Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is*

Tomas Havranek[a,b] and Zuzana Irsova[b]

[a]Czech National Bank
[b]Charles University, Prague

Revision: January 5, 2011

Abstract

In the last decade, more than 100 researchers have examined productivity spillovers from foreign affiliates to local firms in upstream or downstream sectors. Yet results vary broadly across methods and countries. To examine these vertical spillovers in a systematic way, we collected 3,626 estimates of spillovers and reviewed the literature quantitatively. Our meta-analysis indicates that model misspecifications reduce the reported estimates and journals select relatively large estimates for publication. Taking these biases into consideration, the average spillover to suppliers is positive and economically significant, whereas the spillover to buyers is negligible. Greater spillovers are received by countries that have underdeveloped financial systems and are open to international trade. Greater spillovers are generated by investors who come from distant countries and have only a slight technological edge over local firms.

Keywords: Foreign direct investment; Productivity; Spillovers; Meta-analysis; Publication selection bias

JEL Codes: C83; F23

1 Introduction

Few topics in international economics have been examined as extensively as productivity spillovers from foreign affiliates to domestic firms. The evidence for spillovers had been mixed until Javorcik (2004a) redirected the attention of researchers from horizontal (within-sector) to vertical (between-sector) spillovers. Since then, there has been a virtual explosion of studies on vertical spillovers, and empirical research in this area is still growing at an exponential rate with more than a score of studies published in the last two years alone. A consensus has emerged

*We thank Beata Javorcik, Tom Stanley, the editor, and two anonymous reviewers for their helpful comments. The paper benefited from discussions at the Meta-Analysis of Economics Research Colloquium, Conway, 2010, and the Annual Global Development Conference, Bogota, 2011.

that spillovers from foreign affiliates to their suppliers in host countries are positive and significant, yet the estimated size of these spillovers varies broadly. The point estimates of the economic effect of backward linkages reported by the two best known studies, Javorcik (2004a) and Blalock & Gertler (2008), differ by the order of magnitude: Javorcik (2004a) found the effect 30 times greater. Moreover, following the methodology of Javorcik (2004a) and Blalock & Gertler (2008), many other studies conducted for different countries have found insignificant or even negative spillover effects. But despite the striking heterogeneity in the literature, no systematic survey has been done.

The theoretical model of Rodriguez-Clare (1996) implies that spillovers to host-country suppliers increase with larger communication costs between the foreign affiliate and its headquarters, and decrease with greater differences between the host and home countries in terms of the variety of intermediate goods produced. These implications are difficult to test empirically since spillover studies usually lack detailed information on the nationality of foreign investors. Using data from Romania, Javorcik & Spatareanu (2011) distinguish between European, American, and Asian investors; and show that investors from America and Asia generate more spillovers than investors from Europe, possibly because they are more likely to buy local inputs (imports from home countries are more expensive because of transportation costs and tariffs). To examine the predictions of Rodriguez-Clare (1996) more systematically, we take the advantage of 57 vertical spillover studies that provide estimates for many countries and different types of investors.

To take a step beyond single-country case studies and establish robust evidence for spillover effects, we employ the meta-analysis methodology (Stanley, 2001). Meta-analysis, the quantitative method of research synthesis, has been commonly used in economics for two decades (Card & Krueger, 1995; Smith & Huang, 1995; Card *et al.*, 2010). Recent applications of meta-analysis in international economics include Disdier & Head (2008) on the effect of distance on trade, Cipollina & Salvatici (2010) on reciprocal trade agreements, and Havranek (2010) on the trade effect of the euro. Meta-analysis is more than a literature survey: it sheds light on the determinants of the examined phenomenon that are difficult to investigate in primary studies because of data limitations.

In comparison with previous meta-analyses on productivity spillovers (Görg & Strobl, 2001; Meyer & Sinani, 2009), this paper concentrates on vertical instead of horizontal spillovers. We include many more estimates to investigate the full variability in the literature: 3,626 compared with 25 (Görg & Strobl, 2001) and 121 (Meyer & Sinani, 2009). To our knowledge, this makes our paper the largest meta-analysis conducted in economics so far. Moreover, the previous meta-analyses on spillovers used the reported t-statistics to evaluate the statistical significance of spillovers, whereas we use an economic measure of spillovers and employ new synthesis methods. Thus, we are able to estimate the net spillover effect beyond publication bias and misspecifications that are corrected by some studies.

The remainder of the paper is structured as follows: Section 2 briefly describes how spillovers are estimated and explains how we collected the estimates. Section 3 examines the extent of publication selection in the literature. Section 4 introduces variables which may explain heterogeneity in vertical spillovers. Section 5 examines how spillover estimates are affected by these variables, and quantifies the underlying effect beyond publication bias and misspecifications. Section 6 concludes.

2 The Spillover Estimates Data Set

Studies on foreign direct investment (FDI) spillovers usually examine the correlation between the productivity of domestic firms and their linkages with foreign affiliates.[1] With an allusion to the production chain, the linkages are usually classified into horizontal (within-sector: from FDI to local competitors) and vertical (between-sector); vertical linkages are further bifurcated into downstream (backward: from FDI to local suppliers) and upstream (forward: from FDI to local buyers). Most researchers use data from one country and estimate a variant of the following model, the so-called FDI spillover regression:

$$\ln \text{Productivity}_{ijt} = e_0^h \cdot \text{Horizontal}_{jt} + e_0^b \cdot \text{Backward}_{jt} + e_0^f \cdot \text{Forward}_{jt} + \boldsymbol{\alpha} \cdot \text{Controls}_{ijt} + u_{ijt}, \quad (1)$$

where i, j, and t denote firm, sector, and time subscripts, *Controls* denote a vector of either sector- or firm-specific control variables, and $\boldsymbol{\alpha}$ is the vector of the corresponding regression coefficients. The variable *Horizontal* is the ratio of foreign presence in firm i's own sector, *Backward* is the ratio of firm i's output sold to foreign affiliates, and *Forward* is the ratio of firm i's inputs purchased from foreign affiliates. Because firm-level data on linkages with foreign affiliates are usually unavailable the vertical linkages are computed at the sector level: *Backward* becomes the ratio of foreign presence in downstream sectors, *Forward* becomes the ratio of foreign presence in upstream sectors; the weight of each upstream or downstream sector is determined by the input-output table of the country.

The relative homogeneity of FDI spillover regressions allows us to meta-analyze the economic effect of spillovers. Since the response variable is in logarithm and linkage variables are ratios the estimates of coefficients e_0^h, e_0^b, and e_0^f can be interpreted as the semi-elasticities and thus constitute the natural common metric for the spillover literature. By construction it is assumed in most specifications that semi-elasticity is constant across different values of foreign presence. In meta-analysis semi-elasticity was previously used as a common metric by Rose & Stanley (2005) and Feld & Heckemeyer (2011), among others. In our case semi-elasticity is convenient for interpretation since it approximates the percentage increase in the productivity of domestic firms following an increase in the foreign presence of one percentage point:

$$e_0 \approx (\%\text{ change in productivity})/(\text{change in foreign presence}), \quad \text{foreign presence} \in [0, 1]. \quad (2)$$

For instance, the estimate $e^b = 0.1$ implies that a 10-percentage-point increase in foreign presence is associated with a 1% increase in the productivity of domestic firms in upstream sectors.

[1] See Smeets (2008) and Keller (2009) for recent surveys of the broader literature on international technology diffusion.

The estimates are directly comparable across studies that use the log-level specification. Estimates from studies that define foreign presence on the interval [0, 100] instead of [0, 1] are normalized. Within this basic framework, however, researchers use different methodologies and data sets, which cause substantial differences in results. We address these differences in Section 4 by introducing variables that capture method and structural heterogeneity.

The term "spillover" is overused in the literature; both horizontal and vertical semi-elasticities in (1) also capture effects other than knowledge externalities. As for horizontal linkages, the entry of foreign companies can lead to greater competition in the sector. Greater competition can either increase (through reducing inefficiencies) or decrease (through reducing market shares) the productivity of domestic firms. Neither case represents a knowledge transfer, and the coefficient e_0^h thus captures the net effect of knowledge spillovers and competition on productivity. As for vertical linkages, in the supplier-customer relationship the recipient of knowledge is clearly identifiable, and foreigners may be able to internalize the resulting benefits for suppliers (Blalock & Gertler, 2008; Keller, 2009). Anecdotal evidence suggests that compensations may indeed occur, though usually in an indirect form. For instance, in transition countries multinational companies are known to be hard bargainers: the discounted price of inputs that they often require likely reflects the future assistance and considerable prestige associated with such orders. For simplicity, we follow the convention to call productivity semi-elasticities "spillovers." The key takeaway, however, is that even positive and economically significant estimates of semi-elasticities do not necessarily call for governments to subsidize FDI.

