A Meta-Analysis of FDI and Productivity Spillovers in Developing Countries

Journal name: *Journal of Economic Surveys*

Authors: Binyam Afewerk Demena*,**
Peter A.G. van Bergeijk*

* International Institute of Social Studies, Erasmus University Rotterdam, Kortenaerkade 12, 2518 AX, The Hague, The Netherlands

** Corresponding author: demena@iss.nl

Abstract

This meta-analysis reviews the intra-sector heterogeneity of productivity spillovers from foreign direct investment (FDI) in 31 developing countries through a larger more comprehensive dataset. We investigate how the inconsistencies in the reported spillover findings are affected by publication bias, characteristics of the data, estimation techniques and empirical specification, analyzing 1,450 spillover estimates from 69 empirical studies published in 1986 - 2013. Our findings suggest that reported FDI spillover estimates are affected by publication bias. In combination with model misspecification of the primary studies, the bias overstates the genuine underlying meta-effect, but the meta-effect remains economically and statistically significant. Our results emphasize that spillovers and their sign largely depend systematically on specification characteristics of the primary studies and publication bias. Publication bias is not caused by "best practice" choices. Future research needs to cover more developing countries and to investigate not only whether spillovers occur, but also to explore inside the black box of how spillovers actually emerge.

Keywords: meta-analysis, foreign direct investment, spillover, publication bias, genuine empirical effect, developing countries.

1. Introduction

Developing countries increasingly use policies to attract foreign direct investment (FDI) and often provide substantial incentives. Almost 90% of the national regulatory investment regime changes introduced in more than 100 countries in 1991-2012 offer more favourable conditions for FDI (UNCTAD, 2013). One driver behind the FDI-friendlier regimes is the expectation that FDI inflows indirectly boost productivity of domestic firms. FDI is assumed to transfer knowledge of processes and products (so-called productivity spillovers) and this may enhance technological capabilities of domestic firms (Wooster and Diebel, 2010).

An important question is whether FDI actually generates productivity spillovers in the context of developing countries: the empirical evidence is inconclusive and contradictory. Figure 1 illustrates the extent of empirical disagreement for the 74 primary studies that we collected using the Stanley et al. (2013) recommendations. These studies were published in 1983-2013 by 96 researchers, deal with 31 developing countries[1] and provide 1,545 estimates of spillover parameters. Figure 2 presents details of the distribution of these estimates. Approximately only one third of the empirical estimates validate a significantly positive effect. About half the spillover estimates are insignificant; whereas one in six find a significantly negative effect. It should be noted that while we refer to this field as findings regarding developing countries, these studies (and thus our findings) are not necessarily representative for the group of developing countries. The reason is clear: the productivity spillovers of FDI have not yet been investigated for a sufficiently large group of developing countries to make such claims.

<Figures 1 and 2>

Our article aims to estimate the size, sign and significance of the spillover effect and to analyse the determinants behind the heterogeneous results of the primary empirical studies. We demonstrate application of the reporting guidelines of the MAER-Net required by the *Journal of Economic Surveys* for conducting meta-analysis. We use meta-analysis to combine, summarize and investigate the reported spillover estimates. Meta-analysis is a statistical approach to analyse an existing literature of reported empirical findings for a similar hypothesis, research question, empirical effect and/or phenomenon (Stanley and Doucouliagos, 2012). Unlike qualitative reviews (e.g., Fan, 2002), meta-analysis enables us to filter out publication bias and misspecification and to estimate the genuine spillover effect.

