, data on aggregate FDI flows do not provide us with information on the primary motives of investment across different countries, which may vary in the importance of knowledge spillover and intellectual property depending on the country's industry specialisation. Potentially only a small fraction of aggregate FDI involves technologically advanced business, or dedicated to R&D, especially when considering FDI into developing countries that attract large amounts of resource-based FDI. It is possible that the negative effect of relatively IPR-insensitive foreign investment, such as mineral extraction and utilities, is crowding out a technologically advanced, IPR-sensitive FDI. With strong IPR regimes, firms are more able to secure their technologies that give them a comparative advantage. Hence with a stronger IPR regime, foreign subsidiaries will tend to receive sensitive knowledge and technology from their parent multinational firm (Stephan, 2011). (For a similar view-point, see Branstetter et al., 2011.)

In line with this view, we hypothesise that developing countries with stronger IPR will attract FDI more conducive to technological knowledge spillovers. In column (2) of Table 8.1 we include interaction terms between IPR index of the developing country and the FDI flow variables. We find positive and statistically significant coefficients of both current and lagged FDI interaction terms. Typically, in linear models, the interaction effect can be taken directly from parameter associated with the interaction term. However, in non-linear models the interaction effect also depends on the values and parameters of the variables included in the interaction term. The interaction effect is represented as follows:

$$\frac{\partial^2 E[C|\boldsymbol{X}]}{\partial IPR \partial \ln(FDI)} \cdot \frac{1}{E[C|\boldsymbol{X}]} = \beta_1\beta_2 + \beta_{12}[1 + \beta_1 IPR + \beta_2 \ln(FDI) + \beta_{12} IPR \times \ln(FDI)] \tag{8.1}$$

where β_1, β_2, and β_{12} are the coefficients on IPR_{it} , $\ln FDI_{ijt}$, and the interaction term, respectively. The derivation of this effect is in Appendix A3. This essentially implies that the interaction effect will vary across countries with different levels of FDI and IPR. Therefore, to analyse whether strengthening IPRs has an impact on the effect FDI has on knowledge flows, we evaluate the interaction term effect at the average $\ln FDI_{ijt}$ and three alternative IPR strengths: low, medium, and high. We find a positive interaction effect for all three levels.[57] This finding supports our hypothesis; stronger intellectual property rights stimulate knowledge spillovers via FDI *relative to countries with poor IPR*. It is not the magnitude of FDI inflows that matter for knowledge diffusion, but the magnitude of the 'correct' type of FDI. This conclusion is also consistent with the findings of Branstetter (2006). Branstetter provides evidence that international knowledge flows are considerably stronger through Japanese firms with subsidiaries in the U.S. who possess a technological advantage over its American competitors, than 'total' Japanese FDI into the U.S.

In column (2), to determine whether there an increase in FDI has a positive or negative effect on patent citations we analyse the following elasticity:[58]

$$\frac{\partial E[C|\boldsymbol{X}]}{\partial \ln(FDI)} \cdot \frac{1}{E[C|\boldsymbol{X}]} = (\beta_2 + \beta_{12} IPR) \tag{8.2}$$

We find that an increase FDI remains to have a contemporaneous negative net effect on bilateral patent citations for the three levels of IPR, but the lagged effect is positive for countries with medium and high IPR. This suggests that the negative crowding out effect initially dominates positive influence that increasing IPR has on attracting knowledge

[57] The interaction term is also positive for the lagged FDI terms.
[58] The derivation of this elasticity is included in the derivation of the previous elasticity.

flows, but in the "medium term" there may be a positive net effect of FDI on knowledge flows.

FDI flows can take on positive or negative values, where a negative FDI flow represents a net decrease of OECD country j's FDI presence in developing country i. Because we model the continuous variables by taking the natural logarithm, the country pairs with negative or zero value FDI flows are dropped from the sample. To allow these observations to enter the model we decompose FDI flows into three separate variables: positive, negative, and zero values. We then take the natural logarithm of the positive-FDI and absolute value of the negative-FDI variables, and include a dummy variable for zero value FDI. This allows a great deal of flexibility in the model, and almost doubles the sample size (see Table 8.1, column 3). If a negative net inflow of FDI represents a decrease in the amount of foreign knowledge available to the receiving developing country, then we might expect a negative effect on the number of patent citations. Comparing column (1) and (3), this variation of the model makes little difference to the magnitude and significance of the rest of the variables, but the FDI terms are no longer statistically different from zero. Column (4) is our preferred model where we include interaction terms between the positive and negative FDI variables and the IPR index of country i. The estimated coefficients and calculated effects of the positive-FDI interaction terms are analogous to column (2). Similarly, we find that the negative-FDI interaction effects are also positive. We suggest this could be picking up the fact that, for there to be a negative net inflow of FDI, there must be positive FDI stocks in the recipient country originating from this source country, implying relatively large amounts of foreign investment from this source country in previous years. In this case, the positive effect could be also be interpreted as a lagged effect of previous FDI from the OECD country.

In columns (5) and (6) we accommodate for information that is lost because the model drops country pairs where the developing country did not apply for any patents in the period. In this case the natural logarithm of P_{it} is not defined, and the technological proximity term is also undefined by construction. Obviously no citations can be made if there are no patents applied for. However, these observations are relevant "zeros" in the sample, and may be explained by our model. To accommodate for these observations we follow a common ad hoc solution by adding 1 to the number of patents applied for by the developing countries. The technological proximity term is set to zero for the country pairs where the developing country did not apply for any patents. This adds more than

700 observations to the sample. However, we find no important differences in the main results.

8.1.2. Gravity Effects

The "inventive mass" of the recipient country has the expected positive and significant effect on patent citations. Developing countries with high levels inventive activity on average cite more patents from the developed OECD countries. Similarly, OECD countries with larger stocks of patents (knowledge) are cited more often by patents invented in developing countries.

There is weak evidence (at the 10% level) that developing countries' GDP has a negative effect on the number of citations to OECD patents. If GDP is regarded as a gravity model "mass" term, this finding is in contrast to theory that suggests flows should increase with the size of the economies. An alternative interpretation may be that larger countries need to specialise less, and are therefore more self-reliant in terms of access to variety of technological knowledge. As an economy increases in size, firms can utilise more knowledge from their domestic counterparts as opposed to seeking foreign sources. Montobbio and Sterzi (2012) describes a similar negative effect as the "supply effect"; a larger economy can supply more knowledge to domestic firms. The GDP of the knowledge source (OECD country) has a positive but insignificant effect on the number of patent citation in this Poisson model. As we will see, the estimated effects of GDP are sensitive to the estimation technique used.

Geographic distance has a negative impact on bilateral patent citations. A 10% increase in the distance between two countries decreases the number of bilateral patent citations by 1%. This is precisely the magnitude MacGarvie (2005) found when studying patent citations between 10 developed countries. Similarly, Li (2009) finds an elasticity of around -0.13%. Using the same interpretation as above, a 10% increase in distance decreases knowledge flows by 1.3%. Or as the author put it, "halving distance will increase knowledge flows by 6.5%" (p.13). Our finding that distance has a negative effect on bilateral patent citations is also consistent with other empirical research on knowledge spillovers, using patent data (see Jaffe et al., 1993; Jaffe and Trajtenberg, 1999; Eaton and Kortum, 1999; Maurseth and Verspagen, 2002; Peri, 2002, 2005; Thompson, 2006; and Picci, 2010), productivity measures (see Keller, 2002b), and other measures of knowledge flows such as book translations (see Sin, 2012). We reiterate some interesting viewpoints from these analyses.

Although knowledge is intangible and largely exempt from transport costs (by their usual interpretation), geographic distances impact the flow of knowledge in other ways. As highlighted by Montobbio and Sterzi (2012), a large part of knowledge is non-codifiable, therefore usually requires face-to-face exchanges to be passed on. Even despite modern technologies that facilitate easier communication through video conferencing, etc., there is evidence that the distance effect is not diminishing (see, for example Li, 2009; and Montobbio and Sterzi, 2012), although Keller (2002b) and MacGarvie (2005) present results to the contrary. As in the trade literature (see Disdier and Head, 2008), there remains an unsolved puzzle on the persistent negative effect of distance on knowledge diffusion.

8.1.3. Other Factors Effecting Knowledge Spillover

Our estimates suggest that how close two countries are in technological space is important for knowledge flows between them. The proxy we use for technological proximity has a large and significant positive effect on patent citations.[59] This suggests that countries that have similar compositions of their knowledge portfolios also share more knowledge between them. This is consistent with Peri (2005), who finds that "regions with technological specialisations in completely different fields have knowledge flows only 5% as large as those with identical technological specialisation" (p.315). Hu and Jaffe (2003), MacGarvie (2005), Picci (2010), and Montobbio and Sterzi (2012) are examples of studies that show similar results. Because technologically proximate countries are likely to be geographically close, failing to include such a technological proximity variable may lead to overestimating the effect of geography.

Whether countries share a common language or past colonial relationships is expected to: a) facilitate communication and b) represent similar cultural values between the two countries. Hence these variables are 'attracting' variables, expected to facilitate bilateral knowledge flows. Sharing a common official language indeed has the expected positive and significant effect on bilateral knowledge flows, whereas having a past colonial relationship has a negative effect.

[59] Although we must interpret the magnitude of this effect with caution. Firstly, this variable is a constructed measure, and secondly, we do not take the natural logarithm of this variable; therefore it does not have the same elasticity interpretation.

The strength of intellectual property rights of both the reporting, and partner country show to have a negative effect on the number of patent citations. However, only the recipient developing country's IPR is significantly different to zero. This finding does not have any intuitive interpretation. (While one might expect the relationship to be positive, it is worth noting that our models control for the inventive stocks in each country, which are positively correlated with the country's IPR.) Nevertheless, as we have found, the strength of intellectual property rights appear to matter for the estimated impact of FDI on knowledge flows.

As expected, there is evidence that traded goods diffuse valuable technological knowledge. The amount of imports the developing country receives from a developed country has a positive and significant effect on how much it learns from the exporting country. Developing countries seem to be able to utilise the embedded knowledge in imported goods to stimulate their own innovations of economically valuable products. However, in this model, this effect is moderate. Increasing imports by 10% leads to a 0.6% increase in patent citations.

8.2. Negative binomial and Poisson FE models

We now compare the Poisson model results to the negative binomial and the Poisson FE models. The regression estimates are reported in Tables 8.2 and 8.3, respectively.

Table 8.2 Negative binomial estimates for the developing sample

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.853**	0.853**	0.871**	0.873**		
	(0.062)	(0.075)	(0.061)	(0.070)		
ln $(P_{it} + 1)$					0.915**	0.936**
					(0.062)	(0.071)
ln P_{jt}	0.646**	0.628**	0.644**	0.613**	0.624**	0.593**
	(0.158)	(0.158)	(0.159)	(0.157)	(0.159)	(0.158)
ln GDP_{it}	-0.413	-0.401	-0.820*	-0.795*	-0.958**	-1.014**
	(0.331)	(0.373)	(0.329)	(0.346)	(0.327)	(0.344)
ln GDP_{jt}	0.395	0.367	0.243	0.364	0.285	0.421
	(0.576)	(0.582)	(0.564)	(0.564)	(0.569)	(0.569)
ln $Dist_{ij}$	-0.101*	-0.093†	-0.115*	-0.130*	-0.107†	-0.121*
	(0.051)	(0.053)	(0.055)	(0.059)	(0.055)	(0.059)
$Lang_{ij}$	0.108	0.077	0.155†	0.143	0.136	0.125
	(0.084)	(0.088)	(0.089)	(0.092)	(0.089)	(0.091)
$Colony_{ij}$	0.000	0.026	-0.057	-0.037	-0.041	-0.022
	(0.088)	(0.090)	(0.097)	(0.099)	(0.098)	(0.100)

$Prox_{ijt}$	1.945**	1.969**	1.794**	1.883**	1.970**	2.064**
	(0.236)	(0.241)	(0.234)	(0.237)	(0.231)	(0.234)
IPR_{it}	-0.176**	-0.085	-0.142**	-0.089	-0.140**	-0.089
	(0.049)	(0.157)	(0.050)	(0.119)	(0.051)	(0.118)
IPR_{jt}	-0.553**	-0.590**	-0.468**	-0.428*	-0.441*	-0.393*
	(0.176)	(0.181)	(0.180)	(0.184)	(0.181)	(0.184)
ln $Import_{ijt}$	0.034	0.058	0.010	0.013	0.006	0.010
	(0.045)	(0.046)	(0.046)	(0.049)	(0.047)	(0.049)
ln FDI_{ijt}	-0.005	0.022				
	(0.018)	(0.045)				
ln FDI_{ijt-1}	0.012	0.014				
	(0.017)	(0.022)				
ln FDI_{ijt}			0.016	0.028	0.017	0.024
(positive)			(0.017)	(0.032)	(0.017)	(0.032)
ln FDI_{ijt-1}			-0.007	0.020	-0.007	0.012
(positive)			(0.012)	(0.018)	(0.012)	(0.018)
ln FDI_{ijt}			0.032†	0.011	0.033†	0.008
(negative)			(0.018)	(0.047)	(0.018)	(0.047)
ln FDI_{ijt-1}			-0.020	0.051	-0.021	0.040
(negative)			(0.015)	(0.034)	(0.015)	(0.034)
ln FDI_{ijt}			0.643*	0.575†	0.631*	0.561†
(zero)			(0.279)	(0.295)	(0.274)	(0.288)
ln FDI_{ijt}* IPR_{it}		-0.009				
		(0.012)				
ln FDI_{ijt-1}* IPR_{it}		0.000				
		(0.005)				
ln FDI_{ijt}* IPR_{it}				-0.007		-0.005
(positive)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				-0.006		-0.003
(positive)				(0.005)		(0.005)
ln FDI_{ijt}* IPR_{it}				0.002		0.003
(negative)				(0.013)		(0.013)
ln FDI_{ijt-1}* IPR_{it}				-0.019*		-0.016†
(negative)				(0.009)		(0.009)
Observations	1,040	1,001	1,951	1,885	2,766	2,700
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Chi2 (model)	43758	48830	37207	35367	56029	52385
Log likelihood	-2838	-2720	-3949	-3772	-3969	-3793
AIC	5882	5649	8114	7768	8184	7839
BIC	6392	6165	8716	8389	8912	8588

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$, † $p<0.1$

When the NB model is used to estimate the same model specifications, the results do not differ much from the Poisson model. The notable differences are that past colonial relationships do not seem to have any effect in the NB model; intellectual property rights of the developed partner country have a negative and significant effect on patent citations; and the level of imports continues to have a small positive effect. However this effect is not statistically different to zero in the NB model. Turning to the variable of particular interest, FDI does not have any significant effect on bilateral patent citations in models (1) and (2) in Table 8.2. Similarly, for models (3) – (6), only the zero-value FDI has a significant effect among the FDI variables, having a positive effect. As far as we are aware, this has no intuitive interpretation.