A vast majority of the recent studies on FDI spillovers concentrate on vertical linkages, and vertical linkages are also the main focus of this paper. The two meta-analyses on horizontal spillovers, however, could not have used the recently developed meta-analysis methods and did not attempt to estimate the spillover effect implied by the literature. For this reason, additionally we present a partial meta-analysis of horizontal spillovers. In the partial meta-analysis, we include only those semi-elasticities that are estimated in the same regression with vertical spillovers.

We employed the following strategy for literature search: After reviewing the references of literature surveys (Görg & Greenaway, 2004; Smeets, 2008; Meyer & Sinani, 2009) and a few recent empirical studies, we elaborated a baseline search query that was able to capture most of the relevant studies. The baseline search in EconLit yielded 108 hits. Next, we searched three other Internet databases (Scopus, RePEc, and Google Scholar) and added studies that were missing from the baseline search. Finally, we investigated the RePEc citations of the most influential study, Javorcik (2004a). The three steps provided 183 prospective studies, which were all examined in detail. The last study was added on 31 March 2010.

Studies that failed to satisfy one or more of the following criteria were excluded from the meta-analysis. First, the study must report an empirical estimate of the effect of vertical linkages on the measure of the productivity of domestic firms. Second, the study must define vertical linkages as a ratio. Third, the study must report information on the precision of estimates (standard errors or t-statistics), or authors must be willing to provide it. Most of the identified studies, although related to the FDI spillover literature, did not estimate vertical spillovers. We excluded a few studies that estimated vertical spillovers but did not define linkages as a ratio and thus could not be used to compute semi-elasticity (for example, Kugler, 2006; Bitzer *et al.*, 2008). We often had to ask the authors for sample means of linkage variables or for clarification of their methodology: about 20% of the studies could be included thanks to cooperation from the authors.[2] No study was excluded on the basis of language, form, or place of publication; we follow Stanley (2001) and rather err on the side of inclusion in all aspects of data collection. We therefore also use studies written in Spanish and Portuguese, Ph.D. dissertations, articles from local journals, working papers, and mimeographs; and control for study quality in the analysis. A detailed description of the studies included in the analysis, as well as the complete list of excluded studies (with reasons for exclusion) are available in an online appendix at `meta-analysis.cz/spillovers`.

Following the recent trend in meta-analysis (Disdier & Head, 2008; Doucouliagos & Stanley, 2009; Cipollina & Salvatici, 2010), we use all estimates reported in the studies. If we arbitrarily selected the "best" estimate from each study, we could introduce an additional bias, and if we used the average reported estimate, we would discard a lot of information. Because the coding of the literature involved the manual collection of thousands of estimates with dozens of variables reflecting study design, to eliminate errors both of us collected all data independently. The simultaneous data collection took three months and the resulting disagreement rate, defined as the ratio of data points that differed between our data sets, was 6.7% (of more than 200,000 data points). After we had compared the data sets, we reached a consensus for each discordant data point. The retrieved data set with details on coding for each study is available in the online appendix.

A few difficult issues of coding are worth discussing. To begin with, some studies (3.7% of the observations; for instance, Girma & Wakelin, 2007) use the so-called regional definition of vertical spillovers. Researchers using the regional definition approximate vertical linkages by the ratio of foreign firms in the region, without using input-output tables. Such an approach does not distinguish between backward and forward linkages. Because the results are interpreted as vertical productivity spillovers from FDI, we include them in the analysis but create a dummy variable for this aspect of methodology. Next, many researchers use more variables for the same type of spillover in one regression. For example, Javorcik (2004a) separately examines the effect of fully owned foreign affiliates and the effect of investments with joint foreign and domestic ownership. Since the distinction between those coefficients is economically important, we use both of them and create dummies for affiliates with full foreign ownership, partial ownership, and for more estimates of the same type of spillover taken from one regression. Finally, some studies report coefficients that cannot be directly interpreted as semi-elasticities. This concerns, most notably, specifications different from the log-level (1.7% of the observations); for these different specifications we evaluated semi-elasticity at sample means. Other studies use the

[2]We are grateful to Joze Damijan, Ziliang L. Deng, Adam Gersl, Galina Hale, Chidambaran Iyer, Molly Lesher, Marcella Nicolini, Pavel Vacek, and Katja Zajc-Kejzar for sending additional data, or explaining the details of their methodology, or both.

interactions of linkage variables with other variables, typically absorption capacity (7.2% of the observations). Instead of omitting those estimates, we evaluate the marginal effects of foreign presence at sample means and control for this aspect in the multivariate analysis.

The resulting data set includes 3,626 estimates of semi-elasticity taken from 57 studies. The median number of estimates taken from one study is 45, and for each estimate we codified 55 variables reflecting study design. To put these numbers into perspective, consider Nelson & Kennedy (2009), who review 140 meta-analyses conducted in economics. They report that a median analysis includes 92 estimates (the maximum is 1,592) taken from 33 primary studies and uses 12 explanatory variables (the maximum is 41).

The oldest study in our sample was published in 2002 and the median study in 2008: in other words, a half of the studies was published in the last three years, which suggests that vertical spillovers from FDI are a lively area of research. The whole sample receives approximately 400 citations per year in Google Scholar, which further indicates the popularity of FDI spillover regressions. The median time span of the data used by the primary studies is 1996–2002, and all the studies combined use almost six million observations from 47 countries. While we cannot exploit the full variability of these primary observations, we benefit from the work of 107 researchers that have analyzed these data thoroughly. The richness of the data sets and methods employed enables us to systematically examine the heterogeneity in results and to establish robust evidence for the effect of foreign presence on domestic productivity.

Several estimates of semi-elasticity do remarkably differ from the main population and remain so even after a careful re-checking of the data; a similar observation applies to the precision of the estimates (the inverse of standard error). Such extreme values, most of which come from working papers and mimeographs, might lead to volatile results and degrade the graphical analysis. To account for outliers, some other large meta-analyses use the Grubbs test (Disdier & Head, 2008; Cipollina & Salvatici, 2010). But because we use precision to filter out publication bias, outlying values in precision could also invalidate the results. Thus, to detect outliers jointly in semi-elasticity and its precision, we use the multivariate method of Hadi (1994). By this procedure, run separately for each type of spillover, 4.87% of the observations are identified as outliers. It is worth noting that some researchers argue for using all observations in meta-analysis (Doucouliagos & Stanley, 2009). Nevertheless, under the assumption that better-ranked outlets publish more reliable results, the estimates identified here as outliers are of lower quality compared to the rest of the sample,[3] and although in the remainder of the paper we report the results for the data set without outliers, the inclusion of outliers does not affect the inference.

[3]Studies that produce outliers have a significantly lower impact factor compared with the rest of the sample: the p-value of the t-test is 0.02 when the recursive RePEc impact factor is used. The advantage of the RePEc ranking is that it also includes working paper series; nevertheless, the results are similar when we use the Journal Citation Report (Thompson) impact factor, Scientific Journal Ranking (Scopus) impact factor, or eigenfactor score (`www.eigenfactor.org`).

3 The Importance of Publication Bias

A mean estimate reported in the literature will be a biased estimate of the true spillover if some results are more likely than others to be selected for publication. Publication selection bias, which has long been recognized as a serious issue in empirical economics research (De Long & Lang, 1992; Card & Krueger, 1995; Ashenfelter & Greenstone, 2004; Stanley, 2005), arises from the preference of editors, referees, or authors themselves for results that are statistically significant or consistent with the theory. Publication bias is likely to be stronger in areas with less theory competition, where a particular sign of estimates is inconsistent with any major theory; this hypothesis is supported empirically by Doucouliagos & Stanley (2008). Selection for significance amplifies this bias and creates a bias of its own every time the underlying effect is different from zero because the estimates with the wrong sign are less likely to be significant.

The consequences of publication selection differ at the study and literature levels. For a well-known example, consider the effect of currency unions on within-union trade: it may be reasonable for an individual researcher not to base discussion on negative estimates since they likely result from model misspecification (in other words, there is no major theory consistent with the negative effect of common currency on trade). If, however, all researchers discard negative estimates, but some report large positive estimates that are also due to misspecification, the average impression from the literature will be biased towards a greater positive effect. This is precisely what the recent meta-analyses find (Rose & Stanley, 2005; Havranek, 2010). Publication bias affects both narrative and quantitative literature surveys, but only the quantitative methods can identify the bias and estimate the true effect beyond.

While in the first meta-analysis of spillovers from FDI Görg & Strobl (2001) identified publication bias among horizontal spillovers, in the last decade the selection for significance or positive signs has been more likely among backward spillovers. The change is due to increased theory competition for horizontal spillovers after the skeptical study of Aitken & Harrison (1999) was published, and to the last decade's consensus that backward spillovers are more important than forward and horizontal, following Javorcik (2004a) and Blalock & Gertler (2008).