To date six relevant meta-analyses were published: Görg and Strobl (2001), Meyer and Sinani (2009), Wooster and Diebel (2010), Iršová and Havránek (2013), Mebratie and Bergeijk (2013) and Demena (2015). We use an enriched version of Demena's (2015) dataset in this paper, covering more than 40 potential research dimensions (including data characteristics, estimation techniques and empirical specifications) and four categories of journal and study qualities. While we add many potentially explanatory variables to provide a comprehensive test, we also enforce limitations. In particular we focus on developing countries, because pooling developing and developed countries is inappropriate (Wooster and Diebel, 2010). Our meta-analysis comprises all observations that investigated intra-sector spillovers

from FDI for developing countries, providing a more comprehensive version of previous meta-analyses. This potentially leads to a better reliability of the findings and helps us to investigate whether the selection of studies has an impact on bias; a topic not researched in previous meta-analyses of FDI spillovers. Indeed, the latest studies suggested that publication bias was not a problem for this literature, but we uncover clear evidence of publication bias in our larger sample.[2] The spillover effects in the primary studies overstate the magnitude of the genuine empirical effect, although these effects remain economically important after controlling for selection bias and misspecification.

The remainder of the paper is structured as follows: Section 2 starts with the primary studies. Section 3 discusses the data and the empirical approach. Section 4 presents detailed results. Section 5 concludes and suggests lessons for future research.

2. Review of the Literature

Many studies deal with productivity spillovers from FDI.[3] We organize the literature by decades in order to illustrate changes over time in spillover findings and study characteristics.

2.1. 1980s: positive spillover effect in cross-sections

The first empirical study for developing countries was Blomström and Persson (1983). They study Mexican plants and report that FDI has a significantly positive effect on productivity of domestic firms. Blomström (1986) – also for Mexican firms – found a similar effect. These studies used industry-level cross-sectional data, and thus face an identification problem. Firstly, given the assumption that foreign firms are more productive and more technologically advanced, industry-level aggregated data that include foreign firms in the aggregate regression create an upward spillover effect - aggregation bias. Secondly, if FDI gravitates towards productive industries, then the observed spillover result from cross-sectional data will overstate the impact of foreign firms due to potential endogeneity of FDI. Thus, it is unclear whether the evidence is due to the presence of FDI or the own-productivity of domestic firms.

2.2. 1990s: Emerging Contradictory Evidence

An important study is the first use of panel firm-level data (Haddad and Harrison, 1993) with insignificant productivity spillover for Morocco. Aitken and Harrison (1999) report negative effects for Venezuela. The former attributed the absence of spillovers to the technological gap between domestic and foreign firms; the latter concluded that the positive effect reported in the previous studies was due to the tendency of foreign firms to invest in productive industries. In contrast, Kokko (1994, 1996) for Mexico, Blomström and Sjöholm (1999), Sjöholm (1999a, b) for Indonesia, and Chuang and Lin (1999) for Taiwan report positive productivity spillovers. These studies continued to rely on cross-sections although panel data are better to test the validity of spillovers enabling to control for the behaviour of firms over time.

2.3. 2000s: Continued Contradiction

The studies of the 1980s and 1990s use a pipeline model: they presume that spillovers are independent of domestic firms' capabilities. In the 2000s a shift occurs towards the domestic capability model. This new strand of the literature assumes that spillovers do not occur automatically, but may depend on the capabilities of domestic firms. As illustrated in Figure 1, the body of evidence increased substantially overtime, possibly due to growing availability of datasets. The findings, however, remain contradictory. Blyde et al. (2004) for Venezuela, Bwalya (2006) for Zambia, Marin and Bell (2006) and Chudnovsky et al. (2008) for Argentina, and Mebratie and Bedi (2013) for South Africa did not find any spillovers. Jordaan (2008a, b) for Mexico and Waldkirch and Ofosu (2010) for Ghana find negative FDI effects. In contrast to these findings, studies on Asian countries report positive spillovers (Taymaz and Yilmaz 2008 for Turkey, Nguyen 2008 for Vietnam, Khalifah and Adam, 2009 for Malaysia).

2.4. Trends

Figure 1 illustrates substantial variations in the reported spillover estimates over time with an increasing trend, in particular since the mid-2000s. Over the past three decades, the number of spillover estimates has continuously increased: about 71% of the estimates are from studies published after 2005.