When the model is estimated using the Poisson FE model there are few differences to the results (see Table 8.3 below). By using the full country pair fixed effects model, we can no longer estimate the effects of time-invariant variables; hence $Dist_{ij}$, $Lang_{ij}$, and $Colony_{ij}$ are dropped out of the model. The effects of FDI in the Poisson FE model are very similar to the Poisson model and therefore require no further explanation. The Poisson FE model estimates a much larger effect (almost double) of the patent stock of the source country on the number of patent citations. The size of the OECD country now has a positive and significant effect on the flow of knowledge, and this is consistent with a typical gravity mass term. This is an interesting finding because similar studies using a gravity framework to model bilateral knowledge flows (as discussed in our literature review) have not included GDP in the model. Our findings may provide some evidence that the size of the economy as well as the inventive or knowledge "mass" are important for bilateral knowledge flows. How close the country pair is in technological space does not have a significant effect in the Poisson FE model. This is explained by the fact that the proximity measure mostly varies between country pairs but varies little over time. There is no evidence that imports facilitate bilateral knowledge flow using this model.

Table 8.3 Poisson FE estimates for the developing sample

variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.904**	0.954**	0.929**	0.953**		
	(0.068)	(0.077)	(0.070)	(0.073)		
ln (P_{it} + 1)					0.932**	0.954**
					(0.071)	(0.073)
ln P_{jt}	1.020**	1.063**	1.089**	1.134**	1.087**	1.134**
	(0.217)	(0.293)	(0.212)	(0.254)	(0.212)	(0.254)
ln GDP_{it}	-0.240	-0.589*	-0.233	-0.553**	-0.235	-0.549**
	(0.202)	(0.230)	(0.197)	(0.206)	(0.197)	(0.206)
ln GDP_{jt}	1.362**	0.980*	1.497**	1.050*	1.497**	1.049*

	(0.528)	(0.493)	(0.525)	(0.454)	(0.524)	(0.454)
$Prox_{ijt}$	0.626†	0.704†	0.348	0.528	0.356	0.536
	(0.361)	(0.367)	(0.375)	(0.335)	(0.374)	(0.334)
IPR_{it}	-0.158**	-0.522**	-0.156**	-0.452**	-0.156**	-0.458**
	(0.038)	(0.160)	(0.039)	(0.114)	(0.039)	(0.115)
IPR_{jt}	0.124	0.128	0.029	0.081	0.030	0.082
	(0.133)	(0.135)	(0.112)	(0.106)	(0.112)	(0.106)
ln $Import_{ijt}$	-0.153*	0.036	-0.100	0.068	-0.100	0.067
	(0.071)	(0.085)	(0.077)	(0.081)	(0.077)	(0.081)
ln FDI_{ijt}	-0.039*	-0.149**				
	(0.020)	(0.037)				
ln FDI_{ijt-1}	0.039	0.008				
	(0.027)	(0.024)				
ln FDI_{ijt}			-0.042*	-0.137**	-0.042*	-0.139**
(positive)			(0.018)	(0.028)	(0.018)	(0.028)
ln FDI_{ijt-1}			-0.002	-0.014	-0.002	-0.013
(positive)			(0.015)	(0.020)	(0.015)	(0.020)
ln FDI_{ijt}			-0.024	-0.190**	-0.024	-0.191**
(negative)			(0.018)	(0.040)	(0.018)	(0.041)
ln FDI_{ijt-1}			-0.017	0.009	-0.017	0.008
(negative)			(0.015)	(0.029)	(0.015)	(0.029)
ln FDI_{ijt}			-0.377	-0.548*	-0.377	-0.546*
(zero)			(0.235)	(0.219)	(0.234)	(0.220)
ln FDI_{ijt}* IPR_{it}		0.028**				
		(0.011)				
ln FDI_{ijt-1}* IPR_{it}		0.009**				
		(0.003)				
ln FDI_{ijt}* IPR_{it}				0.023**		0.023**
(positive)				(0.008)		(0.008)
ln FDI_{ijt-1}* IPR_{it}				0.009*		0.009*
(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				0.038**		0.038**
(negative)				(0.008)		(0.008)
ln FDI_{ijt-1}* IPR_{it}				-0.002		-0.002
(negative)				(0.005)		(0.005)
Observations	707	664	1,118	1,056	1,152	1,092
Number of id	236	230	359	353	366	362
Country dummy	No	No	No	No	No	No
Year dummy	No	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes	Yes
Chi2 (model)	8497	14445	9259	20925	9282	20610
Log likelihood	-2740	-2338	-3589	-3080	-3602	-3096
AIC	5501	4699	7203	6194	7231	6226
BIC	5547	4753	7268	6278	7296	6311

Robust standard errors in parentheses

** $p<0.01$, * $p<0.05$, † $p<0.1$

8.3. Robustness Checks

In Section 7.2 we suggested that the number of patents applied for in the last five years (instead of ten years) also seemed to be a suitable measure for the knowledge "mass" of the source country. For comparison we ran all the above regressions replacing the P_{jt} variable with this five-year stock measure. There are no significant differences in the magnitude or statistical significance of the estimates. The regression results are reported in the Appendix A3, Tables A 3.1-3.3.

As we mention in Section 3.3, the patent data is severely affected by truncation. Towards the end of the sample there is a significant decrease in the number of patents in the data due to the time lag between applying for a patent and when it is granted. Although we control for time period fixed effects, for robustness we run the analysis excluding the last five-year period of the data to see if the truncation causes bias results. The set of results from the shortened sample did not differ substantially from the original sample. Therefore, we have also relegated the regression output to Appendix A3. Despite no alarming differences, on balance the results from the shortened sample are more highly statistically significant. For example, for the Poisson model, the lagged effect of FDI becomes highly significant. Also, sharing a common language is significant at least at the 5% level for all specifications, the same variable is significant at the 5% level only in specifications (1), (3), and (6) for the full length sample.

8.4. The OECD Sample Estimates

In the previous sections we examine the knowledge spillovers into developing countries, which is the main focus of this thesis. However, it is also interesting to look at the determinants of knowledge spillovers between developed countries. We do this by running the same six specifications with the sample including all OECD countries. For the sake of brevity we only present the results using the Poisson model. Furthermore, a key relationship we are interested in includes the effect of IPR on knowledge flow. Because this variable has little variation over time for developed countries, the Poisson FE model is not suitable to model this relationship.

Table 8.4 reports the estimates from the Poisson model, the columns represent the same specifications we have discussed in the previous tables.

Table 8.4 Poisson model for the OECD sample

variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.364**	0.307**	0.411**	0.347**		
	(0.070)	(0.075)	(0.069)	(0.071)		
ln (P_{it} + 1)					0.411**	0.347**
					(0.069)	(0.071)
ln P_{jt}	0.660**	0.626**	0.678**	0.650**	0.678**	0.650**
	(0.091)	(0.090)	(0.085)	(0.082)	(0.085)	(0.082)
ln GDP_{it}	1.712**	1.901**	1.726**	1.919**	1.726**	1.919**
	(0.256)	(0.272)	(0.263)	(0.272)	(0.263)	(0.272)
ln GDP_{jt}	0.667**	0.756**	0.814**	0.877**	0.814**	0.877**
	(0.192)	(0.196)	(0.188)	(0.188)	(0.188)	(0.188)
ln $Dist_{ij}$	-0.132**	-0.130**	-0.144**	-0.141**	-0.144**	-0.141**
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$Lang_{ij}$	0.075*	0.073*	0.080*	0.081*	0.080*	0.081*
	(0.035)	(0.035)	(0.034)	(0.033)	(0.034)	(0.033)
$Colony_{ij}$	0.295**	0.300**	0.288**	0.297**	0.288**	0.297**
	(0.029)	(0.029)	(0.030)	(0.030)	(0.030)	(0.030)
$Prox_{ijt}$	1.845**	1.947**	1.823**	1.916**	1.823**	1.916**
	(0.240)	(0.231)	(0.219)	(0.210)	(0.219)	(0.210)
IPR_{it}	-0.004	0.436**	-0.026	0.386*	-0.025	0.386*
	(0.035)	(0.165)	(0.035)	(0.155)	(0.035)	(0.155)
IPR_{jt}	0.063	0.048	0.048	0.035	0.048	0.035
	(0.039)	(0.038)	(0.037)	(0.036)	(0.037)	(0.036)
ln $Import_{ijt}$	-0.016	-0.021	-0.018	-0.025	-0.018	-0.025
	(0.025)	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)
ln FDI_{ijt}	0.016	0.168**				
	(0.012)	(0.058)				
ln FDI_{ijt-1}	0.025*	0.013				
	(0.012)	(0.017)				
ln FDI_{ijt}			0.010	0.154**	0.010	0.154**
(positive)			(0.011)	(0.055)	(0.011)	(0.055)
ln FDI_{ijt-1}			0.015	0.008	0.015	0.008
(positive)			(0.010)	(0.016)	(0.010)	(0.016)
ln FDI_{ijt}			0.006	0.167*	0.006	0.167*
(negative)			(0.012)	(0.073)	(0.012)	(0.073)
ln FDI_{ijt-1}			0.016	-0.003	0.016	-0.003
(negative)			(0.011)	(0.032)	(0.011)	(0.032)
ln FDI_{ijt}			0.070	0.179	0.070	0.179
(zero)			(0.143)	(0.155)	(0.143)	(0.155)
ln FDI_{ijt}* IPR_{it}		-0.033**				
		(0.012)				
ln FDI_{ijt-1}* IPR_{it}		0.003				
		(0.003)				
ln FDI_{ijt}* IPR_{it}				-0.031**		-0.031**
(positive)				(0.011)		(0.011)
ln FDI_{ijt-1}* IPR_{it}				0.003		0.003

(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				-0.035*		-0.035*
(negative)				(0.015)		(0.015)
ln FDI_{ijt-1}* IPR_{it}				0.005		0.005
(negative)				(0.007)		(0.007)
Observations	1,453	1,372	2,412	2,282	2,454	2,324
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Log likelihood	-48036	-46152	-61398	-58681	-61397	-58679
AIC	96219	92457	122952	117526	122950	117523
BIC	96610	92854	123404	117996	123402	117994

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

The first four coefficients in Table 8.4 shows that all "mass" terms have a large positive and significant effect on knowledge flows between developed countries. In contrast to the developing country sample, the larger the receiving country's economy is, the more inventors cite patents from the source country. This is in line with Montobbio and Sterzi's (2012) idea of the "demand effect". The larger the economy, the more demand there is for foreign knowledge from domestic firms, all else equal.[60] The estimates on the distance, language and technological proximity variables are consistent with the developing country sample results. Past colonial relationships seem to matter for knowledge flows between developing countries; the effect is positive and significant. The effect of FDI inflows on bilateral knowledge diffusion is also different for this sample. Column (1) shows that lagged FDI flows into a developed country have a positive impact on the number of citations to patents of the investing country. Also the calculated elasticity of FDI (see equation 8.2) from column (2) indicates a positive contemporaneous effect. The alternative specifications in (4) and (6) confirm this effect on both positive and negative FDI flows. Also, in contrast to the developing sample, the interaction term between FDI and IPR has a negative effect on patent citations. For developed countries it appears that strengthening IPRs decreases the amount of knowledge flows through FDI. This finding is consistent with Montobbio and Sterzi (2012) who found an inverse "U" shape effect of IPR on bilateral knowledge flows. Strengthening IPRs facilitates knowledge flows through FDI only if the country starts

[60] The demand effect must outweigh the "supply effect" we discussed for the developing country sample.

from a weak level of IPR (i.e. developing countries). Increasing IPRs for countries with already strong levels of IPRs (i.e. developed countries) has a negative effect on knowledge spillovers from FDI due to monopoly power and higher costs to access the technological knowledge of the foreign subsidiaries.

9. CONCLUSION

Our key research question has been to examine whether foreign direct investment facilitates spillover of technological knowledge into the host country. We also investigate whether the strength of intellectual property rights plays an important role in mediating bilateral knowledge spillovers through FDI.

Our empirical analysis focuses on knowledge spillovers into developing countries given that the diffusion of technology plays such a significant role in determining the long run growth of an economy and income convergence. We find that FDI has a negative effect on knowledge spillovers if we consider the impact of aggregated bilateral FDI flows. We suggest this may be reflecting a crowding out effect of the potentially large portion of FDI that rely on less advanced technology in developing countries. To disentangle the negative effect of aggregate FDI we consider whether the strength of the intellectual property rights in the host country impacts the effect of FDI on knowledge spillovers. Our findings suggest that developing countries with relatively strong intellectual property rights may indeed receive positive knowledge spillovers from FDI; however this effect tends occur only with a lag. We interpret this finding as evidence that strong IPR regimes attract technologically advanced FDI that is more conducive to knowledge spillovers.

In contrast, we provide evidence that aggregate FDI has a positive effect on knowledge flows between developed countries. Interestingly, we find that an increase in IPR in developed countries has a negative effect on bilateral knowledge flows. This suggests that there may be a threshold IPR that must reach in order to induce positive knowledge spillover through FDI. Beyond that threshold, IPR has negative effects, perhaps due to an increase in monopolistic power.

Furthermore, we provide evidence in support of the literature that knowledge flows are geographically and technologically localised. The distance between two countries has a negative effect on the amount of technological knowledge that flows between them. On the other hand, countries that share similar technological specialisations are more likely to utilise technology from one another.