When the literature is free of publication bias the estimates of semi-elasticities will be randomly distributed around the true population effect, e_0. If, however, some estimates fall into the file drawer because they are insignificant or have an unexpected sign, the reported estimates will be correlated with their standard errors (Card & Krueger, 1995; Ashenfelter *et al.*, 1999):

$$e_i = e_0 + \beta_0 \cdot Se(e_i) + u_i, \tag{3}$$

where u_i is a normal disturbance term and β_0 measures the strength of publication bias. For instance, if a statistically significant effect is required, an author who has a small data set may run a specification search until the estimate becomes large enough to offset the high standard errors. Because specification (3) is heteroscedastic by definition, in practice it is usually estimated by weighted least squares (Stanley, 2005, 2008):

$$e_i/Se(e_i) = t_i = e_0/Se(e_i) + \beta_0 + \xi_i. \tag{4}$$

Specification (4), often called the "meta-regression," likewise has a convenient interpretation: if the true semi-elasticity (e_0) is zero and if only positive and significant estimates are reported, the estimated coefficient for publication bias (β_0) will approach two, the most commonly used critical value of the t-statistic. It follows that the estimates of β_0 that are close to two signal very serious selection efforts. Monte Carlo simulations and many recent meta-analyses suggest that this parsimonious specification is also effective in filtering out publication bias and estimating true semi-elasticity, e_0 (Stanley, 2008).

Since we use more estimates from each study, it is important to take into account that estimates within one study are likely to be dependent (Disdier & Head, 2008). Therefore, (4) is likely to be misspecified. A common remedy is to employ the mixed-effects multilevel model, which allows for unobserved between-study heterogeneity (Doucouliagos & Laroche, 2009; Doucouliagos & Stanley, 2009):

$$t_{ij} = e_0/Se(e_{ij}) + \beta_0 + \zeta_j + \epsilon_{ij}, \tag{5}$$

where i and j denote estimate and study subscripts. The overall error term now consists of study-level random effects (ζ_j) and estimate-level disturbances (ϵ_{ij}).

The meta-regression results reported in Column 1 of Table 1 suggest that all types of spillover are free of publication bias if unpublished studies are included together with published studies. This is surprising because publication bias has been found in most areas of economics research even for results collected from working papers (Doucouliagos & Stanley, 2008). If there was publication selection in journals and authors were rationally maximizing the probability of publication, they would likely polish even preliminary versions of their papers.

When we consider only estimates from studies published in refereed journals (Column 2), publication bias is detected for backward spillovers, but not for forward and horizontal spillovers. Although the test for backward spillovers is significant only at the 10% level (p-value = 0.055), the evidence for publication bias is solid considering that this test is known to have low power (Stanley, 2008) and that the magnitude of the coefficient is approximately 1.1, which signals strong selection efforts. An important finding is that the selection is more prominent among the results that are deemed to be more important (backward spillovers) than among the bonus results (forward and horizontal spillovers). Since the important results determine the main message of the study, they are more likely to be polished.

The average estimate of backward spillovers reported in journal articles reaches 0.88. When it is corrected for publication bias, Table 1 shows that the estimate falls to 0.18: publication bias exaggerates the average published estimate fivefold. This simple example shows how dangerous it is to ignore publication bias; therefore, we will correct for publication bias throughout the analysis.[4]

The estimated semi-elasticity beyond publication bias is consistently positive and significant across all specifications for vertical spillovers, but the semi-elasticity for horizontal spillovers is

[4]More discussion of publication bias, including additional evidence and robustness checks, is available in the working-paper version of this article in the online appendix.

Table 1: Test of publication bias and true effect

Backward Spillovers	All	Published	Homogeneous
Intercept (publication bias)	-0.0255	1.083*	-1.481
	(0.496)	(0.656)	(0.942)
$1/Se$ (effect beyond bias)	0.168***	0.178***	0.307***
	(0.0241)	(0.0295)	(0.0380)
Observations	1311	370	568
Studies	55	26	39
Forward Spillovers	All	Published	Homogeneous
Intercept (publication bias)	0.729	-0.437	1.657
	(0.776)	(1.033)	(1.632)
$1/Se$ (effect beyond bias)	0.0872***	0.258***	0.0669**
	(0.0287)	(0.0454)	(0.0288)
Observations	1030	241	591
Studies	44	19	30
Horizontal Spillovers	All	Published	Homogeneous
Intercept (publication bias)	0.363	0.512	0.818
	(0.295)	(0.498)	(0.500)
$1/Se$ (effect beyond bias)	0.00466	0.0137	0.000549
	(0.00722)	(0.00837)	(0.0127)
Observations	1154	305	471
Studies	52	27	37

Note: Standard errors in parentheses. Response variable: t-statistic of the estimate of semi-elasticity. Estimated by the mixed-effects multilevel model using restricted maximum likelihood. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

All = all estimates. Published = only estimates from studies published in refereed journals. Homogeneous = only estimates for which no adjustment was needed, which use the standard definition of spillover variables, and which come from firm-level panel-data studies.

consistently insignificant. To get a flavor of the the likely magnitudes of vertical spillovers before turning to more advanced analysis, we use a relatively homogeneous subset of data that consists only of estimates which come from firm-level panel-data studies, which use the standard definition of spillover variables, and for which no computation of the marginal effect was needed (Column 3 of Table 1). This preliminary estimate suggests that a 10-percentage-point increase in foreign presence is associated with a 3% increase in the productivity of domestic firms in upstream sectors, an effect still three times smaller than the simple average of estimates published in refereed journals. For domestic firms in downstream sectors the increase in productivity is only 0.7%.

Therefore, when we use all available studies and account for publication bias and unobserved heterogeneity, our results suggest that backward spillovers are economically important, forward spillovers are statistically significant but small, and horizontal spillovers are insignificant. Nevertheless, since these effects are averaged across all countries and methods, we need multivariate analysis to explain the vast differences in the reported effects. The reported effects may be systematically influenced by misspecifications or other quality aspects. In the following sections, focusing only on backward spillovers, we will estimate the spillover effect implied by best-practice methodology and describe the determinants of spillovers.

4 Why Spillover Estimates Vary

Our first aim is to test the implications of the theoretical model by Rodriguez-Clare (1996), which indicates that positive backward spillovers are more likely to occur when the costs of communication between the foreign affiliate and its headquarters are high and when the source and host country of FDI are not too different in terms of the variety of intermediate goods produced. We additionally investigate other possible sources of spillover heterogeneity suggested in the recent literature (Crespo & Fontoura, 2007; Smeets, 2008; Meyer & Sinani, 2009), although these are often connected to the Rogriguez-Clare mechanism as well. We label such real differences in underlying spillover coefficients *structural heterogeneity.* Since previous meta-analyses on horizontal spillovers, and economics meta-analyses in general, found that reported results were systematically affected by study design, our second aim is to find out how the use of different methods affects spillover estimates. These differences in reported estimates are labeled *method heterogeneity.* Our notation follows Disdier & Head (2008).

Table A1 in the Appendix presents the descriptions and summary statistics for all variables assumed to explain structural and method heterogeneity. Variables explaining method heterogeneity are divided into four blocks: data characteristics represent the properties of the data used, specification characteristics represent the basic design of the tested models, estimation characteristics represent the econometric strategy, and publication characteristics represent the differences in quality not captured by the data and method variables.

4.1 Structural Heterogeneity

The structural block includes five variables that are computed at the host-country level and three dummy variables that reflect the characteristics of FDI and domestic firms in the particular spillover regressions. For the country-specific variables, we select values from 1999, the median year of the data used in primary studies. This approach can be supported by three reasons: First, because of data limitations it is not feasible to construct the variables as study-specific averages over the data periods of the individual studies. Second, all the studies were published between 2002 and 2010, and most of them use short and similar data periods. Third, we are interested in the relative differences between countries. When studies pool together data for multiple countries in one spillover regression (there are two such studies), we use population-weighted values for all variables.

As suggested by Rodriguez-Clare (1996), communication costs between the foreign affiliate and its headquarters can be approximated by the distance between the host and source countries of FDI, and country similarity can be approximated by the difference in the level of development. Both implications have an intuitive interpretation: On the one hand, investors from distant countries are likely to use more local inputs since it is expensive for them to import inputs from home countries; on the other hand, investors from much more developed countries are likely to use less local inputs since local firms are often unable to produce intermediate goods that would comply with the quality standards of the investors. A higher share of local inputs indicates more linkages with local firms and a greater potential for knowledge transfer.

To create a variable that would reflect the distance between the host country and its source countries of FDI, we need each country's geographic breakdown of inward FDI stocks—but such information is not always directly available. Therefore, as a first step, we use breakdowns of outward FDI positions of OECD countries provided by the OECD's International Direct Investment Statistics. In 1999, OECD countries accounted for more than 85% of the world stock of outward FDI. We additionally obtain breakdowns from the statistical offices of the next three most important source countries of FDI: Hong Kong, Taiwan, and Singapore, which increases the total coverage to 95%. Having information on the destination of 95% of all outward FDI stock in the world, we are able to reconstruct the breakdowns of inward FDI stock with high precision for all 47 countries that have been examined in the spillover literature.