Moreover, the initial results from studies using cross-sectional data were challenged by subsequent panel data studies. This suggests an association of cross-section studies and positive findings, and negative or insignificant findings and panel studies. Panel studies by Aitken and Harrison (1999) for Venezuela report negative effects, whereas Haddad and Harrison (1993) for Morocco, Blalock and Gertler (2008) for Indonesia, and Mebratie and Bedi (2013) for South Africa report no effects. However, recent panel studies (e.g., Kee, 2005 for Bangladeshi; Takii, 2009 for Indonesia; Van Thanh and Hoang, 2010 for Vietnam) find positive effects. It appears likely, therefore, that the dichotomy of cross-section and panel data findings has become less clear (see, Jordaan, 2012).

Importantly, the primary studies include a proxy for the extent of foreign-owned firms' presence. Theoretically, intra-sector spillovers are assumed to occur through the transmission channels of demonstration, labour mobility, competition and/or export effects (Javorcik, 2004; Crespo and Fontoura, 2007; Jordaan, 2012). Unfortunately, the existing empirical work fails to identify the channels underlying the spillovers and instead merely focuses on whether the presence of foreign firms influences the productivity of domestic firms. Thus, there is a gap between theory and empirical analyses.

Furthermore, the empirical studies largely ignore the heterogeneity of both R&D expenditure and the technological levels of domestic firms (only 10% of the studies in the meta-data control for these factors). Clearly, general spillovers cannot exactly indicate how spillovers occur and which domestic firms gain. Some may experience positive impacts, others nothing or even negative impacts. For example, firms with relatively higher technological level can benefit from spillovers via the competition and/or demonstration effects, while firms with lower technological level may not be

in a position to compete or imitate (Hamida, 2013) and may instead benefit from labor mobility.

Thus, the relative importance of the spillover channels varies with the existing firm specific characteristics of domestic firms. Consequently, future investigations of overall spillover effects need to discern the transmission channels through employing various spillover control variables. This would be important to accurately describe and identify the impact of spillover processes, and hence to narrow the heterogeneity of spillover estimates and also to guide policymakers towards relevant channels.

3. Data and Empirical Approach

In order to econometrically analyse the sources of heterogeneity and to understand if FDI generates spillovers, we reviewed existing meta-analyses and primary empirical studies (and their references) as the basis for our literature search. Our search and review for the existing empirical studies aimed to identify all relevant English language unpublished and published studies regarding developing countries.

3.1. Methods, Protocols and Data Construction

In identifying the relevant studies, coding variables and data construction, we follow the MAER-Net prescriptions (Stanley et al., 2013). We identified our relevant primary studies in an extensive search using Google scholar, EconLit and Scopus. The World Bank (2012) database that provides empirical studies conducted using the enterprise survey data was also examined in detail. The search included all potentially relevant published and unpublished empirical studies from 1983 to 2013.

We searched using the broad keywords: "FDI presence effect on host economy", "technology transfer+foreign firms", "productivity spillover+FDI", and "productivity spillover+FDI+developing countries". For example, the keyword "productivity spillover+FDI+developing countries" using the Scopus search engine provided 1,026 records to review. Examinations of titles, abstracts, and keywords were followed by the inspection of the introduction and conclusion. This yielded a list of 233 prospective studies.

Studies were included if they satisfy the following criteria for detailed review: English language empirical micro econometrics that study intra-industry[4] spillover effects and report regression-based[5] coefficients, sample size, t-statistics or standard errors. The imposition of these criteria resulted in our dataset of 74 studies dealing with 31 developing countries for coding. Before transferring the data to a Stata file for analysis, a template for data extraction was designed in excel format. Data on various characteristics of the empirical studies such as spillover measures and effects, data type, estimation techniques, and study control variables were collected. In this case, we have conducted extensive coding of the study characteristics in order to avoid subjectivity and increase the robustness and the reliability of the findings. A second reviewer independently checked the consistency of the data and coding with an initial data collection and coding disagreement rate of 2.9%. After double checking initial data, consensus was reached. In order to ensure comprehensiveness of the dataset, the multiple search process took five months (May – September 2013).