Our results raise interesting questions for country level policy makers. Policies designed to attract FDI may not be justified by positive externalities if the country's IPR regime is not sufficiently strong. We provide evidence that it is not the magnitude of FDI that produces knowledge spillovers. Rather it is more likely that only technologically advanced FDI, which is more sensitive to IPR, contributes to significant knowledge spillovers. Our results also suggest that whether or not increasing the strength of IPR attracts knowledge spillover via FDI is sensitive to the initial level of IPR. An area of future research may be to further analyses the strength of intellectual property rights in developing countries and investigate the effect of FDI on knowledge flows over time. Future research using patent citations as a measure of knowledge flows may also benefit from considering the economic value of the knowledge flow. For example, other areas of research have analysed the value of patents using forward patent citations.

If we can better understand the relationship between intellectual property rights, foreign direct investment, and valuable knowledge spillovers, policy makers may direct resources towards more suitable incentives to attract multinational firms who generate growth stimulating knowledge spillovers.

REFERENCES

Aitken, B. J. & Harrison, A. E. (1999) Do domestic firms benefit from direct foreign investment? *American Economic Review,* 89**,** 605-618.

Alcácer, J., Gittelman, M. & Sampat, B. (2009) Applicant and examiner citations in U.S. patents: An overview and analysis. *Research Policy,* 38**,** 415-427.

Almeida, P. (1996) Knowldge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry. *Strategic Management Journal,* 17**,** 155-165.

Anderson, J. E. (1979) A theoretical foundation for the gravity equation. *American Economic Review,* 69**,** 106-116.

Anderson, J. E. (2011) The gravity model. *Annual Review of Economics,* 3**,** 133-160.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. *The rate and direction of inventive activity: Economic and social factors.* NBER.

Benner, M. & Waldfogel, J. (2008) Close to you? Bias and precision in patent-based measures of technological proximity. *Research Policy,* 37**,** 1556-1567.

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

Borensztein, E., De Gregorio, J. & Lee, J. W. (1998) How does foreign direct investment affect economic growth? *Journal of International Economics,* 45**,** 115-135.

Bottazzi, L. & Peri, G. (2003) Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review,* 47**,** 687-710.

Branstetter, L. (1998) Looking for international knowledge spillovers a review of the literature with suggestions for new approaches. *Annales d'Economie et de Statistique,* 49/50, 517-540.

Branstetter, L. (2001) Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the US and Japan. *Journal of International Economics,* 53**,** 53-79.

Branstetter, L. (2006) Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States. *Journal of International Economics,* 68**,** 325-344.

Branstetter, L., Fisman, R., Foley, C. F. & Saggi, K. (2011) Does intellectual property rights reform spur industrial development? *Journal of International Economics,* 83**,** 27-36.

Breschi, S. & Lissoni, F. (2005). Knowledge networks from patent data. *Handbook of Quantitative Science and Technology Research.* Springer Netherlands.

Burger, M., Oort, F. V. & Linders, G.-J. (2009) On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation. *Spatial Economic Analysis,* 4**,** 167-190.

Cameron, A. C. & Trivedi, P. K. (2005) *Microeconometrics: methods and applications,* Cambridge University Press, Cambridge.

Coe, D. T. & Helpman, E. (1995) International R&D spillovers. *European Economic Review,* 39**,** 859-887.

Coe, D. T., Helpman, E. & Hoffmaister, A. W. (2009) International R&D spillovers and institutions. *European Economic Review,* 53**,** 723-741.

Coe, D. T. & Hoffmaister, A. W. (1999) Are there international R&D spillovers among randomly matched trade partners?-A response to Keller. *IMF Working Papers.*

Criscuolo, P. (2006) The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO. *Scientometrics,* 66**,** 23-41.

Criscuolo, P. & Verspagen, B. (2008) Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy,* 37**,** 1892-1908.

Dernis, H. & Khan, M. (2004). Triadic patent families methodology. OECD Publishing.

Disdier, A. C. & Head, K. (2008) The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics,* 90**,** 37-48.

Döring, T. & Schnellenbach, J. (2006) What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature. *Regional Studies,* 40**,** 375-395.

Duguet, E. & Macgarvie, M. (2005) How well do patent citations measure flows of technology? Evidence from French innovation surveys. *Economics of Innovation and New Technology,* 14**,** 375-393.

Eaton, J. & Kortum, S. (1999) International technology diffusion: Theory and measurement. *International Economic Review,* 40**,** 537-569.

Flowerdew, R. & Aitkin, M. (1982) A method of fitting the gravity model based on the Poisson distribution. *Journal of Regional Science,* 22**,** 191-202.

Frankel, J. A., Stein, E. & Wei, S. J. (1997) *Regional trading blocs in the world economic system*, Peterson Institute.

Genç, M. (2013) Migration and tourism flows to New Zealand. *Quantitative Methods in Tourism Economics*, Á. Matias, P. Nijkamp, & M. Sarmento (Eds.), *forthcoming.*

Ginarte, J. C. & Park, W. G. (1997) Determinants of patent rights: A cross-national study. *Research Policy,* 26**,** 283-301.

Glaeser, E. L. & Kohlhase, J. E. (2003) Cities, regions and the decline of transport costs. *Papers in Regional Science,* 83**,** 197-228.

Görg, H. & Greenaway, D. (2004) Much ado about nothing? Do domestic firms really benefit from foreign direct investment? *The World Bank Research Observer,* 19**,** 171-197.

Griliches, Z. (1992). The search for R&D spillovers. *The Scandinavian Journal of Economics*, 94, 29-47.

Grossman, G. M. & Helpman, E. (1991) *Innovation and growth in the global economy*, MIT Press, Cambridge, MA.

Haddad, M. & Harrison, A. (1993) Are there positive spillovers from direct foreign investment? 1:: Evidence from panel data for Morocco. *Journal of Development Economics,* 42**,** 51-74.

Hall, B. H., Jaffe, A. B. & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. National Bureau of Economic Research, working paper no.8498.

Haworth, J. & Vincent, P. (1979) The stochastic disturbance specification and its implications for log-linear regression. *Environment and Planning A,* 11**,** 781-790.

Hu, A. G. Z. & Jaffe, A. B. (2003) Patent citations and international knowledge flow: the cases of Korea and Taiwan. *International Journal of Industrial Organization,* 21**,** 849-880.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value. National Bureau of Economic Research, working paper no.1815.

Jaffe, A. B. & Trajtenberg, M. (1999) International knowledge flows: Evidence from patent citations. *Economics of Innovation and New Technology,* 8**,** 105-136.

Jaffe, A. B. & Trajtenberg, M. (2002) *Patents, Citations, and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA.

Jaffe, A. B., Trajtenberg, M. & Fogarty, M. S. (2000) Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review,* 90**,** 215-218.

Jaffe, A. B., Trajtenberg, M. & Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108**,** 577-598.

Karemera, D., Oguledo, V. I. & Davis, B. (2000) A gravity model analysis of international migration to North America. *Applied Economics*, 32**,** 1745-1755.

Keller, W. (1998) Are international R&D spillovers trade-related?: Analyzing spillovers among randomly matched trade partners. *European Economic Review,* 42**,** 1469-1481.

Keller, W. (2002a) Trade and the transmission of technology. *Journal of Economic Growth,* 7**,** 5-24.

Keller, W. (2002b) Geographic localization of international technology diffusion. *American Economic Review,* 92**,** 120-142.

Keller, W. (2004) International Technology Diffusion. *Journal of Economic Literature,* 42**,** 752-782.

Keller, W. (2009). International trade, foreign direct investment, and technology spillovers. National Bureau of Economic Research, working paper no.15442.

King, G. (1988) Statistical models for political science event counts: Bias in conventional procedures and evidence for the exponential Poisson regression model. *American Journal of Political Science*, 32(3), 838-863.

Krugman, P. (1992) *Geography and trade*, MIT press.

Lai, R., D'amour, A., Yu, A., Sun, Y., Torvik, V. & Fleming, L. (2011) Disambiguation and co-authorship networks of the US Patent Inventor Database. *Harvard Institute for Quantitative Social Science, Cambridge, MA,* 2138.

Li, Y. (2009) Borders and distance in knowledge spillovers: Dying over time or dying with age?-Evidence from patent citations. *CESifo Working Paper Sereis*, no. 2625.

Linders, G. & De Groot, H. (2006) Estimation of the gravity equation in the presence of zero flows. *Tinbergen Institute*, Discussion Paper 2006-072/3.

MacGarvie, M. (2005) The determinants of international knowledge diffusion as measured by patent citations. *Economics Letters*, 87**,** 121-126.

MacGarvie, M. (2006) Do firms learn from international trade? *Review of Economics and Statistics*, 88**,** 46-60.

Maurseth, P. B. & Verspagen, B. (2002) Knowledge spillovers in Europe: a patent citations analysis. *The Scandinavian journal of economics*, 104**,** 531-545.

Montobbio, F. & Sterzi, V. (2012). The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations. Groupe de Recherche en Economie Théorique et Appliquée.

OECD (1999). Benchmark definition of foreign direct investment. *Third edition*

Park, W. G. (2008) International patent protection: 1960–2005. *Research Policy,* 37**,** 761-766.

Park, W. G. & Ginarte, J. C. (1997) Intellectual property rights and economic growth. *Contemporary Economic Policy,* 15**,** 51-61.

Peri, G. (2002) Knowledge flows and knowledge externalities. *CESifo Working Paper Series*, no.765.

Peri, G. (2005) Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87**,** 308-322.

Picci, L. (2010) The internationalization of inventive activity: A gravity model using patent data. *Research Policy,* 39**,** 1070-1081.

Ramsey, J. B. (1969) Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B***,** 350-371.

Romer, P. M. (1986) Increasing returns and long-run growth. *The Journal of Political Economy***,** 94, 1002-1037.

Santos Silva, J. M. C. S. & Tenreyro, S. (2006) The log of gravity. *Review of Economics & Statistics*, 88**,** 641-658.

Santos Silva, J. M. C. & Tenreyro, S. (2011) Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112**,** 220-222.

Sapra, S. (2005) A regression error specification test (RESET) for generalized linear models. *Economics Bulletin,* 3**,** 1-6.

Schumpeter, J. (1934). *The theory of economic development*. Harvard University Press, Cambridge, MA1934.

Sin, I. (2012) The Gravity of Ideas: How distance affects translations. *Motu Economic and Public Policy Research*.

Singh, J. (2004) Multinational firms and knowledge diffusion: evidence using patent citation data, Unpublished manuscript.

Sjöholm, F. (1996) International transfer of knowledge: the role of international trade and geographic proximity. *Review of World Economics*, 132**,** 97-115.

Smeets, R. (2008) Collecting the pieces of the FDI knowledge spillovers puzzle. *The World Bank Research Observer,* 23**,** 107-138.

Stephan, J. (2011) Foreign direct investment in weak intellectual property rights regimes – the example of post-socialist economies. *Post-Communist Economies*, 23**,** 35-53.

Thompson, P. (2006) Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. *Review of Economics and Statistics*, 88**,** 383-388.

Tinbergen, J. (1962) *Shaping the world economy: Suggestions for an international economic policy*, Twentieth Century Fund, New York.

Trajtenberg, M. (1990) A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics*, 21**,** 172-187.

WIPO (2011). *World Intellectual Property Indicators*. Economics & Statistics Series.

Wooldridge, J. M. (1997) Quasi-likelihood methods for count data. *Handbook of applied econometrics*, 2**,** M.H. Pesaran and P. Schmidt (Eds.). Oxford: Blackwell, 352-406.

Wooldridge, J. M. (2002) *Econometric analysis of cross-section and panel data*, MIT Press, Cambridge, MA.

Zhu, L. & Jeon, B. N. (2007) International R&D spillovers: Trade, FDI, and information technology as spillover channels. *Review of International Economics*, 15**,** 955-976.

A1. Data Appendix

Table A 1.1 Annual number of patents granted by application year

Developing Countries

year	Total	Average	Median	Min	Max
1975	962	6	0	0	395
1980	867	6	0	0	246
1985	982	7	0	0	253
1990	2,069	14	0	0	936
1995	4,132	27	0	0	2,258
2000	9,534	63	1	0	5,050
2005	11,746	78	0	0	6,844
2009	1,439	10	0	0	926

OECD member Countries

year	Total	Average	Median	Min	Max
1975	64,916	2,164	104	0	42,148
1980	65,604	2,187	125	0	38,833
1985	70,464	2,349	142	0	37,697
1990	97,285	3,243	284	0	53,329
1995	140,489	4,683	456	0	81,935
2000	187,050	6,235	609	0	102,983
2005	154,410	5,147	401	0	82,319
2009	10,679	356	17	0	6,152

All Countries excl U.S.

year	Total	Average	Median	Min	Max
1975	23,730	132	0	0	6,077
1980	27,638	154	0	0	9,574
1985	33,749	187	0	0	14,340
1990	46,025	256	0	0	22,126
1995	62,686	348	0	0	29,066
2000	93,601	520	1	0	39,525
2005	83,837	466	1	0	37,657
2009	5966	33	0	0	2,567

All Countries

year	Total	Average	Median	Min	Max
1975	65,878	364	0	0	42,148
1980	66,471	367	0	0	38,833
1985	71,446	395	0	0	37,697
1990	99,354	549	0	0	53,329
1995	144,621	799	0	0	81,935
2000	196,584	1,086	1	0	102,983
2005	166,156	918	1	0	82,319
2009	12118	67	0	0	6,152

Table A 1.2 Annual number of bilateral patent citations by application year

Developing Countries

year	Total	Average	Median	Min	Max
1975	263	0.1	0	0	47
1980	1,218	0.4	0	0	189
1985	2,756	1.0	0	0	384
1990	7,140	2.6	0	0	1,655
1995	21,573	7.8	0	0	5,485
2000	64,381	23.3	0	0	14,277
2005	70,883	25.7	0	0	18,555
2009	359	0.1	0	0	123

OECD member Countries

year	Total	Average	Median	Min	Max
1975	13,328	15.3	0	0	1,714
1980	66,043	75.9	0	0	9,812
1985	121,124	139.2	0	0	20,782
1990	224,586	258.1	1	0	39,516
1995	483,197	555.4	1	0	110,208
2000	827,732	951.4	3	0	173,762
2005	605,523	696.0	2	0	138,385
2009	2,441	2.8	0	0	721

All Countries excl U.S.