It is necessary to take into account that some authors already separate the linkage effects of investors of different nationalities; for example, many studies on China separate ethnic Chinese investors (Hong Kong, Macao, Taiwan) from Western investors. Hence, we use three different breakdowns for China: the first for all investors, the second for Western investors, and the third for ethnic Chinese. Similarly, Javorcik & Spatareanu (2011) use separate linkage variables for European, American, and Asian investors to examine backward spillovers to Romanian firms.

The data on distances come from the CEPII database (`www.cepii.org`) and are computed following the great circle formula. The distance variable is then calculated using inward FDI breakdowns as weights. For example, if 70% of inward FDI stock in Mexico originated in the USA, 20% in Germany, and 10% in Korea, the average distance of foreign affiliates in Mexico from their headquarters would be $0.7 \cdot 1{,}600 + 0.2 \cdot 9{,}500 + 0.1 \cdot 11{,}700 = 4{,}190$ kilometers. We employ a similar approach to calculate the average technology gap of host countries with respect to the stock of inward FDI, measuring the development of the country as GDP per capita. The source of the data, similar to all remaining country-specific variables with the exception of patent rights, is the World Bank's World Development Indicators.

Another important determinant of spillovers is the international experience of domestic firms, which we approximate by the trade openness of the country. Firms with international experience may benefit more from backward linkages since they are used to trading with foreign firms and, for example, have employees with the necessary language skills. Such firms have a higher capacity to absorb spillovers. Firms exposed to international competition are also more likely to produce intermediate goods required by foreign affiliates, and hence, in line with the Rodriguez-Clare mechanism, benefit from greater spillovers.

As a major precondition of positive spillovers, many researchers stress the financial development of the host country (Javorcik & Spatareanu, 2009; Alfaro *et al.*, 2010): if domestic firms have difficulty obtaining credit, they react rigidly to the demand of foreign affiliates, and the sluggish response can result in fewer linkages. On the other hand, if the inflow of FDI eases the existing credit constraints of domestic firms by bringing in scarce capital (Harrison *et al.*, 2004), better credit terms reflect in higher productivity, and the benefits of FDI are more important in countries with tougher credit constraints. We approximate the development of the financial system by the ratio of private debts to GDP.

Countries with weak protection of intellectual property rights are likely to attract relatively low-technology investors (Javorcik, 2004b). If a smaller technology gap contributes to more linkages because of the Rodriguez-Clare mechanism, then the effect of weak intellectual property protection on spillovers may be positive. To approximate the protection of intellectual property, we choose the Ginarte-Park index of patent rights; the source of the data is Walter G. Park's website and Javorcik (2004b). The index is calculated once every five years, and values for 1999 are unavailable. Because Javorcik (2004b) computed the 1995 index for most of the originally missing transition countries that we need, we use the values for 1995.[5]

The other structural variables are dummies capturing the degree of foreign ownership used to define foreign presence or the investigated sector of the domestic economy. Many researchers argue that fully owned foreign affiliates create fewer spillovers compared with joint foreign and domestic projects (Javorcik & Spatareanu, 2008) since joint projects will arguably use technology that is more accessible to domestic firms. Some authors estimate spillovers separately for service sectors, which allows us to test the hypothesis that firms in services, compared with manufacturing firms, are less likely to benefit from linkages. Firms in services may lack international experience since they exhibit lower export propensity.

4.2 Method Heterogeneity

Data characteristics Following Görg & Strobl (2001) we include dummy variables for cross-sectional data and aggregation at the sector level, even though more than 90% of the estimates come from firm-level panel-data studies. Because the size of data sets used by primary studies varies substantially, we control for the number of years and firms to find out whether smaller studies report systematically different outcomes. We include the average year of the data period to control for possible structural changes in the effects of FDI. Finally, because a large part of studies on European countries use data from the same source (the Amadeus database), we include a corresponding dummy variable.

Specification characteristics We construct dummies for the inclusion of forward and horizontal spillover variables in the same regression, the proxy for foreign presence (most studies use share in output, others in employment or equity), the subset of firms used for the estimation of spillovers (whether all firms or only domestic are included), the inclusion of important control variables (sector competition and demand in downstream sectors), the control for absorption capacity, and the use of a lagged, instead of a contemporaneous, linkage variable.

Estimation characteristics Although the majority of studies use total factor productivity (TFP) as the measure of productivity, some estimate spillovers in one step using output, value

[5] If we used values of Ginarte-Park index for year 2000 or the Rule of Law measure reported by Political Risk Services instead, the results would remain similar.

added, or labor productivity as the response variable. When computing TFP, most authors take into account the endogeneity of input demand and use the Levinsohn-Petrin or Olley-Pakes method, but 10% of all estimates are computed using OLS. In the second step, TFP is regressed on the linkage variable, and the estimation is usually performed using firm fixed effects. We create dummies for random effects and pooled OLS as well as for the inclusion of year and sector fixed effects. Approximately a half of the regressions are estimated in differences. A general-method-of-moments (GMM) estimator is employed by 9% of the regressions, and the translog production function instead of the Cobb-Douglas function is employed by 8% of them.

Publication characteristics To control for the different quality of studies, we include a dummy for publication in refereed journals, the recursive RePEc impact factor of the outlet, the number of Google Scholar citations of the study discounted by study age, and the number of RePEc citations of the co-author who is most frequently cited. We also include a dummy variable for studies where at least one co-author is "native" to the examined country; we consider authors to be native if they either were born in the examined country or obtained an academic degree there. Such researchers are more familiar with the data used; on the other hand, they may have vested interests in the results. To account for any systematic difference between the results of researchers affiliated in the USA (for our sample it usually means highly ranked institutions) and elsewhere, we add a dummy for studies where at least one co-author is affiliated with a US-based institution. Finally, publication date (year and month) is included to capture the publication trend: possibly the advances in methodology that are difficult to codify in any other way.

Although we have additionally codified other variables reflecting methodology (among others the degree of aggregation of the linkage variable and the number of input-output tables used), the variation in these variables is too low to bring any useful information.

5 Multivariate Meta-Regression

To investigate the pattern of heterogeneity in the spillover literature, we add the explanatory variables assumed to explain structural and method heterogeneity into (3), and again as in Section 3 divide the resulting equation by the standard error to correct for heteroscedasticity and add the random-effects component to account for within-study dependence. The multivariate meta-regression then takes the following form (Doucouliagos & Stanley, 2009; Cipollina & Salvatici, 2010):

$$t_{ij} = \beta_0 + e_0/Se(e_{ij}) + \boldsymbol{\beta}\mathbf{x}'_{ij}/Se(e_{ij}) + \zeta_j + \epsilon_{ij}, \tag{6}$$

where $\mathbf{x}_{ij} = (x_{1ij}, \ldots, x_{pij})$ is the vector of explanatory variables listed in Table A1, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ is the vector of the corresponding regression coefficients, and the exogeneity assumptions are $\zeta_j | Se(e_{ij}), \mathbf{x}_{ij} \sim N(0, \psi)$ and $\epsilon_{ij} | Se(e_{ij}), \mathbf{x}_{ij}, \zeta_j \sim N(0, \theta)$. β_0 measures the incidence of publication selection, and e_0 represents the true effect, corrected for publication bias, in the reference case ($\mathbf{x}_{ij} = \mathbf{0}$): that is, e_0 is conditional on the values of other variables in the meta-regression, $\mathbf{x}$.

The high degree of unbalancedness of the data makes reliable testing of the exogeneity assumptions difficult.[6] Hence, as a specification check, meta-analysts usually employ OLS with clustered standard errors (Disdier & Head, 2008; Doucouliagos & Laroche, 2009). The principal problem with OLS in meta-analysis is that it gives each estimate the same weight, which causes studies reporting lots of estimates to become overrepresented. The mixed-effects multilevel model, on the other hand, gives each study approximately the same weight if the between-study heterogeneity is high (Rabe-Hesketh & Skrondal, 2008, p. 75). Yet large differences between the estimates based on OLS and on mixed effects may signal a violation of the exogeneity assumptions, and we therefore report both models, although the mixed-effects model is preferred. (For all specifications in our analysis, the significance of between-study heterogeneity is confirmed by likelihood-ratio tests at the 1% level.)

We begin by including all explanatory variables into the regression; this general model is not reported, but is available on request. To obtain a more parsimonious model, we employ the Wald test and exclude the method variables that are jointly insignificant at the 10% level, but keep all structural variables. The results for structural variables are reported in Table 2; the significant method variables are included in all regressions (the results for method variables are reported in Table 3). All structural variables are included in the specification reported in Column 1; the specifications in Columns 2 and 3 omit some of them to avoid the relatively high correlations (the highest one reaches 0.68), but the coefficients do not change a lot. The results are similar even if the effects of the country-specific variables are examined one by one in separate regressions.