It is worthwhile to describe some critical aspects encountered during this process. Multiple estimates are a common standard in economics. This is partly due to the demand from editors and reviewers that applied econometric studies should report multiple models, methods and estimates to ensure robustness (Stanley and Doucouliagos, 2012). This may lead to a best-set, average-set or all-set meta-data.[6] Following Stanley (2001), mainly to evade giving undo weight to a single study, many meta-analysts use either the "best" estimate or the average estimate. In the latter case, however, it is impossible to analyse the impact of different estimation techniques, estimation samples and models. Moreover, we may lose important within study information if we use average estimates. Importantly, choosing the "best" estimate may introduce subjectivity. First, in most cases, authors do not explicitly indicate their best estimate. Second, if they do, author's preference may introduce selection bias. Third, estimates in a comprehensive single paper can be underweighted relative to estimates of researchers who publish a large number of closely related articles as each would count as an individual study to be included as a best-set estimate (Stanley, 2001). For these reasons, we adopt the all-set estimates.

The majority (86%) of the models are estimated in log-linear form, with productivity proxies expressed in logs and FDI linearly: the regression coefficients are semi-elasticities, and the standard errors are directly derived from the regression coefficients. However, when models are estimated using the double-log or linear form, we had to re-calculate the effect size by using sample means (see Gujarati and Porter, 2009).[7] We contacted authors when sample means, observation size, t-values or standard errors were not reported in the primary studies and when we needed clarification of the models, methods, and estimates. We have also collected estimates for interaction variables (see Section 4.3.2). Five studies (95 parameter estimates) were excluded because the authors were unwilling or unable to support missing data[8] (note that Figures 1 and 2 have been based on the full sample of 74 studies including the studies where information is incomplete).

3.2. Meta-Dataset

Our dataset consists of 69 primary studies for which the required data are available (1,450 observations). These studies deal with 31 developing countries and were published in 1986-2013. The median number of parameter estimates in a primary study in our sample is 11 estimates. The mean and maximum are 21 and 100 estimates, respectively.

For each empirical study we coded more than 40 potential research dimensions, and four categories of publication qualities. The dataset includes 43 peer-reviewed journal articles and 26 working papers, dissertations, book chapters, unpublished studies or reports. The oldest study was published in 1986, and the median study appeared in 2008. Half of the research in question was published in the last five years: this topic is very lively and many new investigations appeared.

To put these figures into perspective: Nelson and Kennedy (2009) summarize and assess 140 meta-analyses in economics conducted since 1989. They report that the average number of parameter estimates included was 191, the median was 92, and the largest number of parameter estimates was 1,592. The average and the median of primary studies reviewed were 42 and 33, respectively. The mean and median of explanatory variables included were 14 and 12, respectively. The maximum control

variables was 41. Therefore, compared to other meta-analyses in economics, the current dataset can be regarded as quite comprehensive.

Out of the 1,450 spillover estimates, 16 are found to be larger than 10, in absolute value. Some meta-analysts (Iršová and Havránek, 2013; Mebratie and Bergeijk, 2013) consider that these large estimates are outliers which led them to exclude these estimates from the main analysis. Others, like Stanley and Doucouliagos (2012), however, argue that unusually large estimates may be due to coding errors. In our case (after double-checking by a second independent reviewer) the reported spillovers appear to be genuine. Therefore, we applied the multivariate outlier method proposed by Hadi (1994) in order to use spillover estimates along with their precision to filter publication bias and identify outliers jointly. We identify 14.7% of the observations as outliers. Slightly more than 70% of the outliers have been published in either journals with a zero 5-year impact factor (source: 2013 Institute of Scientific Information, ISI) or are from unpublished studies. Some meta-analysts have argued that better ranked journals can be expected to publish articles that use better methods, and thus produce more reliable findings (Disdier and Head, 2008; Havránek and Iršová, 2011). If this is the case, then these outliers may represent lower quality research as compared to the non-outliers. In this paper we report the findings without outliers, but note that the inclusion of outliers yields similar findings.