year	Total	Average	Median	Min	Max
1975	8,803	2.4	0	0	1,714
1980	42,847	11.9	0	0	9,812
1985	79,814	22.2	0	0	20,782
1990	135,949	37.8	0	0	36,642
1995	256,920	71.3	0	0	64,977
2000	493,139	136.9	0	0	100,094
2005	354,895	98.6	0	0	68,374
2009	1,232	0.3	0	0	201

All Countries

year	Total	Average	Median	Min	Max
1975	13,591	3.7	0	0	1,714
1980	67,261	18.5	0	0	9,812
1985	123,880	34.1	0	0	20,782
1990	231,726	63.8	0	0	39,516
1995	504,770	139.1	0	0	110,208
2000	892,113	245.8	0	0	173,762
2005	676,406	186.3	0	0	138,385
2009	2,800	0.8	0	0	721

USPTO patent example (extract)

United States Patent	**6,746,448**
Weiner , et al.	**June 8, 2004**

Outrigger for bone fixator

Inventors: Weiner; Lon S. (Rumson, NJ)**, Coull; Thomas** (Rancho Palos Verdes, CA)
Assignee: Millennium Medical Technologies, Inc. (Santa Fe, NM)
Appl. No.: 10/233,897
Filed: September 3, 2002

Related U.S. Patent Documents

Application Number	Filing Date	Patent Number	Issue Date
160470	May., 2002		

Current U.S. Class:	**606/54** ; 606/55; 606/59
Current International Class:	A61B 17/60 (20060101); A61B 17/64 (20060101); A61B 017/56 ()
Field of Search:	606/54,55,59

References Cited

U.S. Patent Documents

2333033	October 1943	Mraz
2393694	January 1946	Kirschner
4548199	October 1985	Agee
4611586	September 1986	Agee et al.
4628919	December 1986	Clyburn
4662365	May 1987	Gotzen et al.
4730608	March 1988	Schlein
4848327	July 1989	Perdue
4889111	December 1989	Ben-Dov
4922896	May 1990	Agee et al.
5087258	February 1992	Schewior
5122140	June 1992	Asche et al.
5139500	August 1992	Schwartz
5527309	June 1996	Shelton

Table A 1.3 Countries included in sample

OECD	Developing countries		
Australia	Algeria	Guatemala	Paraguay
Austria	Angola	Guyana	Peru
Belgium	Argentina	Haiti	Philippines
Canada	Bangladesh	Honduras	Romania
Chile	Benin	India	Russian Federation
Czech Rep.	Bolivia	Indonesia	Rwanda
Denmark	Botswana	Iran	Saudi Arabia
Finland	Brazil	Iraq	Senegal
France	Bulgaria	Israel	Sierra Leone
Germany	Burkina Faso	Jamaica	Singapore
Greece	Burundi	Jordan	Somalia
Hungary	Cameroon	Kenya	South Africa
Iceland	Central African Rep.	Liberia	Sri Lanka
Ireland	Chad	Lithuania	Sudan
Italy	China	Madagascar	Swaziland
Japan	China, Hong Kong	Malawi	Syria
Luxembourg	Colombia	Malaysia	Taiwan
Mexico	Congo	Mali	Thailand
Netherlands	Costa Rica	Malta	Togo
New Zealand	Côte d'Ivoire	Mauritania	Trinidad and Tobago
Norway	Cyprus	Mauritius	Tunisia
Poland	Dem. Rep. of the Congo	Morocco	Uganda
Portugal	Dominican Rep.	Mozambique	Ukraine
Rep. of Korea	Ecuador	Myanmar	United Rep. of Tanzania
Slovakia	Egypt	Nepal	Uruguay
Spain	El Salvador	Nicaragua	Venezuela
Sweden	Ethiopia	Niger	Viet Nam
Switzerland	Fiji	Nigeria	Zambia
Turkey	Gabon	Pakistan	Zimbabwe
United Kingdom	Ghana	Panama	
USA	Grenada	Papua New Guinea	

Table A 1.4 Definition and data source of explanatory variables

Variable name	Definition and source
P_{it}	Number of patents from country *i* applied for at USPTO in period *t*. Source Harvard Patent data set.[61]
P_{jt}	Measure of USPTO patent stock of country *j* at period *t* Source Harvard Patent data set.
$Dist_{ij}$	Km, simple distance which uses latitudes and longitudes of the most populated cities. Source: CEPII data set
FDI_{ijt}	Bilateral Foreign Direct Investment inflows to country *i* from country *j*. Where inflows to country *i* are unavailable, outflows from *j* to *i* are used. Millions of U.S. dollars. Source OECD international direct investment database
GDP_{it}	Millions of constant U.S. dollars (year 2000 prices). Source World Bank
GDP_{jt}	Millions of constant U.S. dollars (year 2000 prices). Source World Bank
$Import_{ijt}$	Imports of country *i* from *j*. Source UN Comtrade using SITC Rev.1 classification
IRR_{it}	Ginarte and Park Intellectual Property Rights strength index. Park and Ginarte (1997), Park (2008)
IRR_{jt}	Ginarte and Park Intellectual Property Rights strength index. Park and Ginarte (1997), Park (2008)
$Lang_{ij}$	Dummy variable equal to 1 if the country pair shares a common official language. Source CEPII dataset.
$Colony_{ij}$	Dummy variable equal to 1 if the country pair has ever had a colonial relationship. Source CEPII data set.
$Prox_{ijt}$	The technological proximity of the *ij* country-pair. Calculated according to Section 5. Source Harvard Patent data set

[61] See Lai et al. (2011)

Table A 1.5 Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Developing Countries Sample					
year				1975	2010
C_{ijt}	19,320	44	1,136	0	101,787
P_{it}	22,080	212	1,705	0	34,146
P_{ji}	22,080	15,145	56,384	0	479,319
$Dist_{ij}$	21,840	7,810	3,934	394	19,630
FDI_{ijt}	5,240	81,207	456,254	3,236,639	11,800,000
GDP_{it}	20,220	60,043	187,820	114	3,246,008
GDP_{jt}	20,976	729,206	1,659,834	3,807	11,500,000
$Import_{ijt}$	15,391	335	2,388	0	141,580
IPR_{it}	18,240	2.40	0.85	0	4.67
IPR_{jt}	17,112	3.56	1.00	0	4.88
$Prox_{ijt}$	11,331	0.55	0.22	0	1
$Lang_{ij}$	21,840	0.11	0.31	0	1
$Colony_{ij}$	21,840	0.03	0.17	0	1
OECD countries Sample					
year				1975	2010
C_{ijt}	6090	1,849	19,915	0	865,077
P_{it}	6960	15,145	56,387	0	479,319
P_{ji}	6960	15,145	56,387	0	479,319
$Dist_{ij}$	6960	5,975	5,457	60	19,335
FDI_{ijt}	3,552	590,000	2,627,688	5,749,274	56,200,000
GDP_{it}	6,612	729,206	1,659,920	3,807	11,500,000
GDP_{jt}	6,612	729,206	1,659,920	3,807	11,500,000
$Import_{ijt}$	5,533	2,901	11,585	0	291,560
IPR_{it}	5,394	4	1	0	5
IPR_{jt}	5,394	3.56	1.00	0	4.88
$Prox_{ijt}$	5,574	0.76	0.17	0	1
$Lang_{ij}$	6,960	0.07	0.26	0	1
$Colony_{ij}$	6,960	0.03	0.18	0	1

A2. Specification Testing Appendix

Table A 2.1 FDI lag structure; developing sample Poisson estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	1.007**	0.980**	1.023**	1.022**	1.038**	1.011**	0.986**	0.996**
	(0.042)	(0.045)	(0.046)	(0.047)	(0.045)	(0.045)	(0.045)	(0.047)
ln P_{jt}	0.203†	0.154	0.169	0.178	0.127	0.174	0.119	0.106
	(0.112)	(0.120)	(0.133)	(0.137)	(0.142)	(0.155)	(0.165)	(0.173)
ln GDP_{it}	-0.749**	-0.548*	-0.754**	-0.754**	-0.670**	-0.553*	-0.383	-0.490
	(0.246)	(0.258)	(0.260)	(0.276)	(0.257)	(0.280)	(0.289)	(0.311)
ln GDP_{jt}	0.833*	0.728*	0.672†	0.648†	0.698†	0.381	0.515	0.492
	(0.341)	(0.361)	(0.361)	(0.380)	(0.398)	(0.437)	(0.451)	(0.471)
ln $Dist_{ij}$	-0.129**	-0.136**	-0.132**	-0.130**	-0.142**	-0.127**	-0.129**	-0.118**
	(0.037)	(0.038)	(0.040)	(0.040)	(0.040)	(0.041)	(0.043)	(0.045)
$Lang_{ij}$	0.154**	0.185**	0.161**	0.147**	0.157**	0.178**	0.186**	0.221**
	(0.054)	(0.056)	(0.057)	(0.056)	(0.057)	(0.061)	(0.061)	(0.065)
$Colony_{ij}$	-0.183**	-0.158**	-0.150**	-0.126*	-0.131*	-0.128*	-0.126*	-0.121*
	(0.053)	(0.053)	(0.055)	(0.055)	(0.051)	(0.052)	(0.052)	(0.053)
$Prox_{ijt}$	1.621**	1.597**	1.655**	1.635**	1.650**	1.671**	1.655**	1.640**
	(0.169)	(0.171)	(0.176)	(0.183)	(0.183)	(0.189)	(0.193)	(0.201)
IPR_{it}	-0.281**	-0.287**	-0.301**	-0.315**	-0.318**	-0.311**	-0.305**	-0.306**
	(0.046)	(0.047)	(0.047)	(0.048)	(0.043)	(0.041)	(0.042)	(0.044)
IPR_{jt}	-0.536**	-0.625**	-0.639**	-0.742**	-0.706**	-0.818**	-0.755*	-0.689*
	(0.169)	(0.182)	(0.190)	(0.211)	(0.237)	(0.283)	(0.304)	(0.325)
ln $Import_{ijt}$	0.129**	0.137**	0.136**	0.148**	0.131**	0.125**	0.124**	0.139**
	(0.032)	(0.034)	(0.036)	(0.036)	(0.037)	(0.037)	(0.037)	(0.038)
ln FDI_{ijt}	-0.031*	-0.031*	-0.037*	-0.035*	-0.043**	-0.040**	-0.044**	-0.047**
	(0.014)	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
ln FDI_{ijt-1}		-0.000†	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*	-0.000*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			0.000	0.000	0.000	0.000*	0.000*	0.000*
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				-0.000	-0.000*	-0.000**	-0.000**	-0.000**
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					0.000**	0.000**	0.000**	0.000**
					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						0.000**	0.000**	0.000**
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							-0.000	-0.000
							(0.000)	(0.000)

ln FDI_{ijt-7}								0.000
								(0.000)
Obs	4,202	3,694	3,312	2,975	2,647	2,330	2,057	1,837
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No	No	No
R^2	0.974	0.975	0.975	0.976	0.977	0.977	0.978	0.978
Log likelihood	-11862	-10897	-10007	-9426	-8703	-8132	-7503	-6978
AIC	23959	22025	20247	19079	17624	16481	15223	14160
BIC	24708	22739	20955	19763	18265	17109	15837	14723

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 2.2 FDI lag structure; developing sample Poisson FE estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	1.115**	1.109**	1.114**	1.120**	1.138**	1.107**	1.110**	1.117**
	(0.075)	(0.069)	(0.070)	(0.065)	(0.068)	(0.058)	(0.060)	(0.064)
ln P_{jt}	0.495**	0.507**	0.402**	0.428**	0.432**	0.478**	0.464**	0.447**
	(0.125)	(0.136)	(0.133)	(0.141)	(0.118)	(0.123)	(0.129)	(0.136)
ln GDP_{it}	-0.507†	-0.412	-0.390	-0.333	-0.299	-0.061	-0.041	-0.072
	(0.278)	(0.286)	(0.293)	(0.288)	(0.272)	(0.330)	(0.329)	(0.348)
ln GDP_{jt}	2.071**	2.141**	2.282**	2.479**	2.361**	1.858**	1.986**	2.139**
	(0.449)	(0.425)	(0.357)	(0.382)	(0.325)	(0.412)	(0.411)	(0.510)
$Prox_{ijt}$	0.557*	0.538*	0.591*	0.560†	0.574†	0.540†	0.545†	0.505
	(0.254)	(0.268)	(0.295)	(0.306)	(0.300)	(0.317)	(0.320)	(0.330)
IPR_{it}	-0.206**	-0.222**	-0.245**	-0.278**	-0.278**	-0.255**	-0.261**	-0.275**
	(0.063)	(0.065)	(0.066)	(0.075)	(0.066)	(0.057)	(0.058)	(0.063)
IPR_{jt}	0.334†	0.301	0.374†	0.299	0.287	0.221	0.208	0.207
	(0.191)	(0.205)	(0.209)	(0.221)	(0.197)	(0.222)	(0.238)	(0.298)
ln $Import_{ijt}$	-0.067	-0.078	-0.086	-0.066	-0.076	-0.118†	-0.125†	-0.114
	(0.064)	(0.070)	(0.068)	(0.065)	(0.065)	(0.067)	(0.070)	(0.076)
ln FDI_{ijt}	-0.026*	-0.026†	-0.035**	-0.036**	-0.044**	-0.031*	-0.037**	-0.040**
	(0.013)	(0.014)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)
ln FDI_{ijt-1}		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			0.000	0.000	0.000**	0.000**	0.000**	0.000**
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				-0.000*	-0.000**	-0.000**	-0.000**	-0.000**
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					0.000**	0.000**	0.000**	0.000**

					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						0.000**	0.000**	0.000**
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							-0.000	-0.000
							(0.000)	(0.000)
ln FDI_{ijt-7}								-0.000
								(0.000)
Obs	3,310	2,950	2,642	2,392	2,154	1,956	1,758	1,592
Number of id	371	341	322	291	263	247	224	206
Country dummy	No	No	No	No	No	No	No	No
Year dummy	No	No	No	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chi2 (model)	4490	5222	4958	4713	6969	7207	7399	7385
Log likelihood	-9594	-8850	-7998	-7502	-6801	-6333	-5873	-5454
AIC	19205	17721	16019	15028	13628	12693	11776	10941
BIC	19260	17780	16083	15097	13702	12771	11858	11026