There are two structural variables that are individually insignificant, and they are also jointly insignificant with the previously excluded method variables. Omitting all jointly insignificant variables yields our preferred "specific" model; that is, the model without redundant variables. The specific model is then re-estimated using OLS with standard errors clustered at the study level. Although three structural variables become less significant using OLS (their new p-values range between 0.1 and 0.2), the coefficients for all structural and control variables retain the same sign, which indicates that the mixed-effects model is correctly specified. Moreover, two of the three less significant structural variables become significant at standard levels when country-level instead of study-level clustering is used for OLS. The pseudo R^2s of about 0.4 show that a lot of heterogeneity still remains unexplained. But such values are common for large meta-analyses because of the microeconomic nature of the data (see, for instance, Disdier & Head, 2008). All of the qualitative results are robust to the inclusion of outliers.

[6]Fixed effects in the panel-data sense are generally inappropriate for meta-analysis since some studies report only one usable estimate; additionally, fixed effects make it impossible to examine the effect of study-level explanatory variables. As Nelson & Kennedy (2009, p. 358) put it: "The advantages of random-effects estimation [in meta-analysis] are so strong that this estimation procedure should be employed unless a very strong case can be made for its inappropriateness."

5.1 Structural Heterogeneity

Our most important finding concerns the effects of the nationality of foreign investors on the magnitude of backward spillovers. The distance between the host and source country of FDI has a robustly positive and significant effect, which suggests that investors from far-off countries create ceteris paribus more beneficial linkages. We thus corroborate the findings of Javorcik & Spatareanu (2011), who report that American and Asian investors in Romania generate greater spillovers than European investors. Furthermore, our results indicate that a high technology gap between foreign affiliates and domestic firms impedes knowledge transfer. Since, however, a very low or even negative technology gap may leave little room for knowledge transfer, we also test for a possible quadratic relationship between spillovers and the technology gap. Contrary to the recent meta-analysis on horizontal spillovers by Meyer & Sinani (2009), who use host-country-level data for GDP as a proxy of the technology gap and do not account for the difference between the host and source country, the quadratic term is insignificant and the linear specification fits the data better.

Table 2: Structural heterogeneity in backward spillovers

	Mixed-effects multilevel				OLS
	Full	Subset 1	Subset 2	Specific	Specific
Distance	0.247***	0.258***		0.249***	0.217***
	(0.0538)	(0.0520)		(0.0536)	(0.0671)
Technology gap	-0.513***		-0.462***	-0.386***	-0.370***
	(0.141)		(0.0880)	(0.103)	(0.131)
Openness	0.441***	0.646***		0.409***	0.266
	(0.125)	(0.0997)		(0.122)	(0.192)
Financial development	-0.344***		-0.591***	-0.339***	-0.219
	(0.122)		(0.0956)	(0.121)	(0.167)
Patent rights	-0.0673	0.0250			
	(0.0514)	(0.0334)			
Fully owned	-0.203***		-0.209***	-0.216***	-0.281***
	(0.0602)		(0.0603)	(0.0566)	(0.0946)
Partially owned	0.0203	0.0804	0.0227		
	(0.0561)	(0.0535)	(0.0564)		
Services	-0.220***	-0.234***	-0.220***	-0.222***	-0.387
	(0.0766)	(0.0771)	(0.0772)	(0.0765)	(0.350)
Pseudo R^2	0.39	0.36	0.38	0.40	0.46
Observations	1308	1308	1311	1311	1311
Studies	55	55	55	55	55

Note: Standard errors in parentheses. Response variable: t-statistic of the estimate of semi-elasticity.
All explanatory variables are divided by the standard error of the estimate of semi-elasticity.
OLS = ordinary least squares with clustered standard errors. The "Specific" model contains only variables that are jointly significant. The intercept, precision, and method variables are included in all specifications (these results are reported in Table 3). *** denotes significance at the 1% level.

We find that firms in countries open to international trade benefit more from FDI, which corresponds to Meyer & Sinani (2009). Thus both horizontal and vertical spillovers seem to be especially important for firms with international experience. On the other hand, the financial development of the host country has a negative effect on spillovers, which supports the view that foreign affiliates help domestic firms ease credit constraints. Indeed, according to the survey evidence reported by Javorcik & Spatareanu (2009) for the Czech Republic, a quarter of suppliers of foreign affiliates claimed that the supplier status helped them to gain more financing.

The results suggest that the protection of intellectual property rights is insignificant for the magnitude of spillovers. On the other hand, the degree of foreign ownership of investment projects is important. The dummy variable for investments with full foreign ownership is consistently negative and significant, suggesting that projects with full foreign ownership generate lower spillovers than projects with partial ownership (according to the specific model the semi-elasticity is lower by about 0.22). The coefficient for the variable capturing partial ownership is positive but insignificant; the insignificance is, however, largely due to the connection with the variable capturing full foreign ownership. When we drop the variable for full ownership from the regression (Column 2 of Table 2) the p-value corresponding to the variable for partial ownership decreases to 0.13. These findings are consistent with the negative effect of the technology gap on spillovers: fully owned foreign affiliates are likely to use more advanced technology, which increases the technology gap. Likewise, the smaller effect on domestic firms in service sectors is consistent with the importance of international experience for the adoption of spillovers.

Our results are in line with the theoretical predictions of Rodriguez-Clare (1996). To illustrate the economic significance of the effects of distance and the technology gap on spillovers, we quantify the implied spillover to Mexican firms generated by FDI from three different source countries: the United States, Germany, and South Korea. We use the results of (6) reported in Table 2 a Table 3, plug in the values of trade openness and financial development for Mexico and the bilateral values of distance and technology gap, and set all other variables in the regression to their sample means. Next, we evaluate the implied estimate of e_0 using the `lincom` command in Stata.

The model suggests that the greatest spillovers are generated by Korean FDI (1.07) followed by German FDI (0.51); investments from the nearby USA generate the least spillovers (−0.13). All these estimates are significant at the 5% level. Since Mexico has a similar technology gap with respect to the USA and Germany, the difference between the estimated spillover effects, 0.64, is largely due to different distances. Likewise, the distance from Mexico to Germany is similar to the distance from Mexico to Korea, and the difference in spillovers, 0.56, is due to different technology gaps. It follows that, under realistic conditions, the origin of FDI is economically important for the effect on domestic firms.

5.2 Method Heterogeneity

Table 3 shows that seventeen variables reflecting the characteristics of the data, specification, estimation, and quality are significant, suggesting that results depend on study design in a systematic way. The results are affected by the level of aggregation, age, and source of the data. The omission of the standard control variables (sector competition, downstream demand), the definition of the response variable, and the method of computing TFP matter. Furthermore,

Table 3: Method heterogeneity in backward spillovers

	Mixed-effects multilevel				OLS
	Full	Subset 1	Subset 2	Specific	Specific
Intercept	0.397	0.242	0.339	0.385	0.670**
	(0.375)	(0.396)	(0.378)	(0.371)	(0.298)
$1/Se$	2.785*	-2.890***	4.250***	1.293	1.554
	(1.643)	(0.523)	(0.952)	(1.190)	(1.563)
Data characteristics					
Aggregated	1.206***	1.213***	1.224***	1.193***	1.187***
	(0.145)	(0.140)	(0.145)	(0.144)	(0.190)
Average year	0.0349***	0.0236***	0.0277***	0.0323***	0.0301***
	(0.00789)	(0.00719)	(0.00754)	(0.00763)	(0.00837)
Amadeus	-0.686***	-0.489***	-0.861***	-0.680***	-0.603***
	(0.0950)	(0.0855)	(0.0874)	(0.0946)	(0.127)
Specification characteristics					
Employment	-0.168*	-0.149*	-0.131	-0.158*	-0.323*
	(0.0929)	(0.0825)	(0.0930)	(0.0921)	(0.171)
Competition	-0.315***	-0.353***	-0.368***	-0.333***	-0.306***
	(0.0673)	(0.0664)	(0.0655)	(0.0649)	(0.106)
Demand	0.567***	0.487***	0.581***	0.596***	0.615***
	(0.0995)	(0.0985)	(0.0944)	(0.0967)	(0.192)
Estimation characteristics					
One step	-0.348***	-0.302***	-0.304***	-0.353***	-0.447***
	(0.0783)	(0.0788)	(0.0779)	(0.0780)	(0.137)
Olley-Pakes	-0.318***	-0.305***	-0.324***	-0.346***	-0.464***
	(0.0824)	(0.0827)	(0.0802)	(0.0794)	(0.154)
OLS	-0.388***	-0.349***	-0.354***	-0.400***	-0.587***
	(0.102)	(0.102)	(0.102)	(0.101)	(0.173)
Pooled OLS	0.155***	0.174***	0.150***	0.155***	0.221***
	(0.0430)	(0.0430)	(0.0433)	(0.0430)	(0.0429)
Sector fixed	0.119***	0.140***	0.135***	0.128***	0.117*
	(0.0401)	(0.0380)	(0.0393)	(0.0393)	(0.0617)
Differences	0.107*	0.0415	0.0211	0.0989*	0.0583
	(0.0578)	(0.0568)	(0.0543)	(0.0569)	(0.0674)
Publication characteristics					
Published	0.276***	0.273***	0.274***	0.283***	0.407***
	(0.0786)	(0.0798)	(0.0777)	(0.0782)	(0.0958)
Study citations	0.0799**	0.0878***	0.108***	0.0820**	0.0421
	(0.0324)	(0.0323)	(0.0320)	(0.0322)	(0.0281)
Native	0.449***	0.466***	0.389***	0.461***	0.449***
	(0.0626)	(0.0634)	(0.0562)	(0.0617)	(0.0522)
Author citations	-0.0682***	-0.0574***	-0.0752***	-0.0739***	-0.0266
	(0.0190)	(0.0152)	(0.0185)	(0.0184)	(0.0214)
Publication date	0.0669**	0.0476**	0.105***	0.0756***	0.0503
	(0.0270)	(0.0239)	(0.0252)	(0.0261)	(0.0351)
Pseudo R^2	0.39	0.36	0.38	0.40	0.46
Observations	1308	1308	1311	1311	1311
Studies	55	55	55	55	55