3.3. Weighted Average Effect, Publication Bias and Genuine Effect

We start with the computation of the weighted average spillover effect:

$$\varepsilon_w = \frac{\sum \varepsilon_i N_i}{\sum N_i} \quad \text{..... (1)}$$

Where ε_i is the reported spillover effect in the i^{th} study with N_i its associated sample size as weights. Copper and Hedges (1994) suggest to use the inverse of the variance weights. However, Adams et al. (1997) argue that the estimates derived from large samples are more precise, and thus should receive a larger weight. Hunter and Schmidt (2004) and Schulze (2004) recommend the use of sample size to weight the effect size.[9] Accordingly, the weighted average spillover effect is 0.16 (statistically significant at a 95% confidence interval: 0.105 - 0.238). Thus, FDI would seem to have a significant positive effect on productivity spillovers.

The next step is to assess whether the effect is genuine or influenced by publication bias. We use funnel plots to get a first indication of the extent of publication bias. A funnel plot is a scatter diagram with the reported spillovers on the horizontal axis and its precision on the vertical axis (usually the reciprocal of the standard error). In the absence of publication bias, a funnel plot should be symmetrical, because small sample size (imprecise estimates/large standard errors) are widely dispersed at the bottom of the funnel (Stanley, 2005). In contrast, large sample studies with usually more precise estimates should be more compactly distributed at the top of the funnel (Stanley and Doucouliagos, 2010). In the case of bias, a funnel plot will be asymmetrical. Asymmetrical plots may also indicate that some parameter estimates are discarded or unreported (Iršová and Havránek, 2013).

However, this method of publication bias detection is only based on visual inspection, and thus prone to subjective interpretation, and therefore unconvincing. A powerful statistical method is the meta-regression model (MRM):

$$\varepsilon_{ji} = \beta_0 + \beta_1 Se_{ji} + u_{ji} \text{…..……..} \quad (2)$$

In the absence of publication bias, the estimated effects (ε_{ji}) should vary randomly around β_0 (its standard error, Se_i will approach zero) and should be independent of their standard errors (Stanley, 2005; Roberts and Stanley, 2005; Doucouliagos and Stanley, 2013). Likewise, the presence of publication bias can be detected if reported estimates correlate with their standard errors. u_{ji} is the error term. According to Stanley and Doucouliagos (2012), dividing equation (2) by Se_{ji} to adjust for potential heteroscedasticity, yields the weighted least squares (WLS) version:

$$t_{ji} \equiv \varepsilon_{ji}/Se_{ji} = \beta_1 + \beta_0(1/Se_{ji}) + e_{ji} \text{……...……} \quad (3)$$

where t_{ji} now represents the t-value of *i*th spillover estimate from the *j*th study and e_{ji} is u_{ji}/Se_{ji}.The slope of equation (3) estimates both the magnitude and direction of a genuine effect (the precision-effect test, PET), whereas the intercept term is used to test for publication bias in the reported effect (the funnel-asymmetry test, FAT).

It is important to account for within-study dependence, as multiple estimates from the same studies are unlikely to be statistically independent; Bateman and Jones (2003) and Doucouliagos and Laroche (2009) recommend the use of a multilevel model (hierarchical model). We prefer the mixed-effects multilevel (MEM) model that accounts for both within and between study variations (see Doucouliagos and Laroche, 2009, Havránek and Iršová, 2011). In this respect, the overall error term (e_{ji}) comprises of the study-level random effects and measurement error disturbance term.[10] Furthermore, we test the robustness of our results by excluding unpublished papers and interaction variables, and by applying clustered data analysis (CDA). All WLS estimations use either inverse variance or sample size weights.

<Table 1 >

3.4. Explaining Heterogeneity

Table 1 lists the potential sources of heterogeneity of the findings in the primary studies that include means and standard deviations. The choice for these variables is based on the discussion of the literature as well as the meta-data at hand. Following the debates in the empirical studies and the approach presented by previous meta-analyses (Havránek and Iršová, 2011; Iršová and Havránek, 2013) we report four categories of potential sources of heterogeneity: data, estimation, specification, and publication characteristics.