Table A 2.3 FDI lag structure; developing sample negative binomial estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	0.884**	0.898**	0.923**	0.906**	0.925**	0.918**	0.935**	0.951**
	(0.035)	(0.037)	(0.039)	(0.041)	(0.046)	(0.051)	(0.050)	(0.053)
ln P_{jt}	0.413**	0.362**	0.340**	0.359**	0.280*	0.208†	0.170	0.111
	(0.091)	(0.095)	(0.105)	(0.112)	(0.117)	(0.125)	(0.131)	(0.144)
ln GDP_{it}	-0.516*	-0.537*	-0.572*	-0.526*	-0.629*	-0.697*	-0.636*	-0.731*
	(0.216)	(0.227)	(0.242)	(0.255)	(0.276)	(0.307)	(0.304)	(0.338)
ln GDP_{jt}	1.252**	1.231**	1.177**	1.189**	1.353**	1.270**	1.630**	1.844**
	(0.316)	(0.329)	(0.345)	(0.366)	(0.386)	(0.416)	(0.426)	(0.462)
ln $Dist_{ij}$	-0.068†	-0.077*	-0.073†	-0.073†	-0.087*	-0.105*	-0.133**	-0.136**
	(0.036)	(0.037)	(0.038)	(0.039)	(0.040)	(0.041)	(0.041)	(0.043)
$Lang_{ij}$	0.015	0.012	-0.009	0.026	0.028	0.023	0.044	0.073
	(0.061)	(0.063)	(0.065)	(0.067)	(0.069)	(0.073)	(0.074)	(0.075)
$Colony_{ij}$	-0.018	0.006	-0.002	-0.006	-0.042	-0.085	-0.111	-0.113
	(0.065)	(0.065)	(0.064)	(0.066)	(0.067)	(0.066)	(0.069)	(0.071)
$Prox_{ijt}$	1.775**	1.736**	1.619**	1.643**	1.686**	1.693**	1.653**	1.561**
	(0.136)	(0.139)	(0.145)	(0.148)	(0.156)	(0.164)	(0.165)	(0.171)
IPR_{it}	-0.123**	-0.122**	-0.141**	-0.152**	-0.144**	-0.159**	-0.134**	-0.128*
	(0.045)	(0.046)	(0.047)	(0.048)	(0.049)	(0.050)	(0.049)	(0.052)

IPR_{jt}	-0.231†	-0.212	-0.315*	-0.385*	-0.339†	-0.412†	-0.275	-0.158
	(0.125)	(0.138)	(0.147)	(0.174)	(0.196)	(0.217)	(0.225)	(0.250)
ln $Import_{ijt}$	0.077**	0.088**	0.105**	0.115**	0.106**	0.100**	0.091**	0.108**
	(0.028)	(0.029)	(0.031)	(0.032)	(0.033)	(0.035)	(0.035)	(0.036)
ln FDI_{ijt}	0.005	0.008	0.012	0.007	0.004	0.002	-0.003	-0.007
	(0.011)	(0.011)	(0.012)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)
ln FDI_{ijt-1}		-0.000*	-0.000†	-0.000†	-0.000*	-0.000*	-0.000*	-0.000*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			-0.000	0.000	-0.000	-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				0.000	0.000	0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					0.000	0.000	0.000	0.000
					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						-0.000	0.000	0.000
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							-0.000*	-0.000
							(0.000)	(0.000)
ln FDI_{ijt-7}								-0.000
								(0.000)
Obs	4,202	3,694	3,312	2,975	2,647	2,330	2,057	1,837
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No	No	No
R^2	0.358	0.356	0.357	0.356	0.351	0.347	0.350	0.346
Log likelihood	-8163	-7503	-6919	-6461	-6029	-5595	-5150	-4788
AIC	16562	15238	14071	13151	12278	11411	10516	9783
BIC	17311	15959	14779	13835	12925	12043	11124	10351

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 2.4 FDI lag structure; OECD sample Poisson estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	0.559**	0.566**	0.563**	0.563**	0.559**	0.562**	0.574**	0.590**
	(0.047)	(0.048)	(0.048)	(0.049)	(0.050)	(0.050)	(0.050)	(0.051)
ln P_{jt}	0.326**	0.338**	0.345**	0.346**	0.335**	0.345**	0.364**	0.355**
	(0.041)	(0.041)	(0.040)	(0.040)	(0.042)	(0.045)	(0.046)	(0.049)
ln GDP_{it}	1.904**	1.872**	1.850**	1.822**	1.778**	1.723**	1.646**	1.527**
	(0.141)	(0.139)	(0.137)	(0.138)	(0.141)	(0.146)	(0.151)	(0.162)
ln GDP_{jt}	1.026**	1.048**	1.071**	1.079**	1.083**	1.081**	1.065**	1.074**
	(0.102)	(0.102)	(0.103)	(0.103)	(0.107)	(0.115)	(0.119)	(0.127)
ln $Dist_{ij}$	-0.165**	-0.161**	-0.155**	-0.151**	-0.148**	-0.150**	-0.151**	-0.152**
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)

$Lang_{ij}$	0.096**	0.100**	0.099**	0.098**	0.095**	0.095**	0.091**	0.088**
	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$Colony_{ij}$	0.291**	0.305**	0.316**	0.322**	0.325**	0.323**	0.319**	0.318**
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
$Prox_{ijt}$	1.825**	1.793**	1.766**	1.754**	1.762**	1.763**	1.764**	1.738**
	(0.134)	(0.135)	(0.133)	(0.132)	(0.135)	(0.136)	(0.137)	(0.142)
IPR_{it}	0.020	0.018	0.023	0.016	0.010	-0.003	-0.016	-0.033
	(0.030)	(0.031)	(0.031)	(0.030)	(0.031)	(0.032)	(0.034)	(0.044)
IPR_{jt}	0.229**	0.226**	0.212**	0.190**	0.169**	0.144**	0.126**	0.137**
	(0.038)	(0.038)	(0.037)	(0.035)	(0.034)	(0.033)	(0.037)	(0.043)
ln $Import_{ijt}$	-0.025	-0.015	-0.003	0.003	0.007	0.006	0.002	0.001
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
ln FDI_{ijt}	0.004	0.003	0.002	0.002	0.003	0.004	0.006	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln FDI_{ijt-1}		-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			-0.000*	-0.000	-0.000	-0.000†	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				-0.000	-0.000	-0.000	-0.000	-0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					-0.000	-0.000	-0.000	-0.000
					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						0.000	-0.000	0.000
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							0.000*	0.000
							(0.000)	(0.000)
ln FDI_{ijt-7}								0.000†
								(0.000)
Obs	7,008	6,560	6,167	5,774	5,361	4,966	4,599	4,222
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No	No	No
R^2	0.996	0.996	0.996	0.997	0.997	0.997	0.997	0.997
Log likelihood	-95468	-89972	-84672	-79622	-74952	-70553	-66697	-62563
AIC	191120	180126	169525	159422	150080	141282	133568	125300
BIC	191751	180744	170130	160015	150660	141855	134128	125852

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 2.5 FDI lag structure; OECD sample Poisson FE estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	0.903**	0.908**	0.908**	0.909**	0.899**	0.888**	0.891**	0.909**
	(0.113)	(0.106)	(0.103)	(0.102)	(0.097)	(0.092)	(0.089)	(0.090)

ln P_{jt}	0.651**	0.677**	0.695**	0.696**	0.699**	0.724**	0.755**	0.766**
	(0.102)	(0.092)	(0.090)	(0.091)	(0.097)	(0.100)	(0.102)	(0.106)
ln GDP_{it}	0.673**	0.605*	0.553*	0.518*	0.493*	0.471*	0.452*	0.423†
	(0.236)	(0.238)	(0.237)	(0.227)	(0.213)	(0.205)	(0.209)	(0.231)
ln GDP_{jt}	0.027	0.040	0.058	0.066	0.026	-0.026	-0.057	-0.002
	(0.408)	(0.378)	(0.366)	(0.350)	(0.343)	(0.341)	(0.339)	(0.362)
$Prox_{ijt}$	-0.260	-0.222	-0.172	-0.102	0.029	0.223	0.324	0.375
	(0.478)	(0.440)	(0.422)	(0.403)	(0.384)	(0.372)	(0.367)	(0.369)
IPR_{it}	-0.052	-0.055	-0.054	-0.054	-0.045	-0.048	-0.076	-0.133
	(0.115)	(0.123)	(0.125)	(0.122)	(0.117)	(0.117)	(0.127)	(0.150)
IPR_{jt}	0.175†	0.156	0.130	0.111	0.099	0.070	0.023	-0.022
	(0.099)	(0.107)	(0.115)	(0.122)	(0.124)	(0.132)	(0.147)	(0.174)
ln $Import_{ijt}$	-0.011	0.030	0.072	0.094†	0.080	0.060	0.041	0.031
	(0.045)	(0.047)	(0.048)	(0.051)	(0.054)	(0.057)	(0.060)	(0.063)
ln FDI_{ijt}	-0.006	-0.007	-0.009	-0.008	-0.004	-0.000	0.004	0.006
	(0.018)	(0.016)	(0.017)	(0.016)	(0.015)	(0.014)	(0.011)	(0.011)
ln FDI_{ijt-1}		-0.000**	-0.000**	-0.000*	-0.000†	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			-0.000	-0.000*	-0.000**	-0.000*	-0.000	-0.000†
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				0.000	0.000	0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					0.000*	0.000	0.000†	0.000†
					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						0.000**	0.000†	0.000*
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							0.000**	0.000
							(0.000)	(0.000)
ln FDI_{ijt-7}								0.000†
								(0.000)
Obs	6,467	6,059	5,704	5,349	4,990	4,661	4,344	4,027
Number of id	606	589	571	543	513	487	458	444
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No	No	No
Chi2 (model)	3961	4074	4599	4569	3192	2278	1719	1763
Log likelihood	-97700	-90034	-84782	-79679	-73848	-67785	-62216	-56889
AIC	195418	180086	169581	159376	147713	135588	124450	113796
BIC	195479	180147	169641	159435	147772	135646	124507	113853

Robust standard errors in parentheses

** p<0.01, * p<0.05, † p<0.1

Table A 2.6 FDI lag structure; OECD sample negative binomial estimates.

variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P_{it}	0.836**	0.824**	0.818**	0.809**	0.796**	0.802**	0.796**	0.787**
	(0.030)	(0.031)	(0.033)	(0.035)	(0.038)	(0.040)	(0.041)	(0.044)
ln P_{jt}	0.537**	0.514**	0.489**	0.456**	0.401**	0.382**	0.382**	0.326**
	(0.038)	(0.041)	(0.043)	(0.045)	(0.044)	(0.046)	(0.047)	(0.048)
ln GDP_{it}	0.517**	0.557**	0.534**	0.524**	0.497**	0.497**	0.519**	0.491**
	(0.138)	(0.132)	(0.139)	(0.142)	(0.147)	(0.157)	(0.165)	(0.175)
ln GDP_{jt}	1.217**	1.271**	1.322**	1.364**	1.454**	1.530**	1.503**	1.536**
	(0.130)	(0.134)	(0.137)	(0.140)	(0.145)	(0.151)	(0.154)	(0.162)
ln $Dist_{ij}$	-0.096**	-0.097**	-0.097**	-0.099**	-0.102**	-0.110**	-0.117**	-0.121**
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
$Lang_{ij}$	0.026	0.025	0.024	0.019	0.025	0.022	0.022	0.021
	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)
$Colony_{ij}$	0.255**	0.259**	0.263**	0.266**	0.269**	0.273**	0.280**	0.279**
	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)
$Prox_{ijt}$	1.390**	1.383**	1.376**	1.400**	1.450**	1.441**	1.491**	1.480**
	(0.086)	(0.089)	(0.091)	(0.093)	(0.094)	(0.096)	(0.098)	(0.100)
IPR_{it}	0.157**	0.158**	0.167**	0.169**	0.178**	0.176**	0.183**	0.173**
	(0.027)	(0.028)	(0.030)	(0.031)	(0.033)	(0.036)	(0.040)	(0.049)
IPR_{jt}	0.133**	0.150**	0.163**	0.163**	0.171**	0.179**	0.160**	0.142*
	(0.031)	(0.033)	(0.035)	(0.038)	(0.041)	(0.045)	(0.051)	(0.061)
ln $Import_{ijt}$	0.056**	0.059**	0.063**	0.061**	0.059**	0.052**	0.045**	0.041**
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)
ln FDI_{ijt}	0.025**	0.025**	0.024**	0.025**	0.026**	0.029**	0.028**	0.030**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
ln FDI_{ijt-1}		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-2}			-0.000	-0.000	-0.000†	-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-3}				0.000	0.000	0.000	0.000	0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-4}					0.000	-0.000	0.000	0.000
					(0.000)	(0.000)	(0.000)	(0.000)
ln FDI_{ijt-5}						0.000	0.000	-0.000
						(0.000)	(0.000)	(0.000)
ln FDI_{ijt-6}							0.000	0.000
							(0.000)	(0.000)
ln FDI_{ijt-7}								0.000
								(0.000)
Obs	7,008	6,560	6,167	5,774	5,361	4,966	4,599	4,222
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No	No	No
R^2	0.337	0.334	0.332	0.330	0.329	0.328	0.326	0.324

Log likelihood	-25852	-24644	-23544	-22455	-21274	-20127	-19057	-17899
AIC	51889	49474	47275	45095	42731	40438	38297	35979
BIC	52527	50105	47900	45707	43337	41037	38882	36557

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 2.7 Alternative knowledge stock measures; developing sample Poisson estimates