Note: Standard errors in parentheses. Response variable: t-statistic of the estimate of semi-elasticity. All explanatory variables are divided by the standard error of the estimate of semi-elasticity. OLS = ordinary least squares with clustered standard errors. The "Specific" model contains only variables that are jointly significant. Structural variables are included in all specifications (these results are reported in Table 2). ***, **, and * denote significance at the 1%, 5%, and 10% levels.

we find an upward trend in the results: other things equal, the use of new data increases the reported semi-elasticity by 0.03 each year. Concerning quality characteristics, unpublished studies report estimates that are systematically lower by 0.28 compared with published studies; frequently cited studies also report higher spillovers.

As documented earlier in Section 2, most researchers assume that semi-elasticity is constant across different values of foreign presence. In other words, an increase in foreign presence from 0% to 10% is assumed to have similar effect on domestic producitivity as an increase from 90% to 100%; the impact of FDI is linear. To test the soundness of this assumption we would ideally need data on mean foreign presence for each specification, but in many studies this information is not provided. Nevertheless, we have information on mean FDI penetration for each country in our sample (measured by the ratio of inward FDI stock to GDP). If the estimated semi-elasticity was systematically affected by countries' FDI penetration, the assumption would likely be unrealistic. When we add FDI penetration variable to the general model, however, the variable is insignificant individually (p-value = 0.44) and also jointly with all other excluded variables. Therefore, we found no evidence of the nonlinearity of spillovers.

The results of the multivariate meta-regression can be used to estimate the underlying true semi-elasticity conditional on study design. We label this approach, which makes use of the estimated meta-regression, spillover estimation based on "best-practice" methods. Best practice, however, is subjective as different researchers may prefer different methodologies. We define best practice following Javorcik (2004a), the study published in the American Economic Review. There are two main reasons for such selection. First, the paper was published in the most selective journal and has the highest number of citations, both total and per-year, of all studies in our sample and is thus the natural benchmark for this literature. Second, the preferred model of Javorcik (2004a) is free of all method choices that are considered misspecifications by the majority of researchers. She uses firm-level data (as opposed to data aggregated at the sector level), computes TFP by a method that accounts for the endogeneity of input demand (as opposed to simple OLS), estimates the regression in differences, and controls for sector fixed effects, sector competition, and demand in downstream sectors.

We further extend the definition of best practice to synthesize an "ideal" study. We prefer results from peer-reviewed studies and plug in sample maxima for study citations, author citations, and average year of the data. Other variables, including all structural variables, are set to their sample means. In other words the best-practice estimate is conditional on some characteristics of methods and quality, but it is an average over all countries and sectors—roughly speaking, as if we took all six million observations used by the studies in our sample and employed the methods of Javorcik (2004a) to estimate the magnitude of backward spillover. Such defined best-practice estimate of the underlying semi-elasticity, e_0, reaches 0.94 and is significant at the 1% level with the 95% confidence interval (0.66, 1.21). The whole procedure yields similar results when outliers are included (1.00) or when OLS is used (0.94).[7]

[7] A similar multivariate analysis, available on request, shows that no country-specific variable matters for the degree of forward spillovers, and that the best-practice estimate of forward spillovers is insignificant. These findings corroborate the view that backward linkages are more important than forward linkages.

Therefore, beyond publication bias and observable misspecifications, our preferred estimate implies that a 10-percentage-point increase in foreign presence is associated with an increase in the productivity of local suppliers of about 9%: a large, economically important effect. The estimate further increases to 1.14 if we plug in the sample maximum of publication date. On the other hand, the use of output instead of TFP as the response variable (e.g., Blalock & Gertler, 2008) lowers the estimate from 0.94 to a still highly significant 0.58. When all variables reflecting quality characteristics are set to their sample means, the best-practice estimate declines from 0.94 to 0.73. When additionally average data characteristics are considered, the estimate further diminishes to 0.62. Finally, when average specification and estimation characteristics are also plugged in, the estimate shrinks to 0.02 and loses significance at conventional levels.

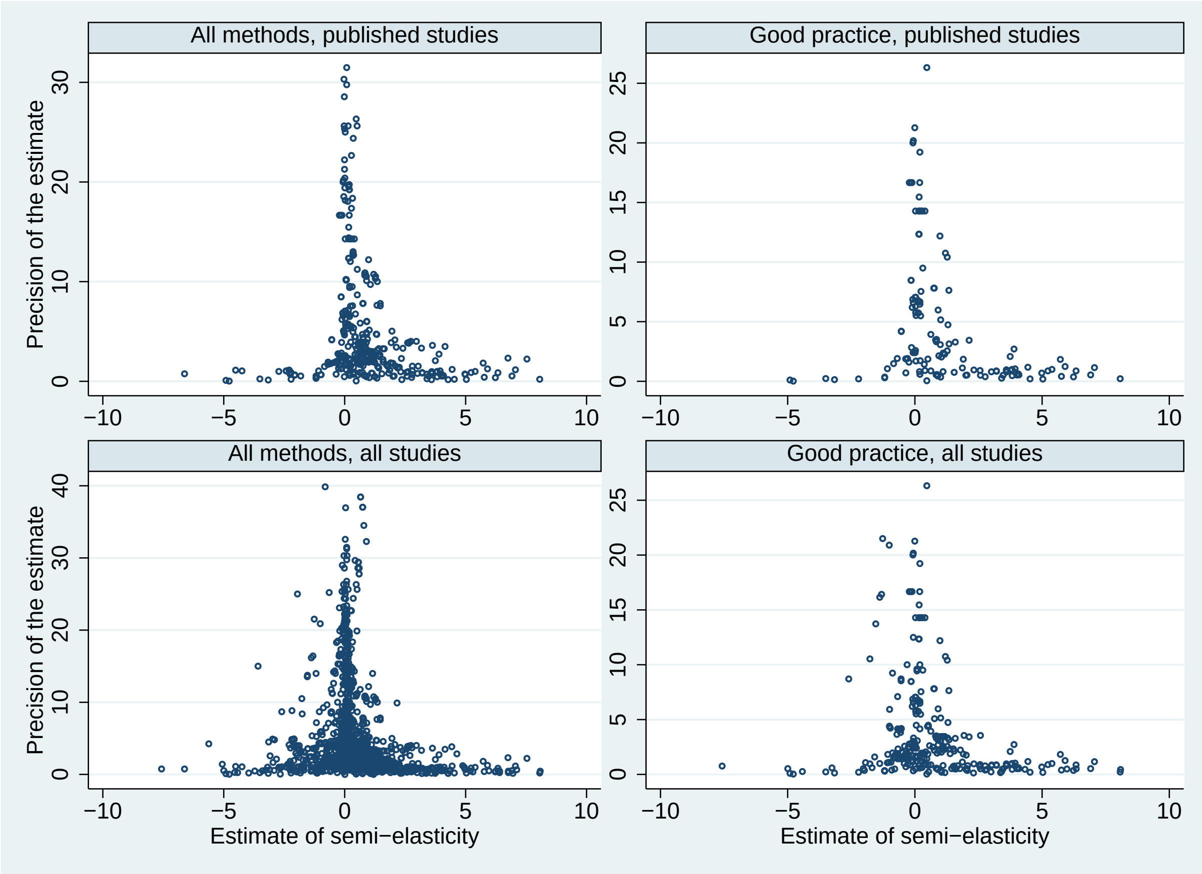
A mirror image of the best-practice estimation (the only exception is that firm-level data are still considered) even gives a significantly negative estimate, −0.42. Our analysis thus suggests that negative estimates are largely due to misspecifications. Indeed, the best-practice estimates are positive and significant for all countries in the sample even if we consider the effect of fully owned foreign affiliates on domestic firms in service sectors.