Data characteristics: We construct dummy variables for the time dimension of the data: panel versus cross-section and the time span (length) of the data, and for the level of observation/aggregation: firm-level (micro-data) versus aggregated industry-

level. We consider the number of observations of the data to test for systematic variation between small and large samples. Finally, we included a dummy variable for the data source (World Bank enterprise survey versus national statistics bureaus).

Estimation characteristics: Roughly, 40% of the studies estimate spillovers using a two-step process with total factor productivity (TFP) as the dependent variable.[11] The other studies employ a one-step procedure based on labour productivity, output or value added. We include a dummy variable for the one-step estimates. Next, we control how TFP is computed: Olley-Pakes (OP, 1996) or Levinsohn-Petrin (LP, 2003) versus OLS, Fixed-effects or other methods to consider the endogeneity of inputs. We also include dummies for the spillover estimation techniques (OLS, Fixed-effects (year, sector, both) or other techniques and for regressions estimated in differences), and the functional form of the models: linear or double log versus log-linear.

Specification characteristics: Empirical studies use several proxy measures for foreign presence in terms of employment, capital, or output share. In our data, the majority of the studies use output (42%) and employment specifications (35%). Most primary studies report estimates related to contemporaneous spillovers while only a few estimate lagged effects, and thus we control for lagged or contemporaneous estimates. Finally, to observe any systematic difference between the theory of the pipeline model and the domestic capability model, we construct dummies for the inclusion of control variables like absorptive capacity, technological gap, exports, and firm size.

Publication characteristics: We also control for study and journal qualities to test if publishing in a peer-reviewed journal is systematically associated with the reported spillover estimates. We construct dummies for the inclusion of publication in a peer-reviewed journal and use author citations in Google Scholar as well as an international journal ranking for development studies.[12] Finally, we control for the publication year of the study.

To investigate the heterogeneity in the reported spillover estimates, we expand equation (3) to include the moderator variables X_{kji}:

$$t_{ji} \equiv \varepsilon_{ji}/Se_{ji} = \beta_1 + \beta_0\left(\frac{1}{Se_{ji}}\right) + \alpha_k X_{kji}/Se_{ji} + e_{ji} \ldots (4)$$

Following the MAER-Net reporting guidelines, equation (4) is estimated using the general-to-specific (GETS) modelling approach. GETS modelling starts with a specification in which all potential moderator variables are included in the general specification (4). Next, the statistically most insignificant variables are removed, one at a time, until we arrive at a reduced/specific specification that contains significant variables only (Charemza and Deadman, 1997; Stanley and Doucouliagos, 2012, Wang and Shailer, 2015, Abdullah et al. 2015).

4. Findings and Discussion

4.1. Funnel Plots

Figure 3 shows the funnel plots of all studies - published and unpublished studies. The plots suggest a positive bias. The top of the funnel plots is usually a good

approximate of the true empirical effect after due allowance for publication bias (Stanley and Doucouliagos, 2010). Consequently, according to Roberts and Stanley (2005, p.27)

> "… for areas of research that contain many studies, the simplest remedy for publication bias is to average the findings from only the largest studies (say, the top 10%)."

Averaging the top 10% (123) estimates provides an average of -0.010 (standard deviation 0.191). In the absence of unbiased reporting, spillover estimates would be expected to vary randomly around this average. However, the average of all 1233 (i.e., excluding outliers) spillover estimates is 0.172: due to publication bias this average of all estimates appears to exceed the average of the most precise estimates by far and also the overall magnitude of the uncorrected weighted average (Section 3.3) is biased towards positive estimates. This kind of publication bias has clear policy implications. For instance, policymakers may expect a 1.7% increase in domestic firms' productivity from a 10-percentage-point increase in FDI. However, the top 10% estimates suggest a 0.01% decrease in productivity of domestic firms.