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}	0.643**	0.631**	0.749**	0.744**	0.746**
	(0.067)	(0.075)	(0.065)	(0.064)	(0.064)
ln P_{jt}(stock)	0.221**				
	(0.074)				
ln P_{jt}(stock dep)		0.128*			
		(0.063)			
ln P_{jt}(5yr)			0.182		
			(0.155)		
ln P_{jt}(L.5yr)				0.413*	
				(0.184)	
ln P_{jt}(10yr)					0.561*
					(0.250)
ln GDP_{it}	-0.737*	-0.386	-0.258	-0.229	-0.256
	(0.308)	(0.279)	(0.300)	(0.287)	(0.289)
ln GDP_{jt}	1.165*	1.187*	0.999	0.302	0.440
	(0.507)	(0.527)	(0.619)	(0.751)	(0.704)
ln $Dist_{ij}$	-0.082†	-0.089*	-0.090*	-0.091*	-0.090*
	(0.043)	(0.042)	(0.042)	(0.042)	(0.042)
$Lang_{ij}$	0.084	0.086	0.089	0.087	0.088
	(0.057)	(0.057)	(0.055)	(0.056)	(0.055)
$Colony_{ij}$	-0.120*	-0.131*	-0.137*	-0.137*	-0.137*
	(0.061)	(0.064)	(0.067)	(0.068)	(0.067)
$Prox_{ijt}$	1.976**	1.988**	2.003**	1.998**	1.989**
	(0.198)	(0.204)	(0.214)	(0.215)	(0.215)
IPR_{it}	-0.443**	-0.535**	-0.476**	-0.492**	-0.484**
	(0.143)	(0.140)	(0.140)	(0.138)	(0.138)
IPR_{jt}	-0.060	-0.037	-0.130	-0.275	-0.316
	(0.177)	(0.181)	(0.208)	(0.215)	(0.219)
ln $Import_{ijt}$	0.150**	0.146**	0.139**	0.140**	0.140**
	(0.035)	(0.036)	(0.037)	(0.036)	(0.036)
ln FDI_{ijt}	-0.111**	-0.138**	-0.128**	-0.134**	-0.131**
	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)
ln FDI_{ijt-1}	-0.010	-0.012	-0.016	-0.017	-0.016
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
ln FDI_{ijt}* IPR_{it}	0.021*	0.026**	0.023*	0.024*	0.023*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)

ln FDI_{ijt-1}* IPR_{it}	0.005*	0.007**	0.010**	0.010**	0.009**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-8.937	-12.391	-12.971	-5.448	-8.496
	(7.904)	(8.090)	(9.211)	(10.398)	(9.784)
Obs	991	991	1,001	1,001	1,001
Country dummy	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No
Chi2 (model)	139988	150292	121380	130796	111756
Log likelihood	-4481	-4535	-4568	-4534	-4534
AIC	9163	9269	9351	9280	9280
BIC	9653	9759	9862	9790	9790
RESET(Y^2)	0.944	0.964	0.562	0.712	0.577
RESET(Y^3)	0.953	0.989	0.844	0.930	c

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

c. The model did not converge

Table A 2.8 Alternative knowledge stock measures; developing sample Poisson FE estimates.

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}	1.134**	1.153**	0.891**	1.144**	0.954**
	(0.035)	(0.039)	(0.080)	(0.044)	(0.077)
ln P_{jt}(stock)	0.018				
	(0.107)				
ln P_{jt}(stock dep)		-0.036			
		(0.089)			
ln P_{jt}(5yr)			0.500**		
			(0.132)		
ln P_{jt}(L.5yr)				0.366*	
				(0.149)	
ln P_{jt}(10yr)					1.063**
					(0.293)
ln GDP_{it}	-1.299**	-1.129**	-0.233	-1.323**	-0.589*
	(0.389)	(0.305)	(0.315)	(0.221)	(0.230)
ln GDP_{jt}	2.287**	2.256**	1.789**	1.622**	0.980*
	(0.344)	(0.353)	(0.381)	(0.490)	(0.493)
$Prox_{ijt}$	1.097*	1.120*	0.898*	0.871*	0.704†
	(0.442)	(0.444)	(0.386)	(0.439)	(0.367)
IPR_{it}	-0.450*	-0.432*	-0.447**	-0.529**	-0.522**
	(0.183)	(0.210)	(0.159)	(0.178)	(0.160)
IPR_{jt}	0.553**	0.539**	0.271*	0.479**	0.128
	(0.147)	(0.148)	(0.124)	(0.137)	(0.135)
ln $Import_{ijt}$	-0.059	-0.065	0.029	-0.040	0.036
	(0.091)	(0.091)	(0.090)	(0.086)	(0.085)
ln FDI_{ijt}	-0.140**	-0.136**	-0.139**	-0.154**	-0.149**

	(0.046)	(0.051)	(0.038)	(0.043)	(0.037)
ln FDI_{ijt-1}	0.002	0.001	0.003	0.006	0.008
	(0.026)	(0.025)	(0.024)	(0.025)	(0.024)
ln FDI_{ijt}* IPR_{it}	0.025*	0.023†	0.023*	0.030*	0.028**
	(0.012)	(0.014)	(0.011)	(0.012)	(0.011)
ln FDI_{ijt-1}* IPR_{it}	0.011**	0.012**	0.010**	0.010**	0.009**
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
Obs	657	657	664	664	664
Number of id	227	227	230	230	230
Country dummy	No	No	No	No	No
Year dummy	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes
Chi2 (model)	25340	20030	18941	36957	23115
Log likelihood	-2548	-2542	-2275	-2475	-2247
AIC	5271	5265	4722	5167	4699
BIC	5325	5319	4776	5221	4753
RESET(Y^2)	0.0629	0.0769	0.00681	0.00770	0.000180
RESET(Y^3)	0.0131	0.0102	0.0243	0.000392	0.000168

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 2.9 Alternative knowledge stock measures; developing sample negative binomial estimates.

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}		0.862**	0.863**	0.851**	0.853**
		(0.091)	(0.077)	(0.076)	(0.075)
ln P_{jt}(stock)		-0.011			
		(0.068)			
ln P_{jt}(stock dep)			0.306*		
			(0.135)		
ln P_{jt}(5yr)				0.414**	
				(0.115)	
ln P_{jt}(L.5yr)					0.628**
					(0.158)
ln P_{jt}(10yr)		-0.379	-0.428	-0.395	-0.401
		(0.383)	(0.377)	(0.373)	(0.373)
ln GDP_{it}		1.735**	1.035*	0.442	0.367
		(0.518)	(0.524)	(0.613)	(0.582)
ln GDP_{jt}		-0.077	-0.083	-0.093†	-0.093†
		(0.055)	(0.054)	(0.054)	(0.053)
ln $Dist_{ij}$		0.078	0.072	0.081	0.077
		(0.089)	(0.088)	(0.088)	(0.088)
$Lang_{ij}$		0.036	0.032	0.026	0.026
		(0.090)	(0.089)	(0.090)	(0.090)
$Colony_{ij}$		2.070**	1.989**	1.986**	1.969**

	(0.246)	(0.244)	(0.240)	(0.241)
$Prox_{ijt}$	-0.069	-0.067	-0.090	-0.085
	(0.160)	(0.157)	(0.157)	(0.157)
IPR_{it}	-0.492*	-0.572**	-0.531**	-0.590**
	(0.195)	(0.192)	(0.182)	(0.181)
IPR_{jt}	0.069	0.066	0.059	0.058
	(0.047)	(0.047)	(0.046)	(0.046)
ln $Import_{ijt}$	0.027	0.026	0.020	0.022
	(0.046)	(0.045)	(0.045)	(0.045)
ln FDI_{ijt}	0.013	0.016	0.011	0.014
	(0.023)	(0.022)	(0.022)	(0.022)
ln FDI_{ijt-1}	-0.011	-0.011	-0.009	-0.009
	(0.012)	(0.012)	(0.012)	(0.012)
ln FDI_{ijt}* IPR_{it}	-0.000	0.000	0.000	0.000
	(0.005)	(0.005)	(0.005)	(0.005)
Constant	-19.546*	-12.123	-5.454	-6.485
	(8.269)	(8.094)	(8.928)	(8.493)
Obs	991	1,001	1,001	1,001
Country dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
FE	No	No	No	No
Chi2 (model)	82313	50114	48013	56215
Log likelihood	-2720	-2722	-2719	-2719
AIC	5643	5656	5651	5649
BIC	6138	6172	6167	6165
RESET(Y^2)	0.330	0.530	0.369	0.464
RESET(Y^3)	c	0.282	0.209	c

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

c. The model did not converge
Note: model (1) did not converge

Table A 2.10 Alternative knowledge stock measures; OECD sample Poisson estimates

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}	0.812**	0.646**	0.290**	0.308**	0.307**
	(0.126)	(0.136)	(0.082)	(0.069)	(0.075)
ln P_{jt}(stock)	-0.507**				
	(0.112)				
ln P_{jt}(stock dep)		-0.404**			
		(0.125)			
ln P_{jt}(5yr)			0.252**		
			(0.091)		
ln P_{jt}(L.5yr)				0.527**	
				(0.056)	
ln P_{jt}(10yr)					0.626**

					(0.090)
ln GDP_{it}	2.331**	2.444**	1.953**	1.903**	1.901**
	(0.246)	(0.343)	(0.295)	(0.258)	(0.272)
ln GDP_{jt}	1.444**	1.304**	1.161**	0.473*	0.756**
	(0.204)	(0.220)	(0.205)	(0.188)	(0.196)
ln $Dist_{ij}$	-0.112**	-0.116**	-0.121**	-0.132**	-0.130**
	(0.023)	(0.023)	(0.022)	(0.021)	(0.021)
$Lang_{ij}$	0.063†	0.068†	0.067†	0.078*	0.073*
	(0.035)	(0.037)	(0.036)	(0.034)	(0.035)
$Colony_{ij}$	0.279**	0.282**	0.296**	0.299**	0.300**
	(0.033)	(0.031)	(0.030)	(0.029)	(0.029)
$Prox_{ijt}$	2.530**	2.432**	2.168**	1.858**	1.947**
	(0.230)	(0.240)	(0.244)	(0.224)	(0.231)
IPR_{it}	0.587**	0.576**	0.532**	0.415**	0.436**
	(0.173)	(0.188)	(0.185)	(0.158)	(0.165)
IPR_{jt}	0.068	0.085†	0.071	0.078*	0.048
	(0.048)	(0.047)	(0.045)	(0.038)	(0.038)
ln $Import_{ijt}$	-0.007	-0.012	-0.008	-0.028	-0.021
	(0.025)	(0.025)	(0.026)	(0.024)	(0.025)
ln FDI_{ijt}	0.235**	0.221**	0.195**	0.170**	0.168**
	(0.061)	(0.067)	(0.065)	(0.056)	(0.058)
ln FDI_{ijt-1}	-0.001	-0.001	0.007	0.012	0.013
	(0.016)	(0.020)	(0.018)	(0.016)	(0.017)
ln FDI_{ijt}* IPR_{it}	-0.048**	-0.045**	-0.040**	-0.032**	-0.033**
	(0.013)	(0.014)	(0.013)	(0.012)	(0.012)
ln FDI_{ijt-1}* IPR_{it}	0.001	0.003	0.001	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	-50.015**	-49.793**	-43.086**	-35.044**	-39.987**
	(4.818)	(5.982)	(5.001)	(4.346)	(4.584)
Obs	1,371	1,371	1,372	1,372	1,372
Country dummy	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No
Chi2 (model)	579558	601890	387871	437762	442203
Log likelihood	-45255	-47269	-49353	-43367	-45253
AIC	94299	99009	100210	89408	92457
BIC	94691	99400	100607	89805	92854
RESET(Y^2)	0.0740	0.0165	0.00768	6.48e-05	0.00111
RESET(Y^3)	2.73e-07	1.15e-06	c	1.15e-06	c

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

c. The model did not converge

Table A 2.11 Alternative knowledge stock measures; OECD sample Poisson FE estimates

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}	1.361**	1.569**	0.859**	1.530**	1.149**
	(0.040)	(0.039)	(0.064)	(0.039)	(0.040)
ln P_{jt}(stock)	-0.726**				
	(0.073)				
ln P_{jt}(stock dep)		-0.812**			
		(0.111)			
ln P_{jt}(5yr)			0.637**		
			(0.082)		
ln P_{jt}(L.5yr)				0.600**	
				(0.090)	
ln P_{jt}(10yr)					1.058**
					(0.128)
ln GDP_{it}	1.897**	1.588**	0.273	-0.933*	-0.516†
	(0.344)	(0.440)	(0.265)	(0.415)	(0.269)
ln GDP_{jt}	1.495**	1.004**	0.521†	-0.563	-0.463
	(0.203)	(0.304)	(0.300)	(0.466)	(0.405)
$Prox_{ijt}$	1.741**	1.164**	-0.215	-1.638**	-1.322**
	(0.379)	(0.381)	(0.417)	(0.502)	(0.427)
IPR_{it}	0.319	-0.047	-0.267	-0.968**	-0.747*
	(0.211)	(0.245)	(0.265)	(0.309)	(0.300)
IPR_{jt}	0.076	0.087	0.035	0.242*	0.061
	(0.051)	(0.056)	(0.071)	(0.099)	(0.090)
ln $Import_{ijt}$	0.090	0.142	0.244**	0.324*	0.291**
	(0.111)	(0.116)	(0.085)	(0.144)	(0.110)
ln FDI_{ijt}	0.155*	0.024	-0.096	-0.326**	-0.249*
	(0.078)	(0.090)	(0.091)	(0.110)	(0.101)
ln FDI_{ijt-1}	-0.053*	-0.076*	-0.052*	-0.079**	-0.051*
	(0.025)	(0.039)	(0.025)	(0.027)	(0.023)
ln FDI_{ijt}* IPR_{it}	-0.037*	-0.013	0.011	0.064**	0.047*
	(0.017)	(0.020)	(0.019)	(0.023)	(0.022)
ln FDI_{ijt-1}* IPR_{it}	0.010**	0.015**	0.005*	0.006*	0.004†
	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Obs	1,137	1,137	1,137	1,137	1,137
Number of id	402	402	402	402	402
Country dummy	No	No	No	No	No
Year dummy	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes
Chi2 (model)	50494	254832	91719	40346	133118
Log likelihood	-26696	-34441	-30862	-43793	-32330
AIC	55829	70815	62293	88708	64782
BIC	55890	70875	62353	88768	64843
RESET(Y^2)	0.268	0.201	0.107	0.0839	0.375
RESET(Y^3)	1.52e-07	6.09e-06	0.169	0.0170	0.463

Table A 2.12 Alternative knowledge stock measures, OECD sample negative binomial estimates