In Section 3 we found that estimates published in refereed journals are exaggerated because of publication selection. Now we have found that, in general, papers using better methods produce higher estimates of spillovers. The reader might wonder how the publishing filter works—are some results more likely to be selected for publication because they are positive and significant, or is it the selection of better methods that indirectly pushes the average reported estimate upwards? In the remainder of this section we will argue that the publishing filter is dominated by selection for statistical significance and positive signs.

For the explanation of our argument it is useful to introduce a graphical tool commonly employed to detect publication bias: a funnel plot (Stanley & Doucouliagos, 2010). The funnel plot depicts the size of the estimates of spillovers on the horizontal axis against their precision (the inverse of standard error) on the vertical axis. While the most precise estimates are close to the true effect, the less precise ones are more dispersed; hence the cloud of estimates should resemble an inverted funnel. In the absence of publication bias the funnel is symmetrical since all imprecise estimates have the same chance of being reported. If the publishing filter was characterized by the selection of better studies that yield higher results, the funnel would move to the right for published estimates compared with the funnel for all estimates. Nevertheless, this is no reason for the funnel to become asymmetrical. Estimates should be still randomly distributed around the true effect, and in the size-precision plane they should form a symmetrical inverted funnel.

The funnel plot for estimates published in reviewed journals is depicted in the top-left panel of Figure 1 and is clearly asymmetrical: the negative estimates of backward spillovers are almost completely missing from journals. On the contrary, the funnel plot for all estimates (the bottom-left panel) is symmetrical. The test of the significance of β_0 in specification (3), estimated earlier in Table 1 of Section 3, can be interpreted as a test of the asymmetry of the funnel plot; it follows from rotating the axes of the plot and inverting the values on the new

Figure 1: Funnel plots show publication bias in published studies



Note: "Good practice" denotes semi-elasticities estimated using firm-level data, controlling for sector competition, using firm-level fixed effects, and taking into account the endogeneity of input demand.

horizontal axis. Thus both formal and visual tests suggest that only published results exhibit asymmetry.

But cannot the asymmetry arise if only some journals select papers for their better methods? Other journals (or authors submitting to that journals) might rely on intuition and discard estimates of backward spillovers that would turn out to be negative. Such mixed publishing filter could produce a funnel similar to the top-left panel of Figure 1. To support our argument that intuition is the driving force of publication selection, we will only depict estimates that comply with the most important aspects of best practice: using firm-level data, controlling for sector competition, using firm fixed effects, and taking into account the endogeneity of input demand (we label theses aspects of methodology "good practice").[8] If journals select these estimates for their good practice and not for positive signs, the funnel plot would be symmetrical. But the real funnel for published estimates (the top-right panel of Figure 1) is no less asymmetrical than in the case when coefficients estimated by any method were considered (the top-left panel).

[8]It is not feasible to use the full definition of best practice because only a small fraction of estimates comply with the full definition.

Finally, Stanley *et al.* (2008) suggest how to test formally whether some aspects of methodology are associated with publication selection. If the aspects of methodology that define best practice cause publication selection, their interactions with the explanatory variable in equation (3), the standard error, will be significant. When we add these interactions to our full model (6), at the 5% level merely one out of nine of these interactions is significant; they are insignificant when considered jointly. Similarly, adding the interaction of a publication dummy with a measure of publication bias to equation (3) shows that the upward bias among the good-practice set of estimates is four times larger for published studies than for unpublished studies.

All in all, our results suggest that publication selection in reviewed journals is dominated by discarding the negative estimates of backward spillovers. We showed that negative results are indeed likely to be wrong and that the net backward spillover is positive and large; thus, somewhat paradoxically, publication selection based on intuition is getting the average published estimate of backward spillover closer to the true effect. Nevertheless, if authors' (or editors' or referees') prior was incorrect, publication selection would lead to an exaggeration of spillovers. This is likely to be the case of the earlier literature on horizontal spillovers where publication bias was found by Görg & Strobl (2001).

6 Conclusion

In a meta-analysis of data from 47 countries we find robust evidence consistent with knowledge transfer from foreign investors to domestic firms in supplier sectors (backward spillovers), but no economically important effect on firms in customer sectors (forward spillovers) or in the same sector (horizontal spillovers). Similar to Görg & Strobl (2001), we detect publication bias in the literature: positive or significant estimates are more likely to be selected for publication. This upward bias is present only among the estimates of backward spillovers from journal articles; unpublished studies and estimates of forward and horizontal spillovers exhibit no selection. On the other hand, misspecifications tend to bias the results downwards. Our results suggest that intuition is the driving force of publication selection: negative estimates are less likely to be reported in journals, even if the researcher avoids all well-known misspecifications.

Taking into consideration publication and misspecification bias, our preferred estimate suggests that a 10-percentage-point increase in foreign presence is associated with an increase in the productivity of domestic firms in supplier sectors of about 9%. Greater spillovers seem to be generated by FDI from distant countries with slight technological advantages over domestic firms. The results are in line with the theoretical model of Rodriguez-Clare (1996) and, in the case of distance, corroborate the findings of Javorcik & Spatareanu (2011) for Romania. Greater spillovers seem to be received by countries that are open to international trade and that have underdeveloped financial systems. In addition, fewer spillovers are generated by fully owned foreign affiliates compared with joint ventures, and fewer spillovers are received by domestic firms in services compared with manufacturing.

Meta-analysis can only filter out misspecifications that have been overcome by a sufficient number of researchers. If a misspecification is shared by the entire literature and influences the estimates in a systematic way, meta-analysis will give biased results. (This problem is important for the point estimate of spillover while less so for spillover determinants.) In this respect, several researchers have emphasized that the traditional definition of linkage variables in spillover regressions is valid only under specific conditions. Concerning backward spillovers, Barrios *et al.* (2011) construct an alternative measure of linkages using, for example, input-output tables for investors' home countries to account for different sourcing behavior. Vacek (2010) constructs firm-level linkage variables that reflect the actual ratio of the output of domestic firms sold to foreign affiliates. Concerning horizontal spillovers, Keller & Yeaple (2009) use an instrumental-variable estimator and take into account that foreign affiliates are active in more than one sector. All of these studies find that using the new measures results in stronger evidence of positive spillovers. These improvements, however, have so far been sparsely applied, and we leave their examination in a meta-regression analysis for further research.

References

Aitken, B. J. & A. E. Harrison (1999): "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela." *American Economic Review* **89(3)**: pp. 605–618.

Alfaro, L., A. Chanda, S. Kalemli-Ozcan, & S. Sayek (2010): "Does foreign direct investment promote growth? Exploring the role of financial markets on linkages." *Journal of Development Economics* **91(2)**: pp. 242–256.

Ashenfelter, O. & M. Greenstone (2004): "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review* **94(2)**: pp. 454–460.

Ashenfelter, O., C. Harmon, & H. Oosterbeek (1999): "A review of estimates of the schooling/earnings relationship, with tests for publication bias." *Labour Economics* **6(4)**: pp. 453–470.

Barrios, S., H. Görg, & E. A. Strobl (2011): "Spillovers Through Backward Linkages from Multinationals: Measurement Matters!" *European Economic Review* **(forthcoming)**.

Bitzer, J., I. Geishecker, & H. Görg (2008): "Productivity spillovers through vertical linkages: Evidence from 17 OECD countries." *Economics Letters* **99(2)**: pp. 328–331.

Blalock, G. & P. J. Gertler (2008): "Welfare gains from Foreign Direct Investment through technology transfer to local suppliers." *Journal of International Economics* **74(2)**: pp. 402–421.

Card, D., J. Kluve, & A. Weber (2010): "Active labour market policy evaluations: A meta-analysis." *Economic Journal* **120(548)**: pp. F452–F477.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-analysis." *American Economic Review* **85(2)**: pp. 238–43.

Cipollina, M. & L. Salvatici (2010): "Reciprocal Trade Agreements in Gravity Models: A Meta-Analysis." *Review of International Economics* **18(1)**: pp. 63–80.

Crespo, N. & M. P. Fontoura (2007): "Determinant Factors of FDI Spillovers—What Do We Really Know?" *World Development* **35(3)**: pp. 410–425.

De Long, J. B. & K. Lang (1992): "Are All Economic Hypotheses False?" *Journal of Political Economy* **100(6)**: pp. 1257–72.

Disdier, A.-C. & K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics* **90(1)**: pp. 37–48.

Doucouliagos, C. & P. Laroche (2009): "Unions and Profits: A Meta-Regression Analysis." *Industrial*

Relations **48(1)**: pp. 146–184.

Doucouliagos, H. & T. Stanley (2008): "Theory Competition and Selectivity: Are All Economic Facts Greatly Exaggerated?" *Economics Series Working Paper 06*, Deakin University.

Doucouliagos, H. & T. D. Stanley (2009): "Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis." *British Journal of Industrial Relations* **47(2)**: pp. 406–428.