Variable	(1) Stock	(2) Stock Dep	(3) 5yr	(4) L.5yr	(5) 10yr
ln P_{it}	0.675**	0.495**	0.594**	0.589**	0.598**
	(0.123)	(0.131)	(0.088)	(0.085)	(0.086)
ln P_{jt}(stock)	-0.138				
	(0.115)				
ln P_{jt}(stock dep)		0.109			
		(0.135)			
ln P_{jt}(5yr)			0.301**		
			(0.091)		
ln P_{jt}(L.5yr)				0.552**	
				(0.075)	
ln P_{jt}(10yr)					0.688**
					(0.112)
ln GDP_{it}	0.977**	0.657†	0.813*	0.828**	0.826**
	(0.346)	(0.384)	(0.323)	(0.311)	(0.314)
ln GDP_{jt}	1.502**	1.516**	0.918**	0.019	0.179
	(0.324)	(0.328)	(0.317)	(0.344)	(0.336)
ln $Dist_{ij}$	-0.088**	-0.084**	-0.092**	-0.098**	-0.099**
	(0.029)	(0.029)	(0.029)	(0.028)	(0.028)
$Lang_{ij}$	0.025	0.023	0.024	0.026	0.027
	(0.042)	(0.042)	(0.042)	(0.040)	(0.040)
$Colony_{ij}$	0.241**	0.244**	0.244**	0.257**	0.249**
	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
$Prox_{ijt}$	1.390**	1.372**	1.273**	1.240**	1.239**
	(0.223)	(0.215)	(0.207)	(0.201)	(0.202)
IPR_{it}	0.659**	0.654**	0.628**	0.579**	0.579**
	(0.196)	(0.187)	(0.195)	(0.193)	(0.195)
IPR_{jt}	0.030	0.037	-0.004	-0.020	-0.040
	(0.076)	(0.075)	(0.075)	(0.068)	(0.070)
ln $Import_{ijt}$	0.034	0.038	0.032	0.027	0.026
	(0.028)	(0.028)	(0.027)	(0.027)	(0.027)
ln FDI_{ijt}	0.284**	0.268**	0.264**	0.245**	0.245**
	(0.073)	(0.069)	(0.073)	(0.073)	(0.074)
ln FDI_{ijt-1}	-0.028	-0.030	-0.030	-0.036	-0.033
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
ln FDI_{ijt}* IPR_{it}	-0.058**	-0.054**	-0.054**	-0.049**	-0.049**
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)
ln FDI_{ijt-1}* IPR_{it}	0.010*	0.010*	0.011*	0.012**	0.012**
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Constant	-35.190**	-31.737**	-28.149**	-18.266**	-21.932**
	(5.918)	(6.341)	(5.624)	(5.695)	(5.499)
Obs	1,371	1,371	1,372	1,372	1,372
Country dummy	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes

FE	No	No	No	No	No
Chi2 (model)	111051	106274	103686	106552	105218
Log likelihood	-6737	-6738	-6730	-6707	-6715
AIC	13628	13629	13616	13568	13584
BIC	14025	14026	14018	13971	13987
RESET(Y^2)	0.273	0.495	0.238	0.647	0.432
RESET(Y^3)	c	c	c	c	c

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

c. The model did not converge

A3. Empirical Results Appendix

The derivation of the elasticities presented in equations 8.1 and 8.2:

Consider our model in non-linear form, rewriting slightly and omitting subscripts for the sake of clarity:

$$E[C|\boldsymbol{X}] = e^{\beta_0 X_0 + \beta_1 IPR} \times e^{(\beta_2 + \beta_{12} IPR)\ln(FDI)}$$

since $(FDI)^{\beta_2} = e^{\beta_2 \ln(FDI)}$, where X_0 is a vector of the other explanatory variables. Then the partial derivative with respect to IPR is, using the chain rule:

$$\frac{\partial E[C|\boldsymbol{X}]}{\partial IPR} = (\beta_1 e^{\beta_0 X_0 + \beta_1 IPR} \times e^{(\beta_2 + \beta_{12} IPR)\ln(FDI)}) \times (\beta_{12} \ln(FDI) e^{\beta_0 X_0 + \beta_1 IPR} \times e^{(\beta_2 + \beta_{12} IPR)\ln(FDI)})$$

$$= (\beta_1 + \beta_{12} \ln(FDI)) e^{\beta_0 X_0 + \beta_1 IPR} \times e^{(\beta_2 + \beta_{12} IPR)\ln(FDI)}$$

$$= (\beta_1 + \beta_{12} \ln(FDI)) E[C|\boldsymbol{X}]$$

Similarly, the partial derivative with respect to ln(FDI) (not FDI, since the interaction effect is defined with respect to the transformed variable) is:

$$\frac{\partial E[C|\boldsymbol{X}]}{\partial \ln(FDI)} = (\beta_2 + \beta_{12} IPR) e^{\beta_0 X_0 + \beta_1 IPR} \times e^{(\beta_2 + \beta_{12} IPR)\ln(FDI)}$$
$$= (\beta_2 + \beta_{12} IPR) E[C|\boldsymbol{X}].$$

What we are interested in is the second order cross partial derivative, which equals the interaction effect. Using the product rule, we take the derivative of $\partial E[C|\boldsymbol{X}]/\partial IPR$ with respect to ln(FDI) and substitute $\partial E[C|\boldsymbol{X}]/\partial \ln(FDI)$ from the above expression, which yields:

$$\frac{\partial^2 E[C|\boldsymbol{X}]}{\partial IPR \partial \ln(FDI)} = \beta_{12} E[C|\boldsymbol{X}] + (\beta_1 + \beta_{12} \ln(FDI)) \frac{\partial E[C|\boldsymbol{X}]}{\partial \ln(FDI)}$$

$$= \beta_{12} E[C|\boldsymbol{X}] + (\beta_1 + \beta_{12} \ln(FDI))(\beta_2 + \beta_{12} IPR) E[C|\boldsymbol{X}]$$

$$= \{\beta_1 \beta_2 + \beta_{12}[1 + \beta_1 IPR + \beta_2 \ln(FDI) + \beta_{12} IPR \times \ln(FDI)]\}\, E[C|\boldsymbol{X}]$$

Finally we divide each side by $E[C|\boldsymbol{X}]$ to express it as an elasticity. This depends on the values of IPR and ln(FDI).

Replacing 5yr of patents applied for source country knowledge "mass".

Table A 3.1 P_{jt} as 5yr stock; developing sample Poisson estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.806**	0.749**	0.792**	0.723**		
	(0.067)	(0.065)	(0.074)	(0.065)		
ln (P_{it} + 1)					0.796**	0.728**
					(0.075)	(0.066)
ln P_{jt}	0.254†	0.182	0.154	0.148	0.155	0.149
	(0.154)	(0.155)	(0.142)	(0.143)	(0.142)	(0.143)
ln GDP_{it}	-0.541†	-0.258	-0.684*	-0.455	-0.690*	-0.460†
	(0.305)	(0.300)	(0.295)	(0.277)	(0.295)	(0.277)
ln GDP_{jt}	0.934	0.999	1.106†	0.974†	1.108†	0.979†
	(0.660)	(0.619)	(0.593)	(0.572)	(0.593)	(0.571)
ln $Dist_{ij}$	-0.134**	-0.090*	-0.111*	-0.071†	-0.110*	-0.070
	(0.043)	(0.042)	(0.044)	(0.043)	(0.044)	(0.043)
$Lang_{ij}$	0.146*	0.089	0.160*	0.097†	0.159*	0.096†
	(0.063)	(0.055)	(0.071)	(0.056)	(0.070)	(0.056)
$Colony_{ij}$	-0.168**	-0.137*	-0.218**	-0.190**	-0.218**	-0.190**
	(0.064)	(0.067)	(0.065)	(0.068)	(0.065)	(0.068)
$Prox_{ijt}$	1.930**	2.003**	2.043**	2.071**	2.055**	2.085**
	(0.252)	(0.214)	(0.260)	(0.225)	(0.258)	(0.224)
IPR_{it}	-0.187**	-0.476**	-0.171**	-0.402**	-0.171**	-0.412**
	(0.034)	(0.140)	(0.037)	(0.122)	(0.037)	(0.122)
IPR_{jt}	-0.211	-0.130	-0.237	-0.180	-0.238	-0.179
	(0.230)	(0.208)	(0.176)	(0.156)	(0.177)	(0.156)
ln $Import_{ijt}$	0.063	0.139**	0.114**	0.165**	0.113**	0.165**
	(0.039)	(0.037)	(0.040)	(0.038)	(0.040)	(0.038)
ln FDI_{ijt}	-0.042**	-0.128**				
	(0.015)	(0.040)				
ln FDI_{ijt-1}	0.025	-0.016				
	(0.019)	(0.019)				
ln FDI_{ijt}			-0.020	-0.088**	-0.020	-0.091**
(positive)			(0.015)	(0.034)	(0.015)	(0.034)
ln FDI_{ijt-1}			0.000	-0.021	0.000	-0.021
(positive)			(0.012)	(0.016)	(0.012)	(0.016)
ln FDI_{ijt}			-0.005	-0.109*	-0.005	-0.112*
(negative)			(0.014)	(0.055)	(0.014)	(0.054)
ln FDI_{ijt-1}			-0.016	0.039	-0.016	0.038
(negative)			(0.013)	(0.031)	(0.013)	(0.031)
ln FDI_{ijt}			0.149	0.100	0.144	0.091
(zero)			(0.204)	(0.232)	(0.204)	(0.233)
ln FDI_{ijt}* IPR_{it}		0.023*				
		(0.010)				
ln FDI_{ijt-1}* IPR_{it}		0.010**				
		(0.002)				

ln FDI_{ijt}* IPR_{it}				0.017*		0.018*
(positive)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				0.008**		0.008**
(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				0.025†		0.025*
(negative)				(0.013)		(0.013)
ln FDI_{ijt-1}* IPR_{it}				-0.011†		-0.011†
(negative)				(0.006)		(0.006)
Obs	1,040	1,001	1,951	1,885	2,766	2,700
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Chi2 (model)	72453	100921	68295	95509	99663	138477
Log likelihood	-5088	-4572	-7327	-6613	-7350	-6640
AIC	10380	9351	14868	13448	14944	13532
BIC	10884	9862	15464	14063	15666	14276

Table A 3.2 P_{jt} as 5yr stock; developing sample negative binomial estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.861**	0.863**	0.874**	0.879**		
	(0.063)	(0.077)	(0.062)	(0.071)		
ln (P_{it} + 1)					0.918**	0.941**
					(0.063)	(0.072)
ln P_{jt}	0.302*	0.306*	0.121	0.156	0.089	0.121
	(0.135)	(0.135)	(0.140)	(0.140)	(0.140)	(0.140)
ln GDP_{it}	-0.423	-0.428	-0.826*	-0.817*	-0.964**	-1.036**
	(0.335)	(0.377)	(0.332)	(0.349)	(0.330)	(0.347)
ln GDP_{jt}	1.114*	1.035*	1.371**	1.344**	1.445**	1.440**
	(0.519)	(0.524)	(0.510)	(0.521)	(0.513)	(0.524)
ln $Dist_{ij}$	-0.090†	-0.083	-0.105†	-0.119*	-0.096†	-0.109†
	(0.052)	(0.054)	(0.056)	(0.059)	(0.056)	(0.059)
$Lang_{ij}$	0.103	0.072	0.156†	0.143	0.138	0.127
	(0.085)	(0.088)	(0.090)	(0.092)	(0.090)	(0.092)
$Colony_{ij}$	0.003	0.032	-0.057	-0.034	-0.041	-0.019
	(0.088)	(0.089)	(0.098)	(0.099)	(0.098)	(0.099)
$Prox_{ijt}$	1.968**	1.989**	1.857**	1.936**	2.038**	2.124**
	(0.239)	(0.244)	(0.242)	(0.246)	(0.239)	(0.241)
IPR_{it}	-0.175**	-0.067	-0.137**	-0.071	-0.135**	-0.072
	(0.051)	(0.157)	(0.051)	(0.119)	(0.052)	(0.118)
IPR_{jt}	-0.524**	-0.572**	-0.422*	-0.398*	-0.390*	-0.358†
	(0.186)	(0.192)	(0.186)	(0.191)	(0.187)	(0.191)
ln $Import_{ijt}$	0.040	0.066	0.018	0.020	0.014	0.017
	(0.045)	(0.047)	(0.047)	(0.049)	(0.047)	(0.049)
ln FDI_{ijt}	-0.005	0.026				
	(0.018)	(0.045)				

ln FDI_{ijt-1}	0.015	0.016				
	(0.017)	(0.022)				
ln FDI_{ijt}		-0.011				
(positive)		(0.012)				
ln FDI_{ijt-1}		0.000				
(positive)		(0.005)				
ln FDI_{ijt}			0.014	0.032	0.015	0.027
(negative)			(0.017)	(0.033)	(0.017)	(0.032)
ln FDI_{ijt-1}			-0.007	0.021	-0.007	0.013
(negative)			(0.012)	(0.018)	(0.012)	(0.018)
ln FDI_{ijt}			0.032†	0.020	0.033†	0.018
(zero)			(0.018)	(0.047)	(0.018)	(0.047)
ln FDI_{ijt}* IPR_{it}			-0.020	0.053	-0.021	0.042
			(0.016)	(0.034)	(0.016)	(0.035)
ln FDI_{ijt-1}* IPR_{it}			0.634*	0.574†	0.621*	0.560†
			(0.277)	(0.294)	(0.271)	(0.287)
ln FDI_{ijt}* IPR_{it}				-0.008		-0.007
(positive)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				-0.006		-0.003
(positive)				(0.005)		(0.005)
ln FDI_{ijt}* IPR_{it}				-0.001		0.000
(negative)				(0.013)		(0.013)
ln FDI_{ijt-1}* IPR_{it}				-0.020*		-0.017†
(negative)				(0.009)		(0.009)
Obs	1,040	1,001	1,951	1,885	2,766	2,700
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Chi2 (model)	42548	45020	38736	34213	58167	49394
Log likelihood	-2842	-2723	-3955	-3777	-3975	-3798
AIC	5891	5656	8127	7778	8196	7849
BIC	6400	6172	8729	8399	8925	8599

Table A 3.3 P_{jt} as 5yr stock; developing sample Poisson FE estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.856**	0.891**	0.884**	0.897**		
	(0.076)	(0.080)	(0.080)	(0.074)		
ln (P_{it} + 1)					0.886**	0.898**
					(0.080)	(0.074)
ln P_{jt}	0.448**	0.500**	0.463**	0.504**	0.463**	0.505**
	(0.106)	(0.132)	(0.102)	(0.111)	(0.102)	(0.111)

ln GDP_{it}	0.099	-0.233	0.094	-0.227	0.093	-0.222
	(0.262)	(0.315)	(0.246)	(0.281)	(0.247)	(0.280)
ln GDP_{jt}	2.172**	1.789**	2.384**	1.935**	2.382**	1.934**
	(0.449)	(0.381)	(0.436)	(0.360)	(0.435)	(0.359)
$Prox_{ijt}$	0.861*	0.898*	0.586	0.726*	0.592	0.733*
	(0.377)	(0.386)	(0.399)	(0.363)	(0.398)	(0.362)
IPR_{it}	-0.157**	-0.447**	-0.152**	-0.365**	-0.152**	-0.371**
	(0.036)	(0.159)	(0.037)	(0.115)	(0.037)	(0.115)
IPR_{jt}	0.284†	0.271*	0.197	0.240*	0.198	0.241*
	(0.146)	(0.124)	(0.125)	(0.107)	(0.125)	(0.107)
ln $Import_{ijt}$	-0.166*	0.029	-0.116	0.049	-0.115	0.048
	(0.075)	(0.090)	(0.081)	(0.084)	(0.081)	(0.084)
ln FDI_{ijt}	-0.048*	-0.139**				
	(0.022)	(0.038)				
ln FDI_{ijt-1}	0.039	0.003				
	(0.027)	(0.024)				
ln FDI_{ijt}			-0.050*	-0.123**	-0.050*	-0.125**
(positive)			(0.019)	(0.029)	(0.019)	(0.029)
ln FDI_{ijt-1}			-0.001	-0.014	-0.001	-0.014
(positive)			(0.015)	(0.020)	(0.015)	(0.020)
ln FDI_{ijt}			-0.030	-0.174**	-0.030	-0.176**
(negative)			(0.019)	(0.046)	(0.019)	(0.046)
ln FDI_{ijt-1}			-0.014	0.007	-0.014	0.006
(negative)			(0.016)	(0.030)	(0.016)	(0.030)
ln FDI_{ijt}			-0.494*	-0.689**	-0.492*	-0.686**
(zero)			(0.239)	(0.233)	(0.238)	(0.233)
ln FDI_{ijt}* IPR_{it}		0.023*				
		(0.011)				
ln FDI_{ijt-1}* IPR_{it}		0.010**				
		(0.003)				
ln FDI_{ijt}* IPR_{it}				0.017*		0.018*
(positive)				(0.008)		(0.008)
ln FDI_{ijt-1}* IPR_{it}				0.010**		0.010**
(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				0.032**		0.033**
(negative)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				0.000		0.000
(negative)				(0.006)		(0.006)
Obs	707	664	1,118	1,056	1,152	1,092
Number of id	236	230	359	353	366	362
Country dummy	No	No	No	No	No	No
Year dummy	No	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes	Yes
Chi2 (model)	7091	11034	7149	15132	7185	15053
Log likelihood	-2786	-2349	-3672	-3132	-3685	-3147
AIC	5592	4722	7370	6298	7395	6329
BIC	5638	4776	7435	6383	7461	6414

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Same models as main text, excluding the last period from sample

Table A 3.4 Exclude 2010 period; developing sample Poisson estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.924**	0.751**	0.913**	0.726**		
	(0.065)	(0.096)	(0.068)	(0.097)		
ln (P_{it} + 1)					0.916**	0.728**
					(0.068)	(0.098)
ln P_{jt}	0.839*	0.767**	0.763*	0.814**	0.762*	0.814**
	(0.326)	(0.291)	(0.307)	(0.293)	(0.307)	(0.293)
ln GDP_{it}	-1.455*	-0.053	-1.740**	-0.200	-1.749**	-0.206
	(0.607)	(0.670)	(0.571)	(0.652)	(0.570)	(0.656)
ln GDP_{jt}	0.223	0.470	0.451	0.445	0.449	0.444
	(1.120)	(0.834)	(0.924)	(0.699)	(0.923)	(0.700)
ln $Dist_{ij}$	-0.179**	-0.137**	-0.139**	-0.099*	-0.139**	-0.099*
	(0.045)	(0.043)	(0.048)	(0.048)	(0.048)	(0.048)
$Lang_{ij}$	0.178**	0.139**	0.167**	0.116*	0.167**	0.116*
	(0.055)	(0.046)	(0.063)	(0.051)	(0.063)	(0.051)
$Colony_{ij}$	-0.202**	-0.181*	-0.248**	-0.210**	-0.247**	-0.210**
	(0.072)	(0.071)	(0.073)	(0.073)	(0.073)	(0.073)
$Prox_{ijt}$	2.025**	2.040**	2.158**	2.139**	2.157**	2.139**
	(0.286)	(0.223)	(0.301)	(0.239)	(0.301)	(0.239)
IPR_{it}	-0.233**	-0.431**	-0.207**	-0.425**	-0.207**	-0.426**
	(0.053)	(0.160)	(0.055)	(0.135)	(0.055)	(0.135)
IPR_{jt}	-0.452	-0.356	-0.408†	-0.348†	-0.408†	-0.348†
	(0.326)	(0.262)	(0.226)	(0.180)	(0.225)	(0.180)
ln $Import_{ijt}$	0.037	0.132**	0.071	0.149**	0.071	0.149**
	(0.043)	(0.040)	(0.046)	(0.044)	(0.046)	(0.044)
ln FDI_{ijt}	-0.033*	-0.123**				
	(0.015)	(0.042)				
ln FDI_{ijt-1}	0.004	-0.064**				
	(0.021)	(0.020)				
ln FDI_{ijt}			-0.004	-0.104**	-0.004	-0.104**
(positive)			(0.016)	(0.032)	(0.016)	(0.032)
ln FDI_{ijt-1}			-0.004	-0.059**	-0.004	-0.059**
(positive)			(0.012)	(0.017)	(0.012)	(0.017)
ln FDI_{ijt}			0.012	-0.128*	0.012	-0.128*
(negative)			(0.014)	(0.054)	(0.014)	(0.054)
ln FDI_{ijt-1}			-0.004	-0.042	-0.004	-0.042
(negative)			(0.017)	(0.038)	(0.017)	(0.038)
ln FDI_{ijt}			0.208	0.138	0.207	0.137
(zero)			(0.195)	(0.236)	(0.195)	(0.236)
ln FDI_{ijt}* IPR_{it}		0.025*				
		(0.011)				
ln FDI_{ijt-1}* IPR_{it}		0.014**				
		(0.003)				

ln FDI_{ijt}* IPR_{it}				0.025**		0.025**
(positive)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				0.013**		0.013**
(positive)				(0.003)		(0.003)
ln FDI_{ijt}* IPR_{it}				0.032*		0.032*
(negative)				(0.013)		(0.013)
ln FDI_{ijt-1}* IPR_{it}				0.006		0.006
(negative)				(0.012)		(0.012)
Obs	650	611	1,094	1,028	1,261	1,195
Country dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No	No
Log likelihood	-3712	-3115	-5090	-4340	-5090	-4345
AIC	7621	6433	10389	8898	10420	8938
BIC	8065	6879	10914	9436	11037	9569

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 3.5 Exclude 2010 period; developing sample negative binomial estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.963**		0.907**	1.012**		
	(0.071)		(0.071)	(0.097)		
ln (P_{it} + 1)					0.919**	1.024**
					(0.071)	(0.097)
ln P_{jt}	0.594**		0.721**	0.647**	0.712**	0.636**
	(0.177)		(0.175)	(0.176)	(0.175)	(0.176)
ln GDP_{it}	-1.875**		-1.763**	-2.427**	-1.834**	-2.515**
	(0.533)		(0.509)	(0.631)	(0.509)	(0.630)
ln GDP_{jt}	0.600		0.726	1.019†	0.729	1.024†
	(0.566)		(0.561)	(0.567)	(0.559)	(0.566)
ln $Dist_{ij}$	-0.146**		-0.111†	-0.136*	-0.110†	-0.137*
	(0.055)		(0.057)	(0.059)	(0.057)	(0.059)
$Lang_{ij}$	0.054		0.030	0.019	0.029	0.018
	(0.083)		(0.089)	(0.091)	(0.089)	(0.091)
$Colony_{ij}$	0.136		0.060	0.080	0.061	0.081
	(0.104)		(0.096)	(0.098)	(0.096)	(0.098)
$Prox_{ijt}$	2.516**		2.283**	2.316**	2.294**	2.326**
	(0.252)		(0.264)	(0.267)	(0.263)	(0.266)
IPR_{it}	-0.196**		-0.131*	-0.025	-0.137*	-0.028
	(0.062)		(0.062)	(0.111)	(0.062)	(0.111)
IPR_{jt}	-0.513**		-0.291†	-0.258	-0.292†	-0.261
	(0.193)		(0.158)	(0.159)	(0.158)	(0.159)
ln $Import_{ijt}$	0.017		0.024	0.038	0.024	0.036

	(0.046)	(0.042)	(0.044)	(0.042)	(0.044)
ln FDI_{ijt}	0.010				
	(0.017)				
ln FDI_{ijt-1}	-0.006				
	(0.017)				
ln FDI_{ijt}		0.006	0.022	0.005	0.023
(positive)		(0.015)	(0.027)	(0.015)	(0.027)
ln FDI_{ijt-1}		-0.007	-0.003	-0.007	-0.001
(positive)		(0.012)	(0.020)	(0.012)	(0.020)
ln FDI_{ijt}		0.015	0.036	0.015	0.037
(negative)		(0.016)	(0.042)	(0.016)	(0.042)
ln FDI_{ijt-1}		0.008	-0.016	0.008	-0.015
(negative)		(0.016)	(0.035)	(0.016)	(0.035)
ln FDI_{ijt}		0.103	-0.045	0.099	-0.051
(zero)		(0.193)	(0.210)	(0.193)	(0.210)
ln FDI_{ijt}* IPR_{it}			-0.008		-0.008
(positive)			(0.007)		(0.007)
ln FDI_{ijt-1}* IPR_{it}			0.000		-0.000
(positive)			(0.005)		(0.005)
ln FDI_{ijt}* IPR_{it}			-0.009		-0.009
(negative)			(0.011)		(0.011)
ln FDI_{ijt-1}* IPR_{it}			0.012		0.011
(negative)			(0.012)		(0.012)
Obs	650	1,094	1,028	1,261	1,195
Country dummy	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes
FE	No	No	No	No	No
Log likelihood	-2022	-2713	-2539	-2715	-2540
AIC	4243	5639	5299	5671	5331
BIC	4691	6169	5842	6293	5967

Robust standard errors in parentheses
** p<0.01, * p<0.05, † p<0.1

Table A 3.6 Exclude 2010 period; developing sample Poisson FE estimates

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln P_{it}	0.945**	0.913**	0.952**	0.918**		
	(0.085)	(0.099)	(0.083)	(0.090)		
ln (P_{it} + 1)					0.955**	0.921**
					(0.083)	(0.090)
ln P_{jt}	0.937**	0.808**	1.009**	0.904**	1.009**	0.904**
	(0.231)	(0.163)	(0.223)	(0.189)	(0.224)	(0.189)
ln GDP_{it}	-0.166	-0.031	-0.172	-0.152	-0.177	-0.156
	(0.364)	(0.450)	(0.357)	(0.406)	(0.356)	(0.407)
ln GDP_{jt}	1.729**	1.411**	1.609**	1.180**	1.609**	1.190**
	(0.542)	(0.398)	(0.513)	(0.364)	(0.513)	(0.364)
$Prox_{ijt}$	0.520	0.059	0.351	0.049	0.351	0.049

	(0.467)	(0.414)	(0.409)	(0.360)	(0.409)	(0.360)
IPR_{it}	-0.192**	-0.480**	-0.177**	-0.376**	-0.177**	-0.374**
	(0.046)	(0.158)	(0.044)	(0.119)	(0.044)	(0.119)
IPR_{jt}	0.248*	0.181	0.173†	0.143†	0.174†	0.145†
	(0.126)	(0.111)	(0.102)	(0.083)	(0.103)	(0.083)
ln $Import_{ijt}$	-0.378**	-0.196*	-0.345**	-0.160†	-0.344**	-0.163†
	(0.091)	(0.096)	(0.087)	(0.089)	(0.087)	(0.089)
ln FDI_{ijt}	-0.027	-0.098**				
	(0.021)	(0.034)				
ln FDI_{ijt-1}	-0.005	-0.057*				
	(0.031)	(0.023)				
ln FDI_{ijt}			-0.024	-0.079**	-0.024	-0.079**
(positive)			(0.018)	(0.026)	(0.018)	(0.026)
ln FDI_{ijt-1}			-0.010	-0.040*	-0.010	-0.040*
(positive)			(0.013)	(0.019)	(0.013)	(0.019)
ln FDI_{ijt}			-0.002	-0.071†	-0.002	-0.071†
(negative)			(0.017)	(0.043)	(0.017)	(0.043)
ln FDI_{ijt-1}			0.002	-0.004	0.002	-0.005
(negative)			(0.018)	(0.030)	(0.018)	(0.030)
ln FDI_{ijt}			-0.318	-0.140	-0.321	-0.145
(zero)			(0.260)	(0.216)	(0.259)	(0.216)
ln FDI_{ijt}* IPR_{it}		0.026*				
		(0.010)				
ln FDI_{ijt-1}* IPR_{it}		0.009*				
		(0.004)				
ln FDI_{ijt}* IPR_{it}				0.020**		0.020**
(positive)				(0.008)		(0.008)
ln FDI_{ijt-1}* IPR_{it}				0.009*		0.009*
(positive)				(0.004)		(0.004)
ln FDI_{ijt}* IPR_{it}				0.020*		0.020*
(negative)				(0.009)		(0.009)
ln FDI_{ijt-1}* IPR_{it}				-0.002		-0.002
(negative)				(0.007)		(0.007)
Obs	468	400	717	613	723	619
Number of id	198	164	302	250	303	251
Country dummy	No	No	No	No	No	No
Year dummy	No	No	No	No	No	No
FE	Yes	Yes	Yes	Yes	Yes	Yes
Log likelihood	-1573	-1130	-1971	-1483	-1972	-1487
AIC	3165	2284	3967	3001	3970	3007
BIC	3207	2332	4027	3076	4030	3083

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$, † $p<0.1$