Feld, L. P. & J. H. Heckemeyer (2011): "FDI and Taxation: A Meta-Study." *Journal of Economic Surveys* **25(2)**: p. 233–272.

Girma, S. & K. Wakelin (2007): "Local productivity spillovers from foreign direct investment in the U.K. electronics industry." *Regional Science and Urban Economics* **37(3)**: pp. 399–412.

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

Görg, H. & E. Strobl (2001): "Multinational Companies and Productivity Spillovers: A Meta-analysis." *The Economic Journal* **111(475)**: pp. F723–39.

Hadi, A. S. (1994): "A Modification of a Method for the Detection of Outliers in Multivariate Samples." *Journal of the Royal Statistical Society, Series (B)* **56**: pp. 393–396.

Harrison, A. E., I. Love, & M. S. McMillan (2004): "Global capital flows and financing constraints." *Journal of Development Economics* **75(1)**: pp. 269–301.

Havranek, T. (2010): "Rose Effect and the Euro: Is the Magic Gone?" *Review of World Economics* **146(2)**: pp. 241–261.

Javorcik, B. S. (2004a): "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages." *American Economic Review* **94(3)**: pp. 605–627.

Javorcik, B. S. (2004b): "The composition of foreign direct investment and protection of intellectual property rights: Evidence from transition economies." *European Economic Review* **48(1)**: pp. 39–62.

Javorcik, B. S. & M. Spatareanu (2008): "To share or not to share: Does local participation matter for spillovers from foreign direct investment?" *Journal of Development Economics* **85(1-2)**: pp. 194–217.

Javorcik, B. S. & M. Spatareanu (2009): "Liquidity Constraints and Firms' Linkages with Multinationals." *World Bank Economic Review* **23(2)**: pp. 323–346.

Javorcik, B. S. & M. Spatareanu (2011): "Does it matter where you come from? Vertical spillovers from foreign direct investment and the origin of investors." *Journal of Development Economics* **(forthcoming)**.

Keller, W. (2009): "International Trade, Foreign Direct Investment, and Technology Spillovers." *NBER Working Papers 15442*, National Bureau of Economic Research.

Keller, W. & S. R. Yeaple (2009): "Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States." *The Review of Economics and Statistics* **91(4)**: p. 821–831.

Kugler, M. (2006): "Spillovers from foreign direct investment: Within or between industries?" *Journal of Development Economics* **80(2)**: pp. 444–477.

Meyer, K. E. & E. Sinani (2009): "When and where does foreign direct investment generate positive spillovers? A meta-analysis." *Journal of International Business Studies* **40(7)**: pp. 1075–1094.

Nelson, J. & P. Kennedy (2009): "The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment." *Environmental & Resource Economics* **42(3)**: pp. 345–377.

Rabe-Hesketh, S. & A. Skrondal (2008): *Multilevel and Longitudinal Modeling Using Stata.* College Station, TX: Stata Press.

Rodriguez-Clare, A. (1996): "Multinationals, Linkages, and Economic Development." *American Economic Review* **86(4)**: pp. 852–73.

ROSE, A. K. & T. D. STANLEY (2005): "A Meta-Analysis of the Effect of Common Currencies on International Trade." *Journal of Economic Surveys* **19(3)**: pp. 347–365.

SMEETS, R. (2008): "Collecting the Pieces of the FDI Knowledge Spillovers Puzzle." *World Bank Research Observer* **23(2)**: pp. 107–138.

SMITH, V. K. & J.-C. HUANG (1995): "Can Markets Value Air Quality? A Meta-analysis of Hedonic Property Value Models." *Journal of Political Economy* **103(1)**: pp. 209–27.

STANLEY, T. & H. DOUCOULIAGOS (2010): "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economic Surveys* **24(1)**: pp. 170–191.

STANLEY, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

STANLEY, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

STANLEY, T. D. (2008): "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

STANLEY, T. D., H. DOUCOULIAGOS, & S. B. JARRELL (2008): "Meta-regression analysis as the socio-economics of economics research." *The Journal of Socio-Economics* **37(1)**: pp. 276–292.

VACEK, P. (2010): "Panel Data Evidence on Productivity Spillovers from Foreign Direct Investment: Firm-Level Measures of Backward and Forward Linkages." *IES Working Paper 18/2010*, Institute of Economic Studies, Charles University, Prague.

A Data Description

Table A1: Summary statistics of regression variables, backward spillovers

Variable	Description	Mean	Std. dev.
t-statistic	The t-statistic of the estimate of semi-elasticity.	0.803	4.997
$1/Se$	The precision of the estimate of semi-elasticity.	5.465	6.640
Structural heterogeneity			
Distance	The logarithm of the country's FDI-stock-weighted distance from its source countries of FDI (kilometers).	7.769	0.621
Technology gap	The logarithm of the country's FDI-stock-weighted gap in GDP per capita with respect to its source countries of FDI (USD, constant prices of 2000).	9.816	0.419
Openness	The trade openness of the country: (exports + imports)/GDP.	0.704	0.330
Financial development	The development of the financial system of the country: (domestic credit to private sector)/GDP.	0.614	0.428
Patent rights	The Ginarte-Park index of patent rights of the country.	2.993	0.800
Fully owned	=1 if only fully owned foreign investments are considered for linkages.	0.069	0.253
Partially owned	=1 if only investments with joint domestic and foreign ownership are considered for linkages.	0.070	0.256
Services	=1 if only firms from service sectors are included in the regression.	0.046	0.209
Data characteristics			
Cross-sectional	=1 if cross-sectional data are used.	0.079	0.269
Aggregated	=1 if sector-level data for productivity are used.	0.033	0.178
Time span	The number of years of the data used.	7.090	3.788
Firms	The logarithm of [(the number of observations used)/(time span)].	7.598	2.040
Average year	The average year of the data used (2000 as a base).	-1.053	3.798
Amadeus	=1 if the Amadeus database by Bureau van Dijk Electronic Publishing is used.	0.223	0.416

Continued on next page

Table A1: Summary statistics of regression variables, backward spillovers (continued)

Variable	Description	Mean	Std. dev.
Specification characteristics			
Forward	=1 if forward spillovers are included in the regression.	0.655	0.475
Horizontal	=1 if horizontal spillovers are included in the regression.	0.866	0.341
Employment	=1 if employment is the proxy for foreign presence.	0.142	0.349
Equity	=1 if equity is the proxy for foreign presence.	0.060	0.238
All firms	=1 if both domestic and foreign firms are included in the regression.	0.252	0.435
Absorption	=1 if the specification controls for absorption capacity using technology gap or R&D spending.	0.070	0.256
Competition	=1 if the specification controls for sector competition.	0.272	0.445
Demand	=1 if the specification controls for demand in downstream sectors.	0.075	0.263
Regional	=1 if vertical spillovers are measured using the ratio of foreign firms in the region as a proxy for foreign presence.	0.037	0.188
Lagged	=1 if the coefficient represents lagged foreign presence.	0.127	0.334
More	=1 if the coefficient is not the only estimate of backward spillovers in the regression.	0.459	0.499
Combination	=1 if the coefficient is a marginal effect computed using a combination of reported estimates.	0.072	0.259
Estimation characteristics			
One step	=1 if spillovers are estimated in one step using output, value added, or labor productivity as the response variable.	0.429	0.495
Olley-Pakes	=1 if the Olley-Pakes method is used for the estimation of TFP.	0.187	0.390
OLS	=1 if OLS is used for the estimation of TFP.	0.107	0.309
GMM	=1 if the system GMM estimator is used for the estimation of spillovers.	0.089	0.285
Random	=1 if the random-effects estimator is used for the estimation of spillovers.	0.031	0.174
Pooled OLS	=1 if pooled OLS is used for the estimation of spillovers.	0.157	0.364
Year fixed	=1 if year fixed effects are included.	0.854	0.353
Sector fixed	=1 if sector fixed effects are included.	0.494	0.500
Differences	=1 if the regression is estimated in differences.	0.456	0.498
Translog	=1 if the translog production function is used.	0.076	0.266
Log-log	=1 if the coefficient is taken from a specification different from log-level.	0.017	0.128
Publication characteristics			
Published	=1 if the study was published in a refereed journal.	0.288	0.453
Impact	The recursive RePEc impact factor of the outlet. Collected in April 2010.	0.238	0.453
Study citations	The logarithm of [(Google Scholar citations of the study)/(age of the study) + 1]. Collected in April 2010.	1.160	1.110
Native	=1 if at least one co-author is native to the investigated country.	0.712	0.453
Author citations	The logarithm of (the number of RePEc citations of the most-cited co-author + 1). Collected in April 2010.	3.114	2.480
US-based	=1 if at least one co-author is affiliated with a US-based institution.	0.397	0.489
Publication date	The year and month of publication (January 2000 as a base).	7.865	1.637

Source of the data: World Development Indicators and `http://www1.american.edu/cas/econ/faculty/park.htm`.