<Figures 3 and 4>

In order to check for a possible additional publication bias from editors and reviewers of journals and for the sake of comparison, Figure 4 depicts peer-reviewed studies only.[13] In case of an additional bias, the entire distribution of the funnel plot of Figure 4 would move more to the right as compared to the funnel plot of Figure 3. The plots are heavier (Figure 3) and thinner (Figure 4), but the shape and location are comparable. The next step is to move beyond eye-o-metrics and test econometrically for the existence and size of bias.

4.2. FAT and PET

To assign greater weight to more precise estimates, we apply inverse variance weights and then use sample size as alternate weights. Table 2 reports the FAT and PET. The FAT indicates the presence of bias. Specifically, when all observations are included in the specification, the preferred MEM (A1) shows positive and significant publication bias. Also when only observations from published studies in peer-reviewed journals (A2) are considered, we detect positive and significant publication bias. To further check robustness we use CDA (A3 and A4) that again provides evidence of publication bias. Corroborating the funnel diagrams, the objective MRA provides clear evidence of publication bias. Columns B1-B4 report results with sample size weights corroborating the corresponding columns A1-A4 in Table 2.[14] The inclusion of unpublished studies may result in inefficient parameter estimates as it increases the variance of the meta-dataset, especially if one assumes that peer-reviewed studies publish more reliable findings. Estimations A2 and A4 confirm and agree with the findings on the larger sample that includes the unpublished studies.[15]

The FAT in A1 versus A2 (similarly A3 versus A4) can also be used to test if journal reviewers and editors are a source of publication bias in selecting and accepting findings. In both Panel 1 and Panel 2 the magnitude of publication bias (FAT) is higher for studies published in a peer-reviewed journal compared to all studies, but this difference is not statistically significant: the primary empirical studies are unlikely to have been affected by an extra bias from journal editors and reviewers.

<Table 2>

The size of publication bias in A1 and A2 exceeds unity (preferred MEM): publication bias is substantial in the primary studies (for practical guidelines, see Stanley and Doucouliagos, 2012; Doucouliagos and Stanley, 2013). Publication bias is important in assessing the genuine spillover magnitude versus the uncorrected weighted average effect. In all the specifications A1-A4, the PET consistently suggests the absence of a genuine effect. So the overall uncorrected weighted spillover effect amounts to a publication (or other) bias. In this case, the size of spillover effect reported in the primary studies is likely to be substantially larger than the actual effect.

Note, however, that these findings are an average across all methods. Therefore, we need a multivariate MRA as our inferences may also depend on other potential sources of heterogeneity such as quality of the primary studies, misspecification, research design or other characteristics. Indeed, the heterogeneity across all the studies is evident by the *Q*-test reported under Table 2. The *Q*-test χ^2 distribution with n-1 degrees of freedom and 1233 observations is 61705 (p-value=0.000). The I^2 test of heterogeneity reports that the variation in the reported spillover effect size due to sampling error is 2%. This motivates us to consider other potential sources of heterogeneity (see Section 3.4).

4.3. Sources of Heterogeneity

Table 3 presents the results of the reduced multivariate MRA using GETS modelling. During this procedure, we observed that more than half of the moderator variables included in the general MRA are not statistically significant. We also note that all moderator variables are not equally important in explaining/contributing to the potential source of heterogeneity. In doing so, we eliminate 15 moderator variables that do not appear to be important for the explanation of the heterogeneity.[16] Considering the general versus the specific multivariate model, a trade-off exists between on the one hand, potential multi-collinearity and loss of degrees of freedom and on the other hand, the inclusion of all moderator variables (Mekasha and Trap, 2013). Thus, following the GETS procedure, 11 variables remain statistical significant at least at 10% significance (A5). Indeed, the joint test of these 11 variables rejects the null hypothesis of a zero joint effect $F_{(11, 1220)} = 21.57$. To account for within study correlation, this specific model is then re-estimated using the preferred MEM model[17] and, for comparison and robustness check, with robust standard errors and CDA. The columns report:

- the specific model without adjusting standard errors (A5),

- MEM model through the restricted maximum likelihood (A6 and A9),
- robust standard errors (A7) and
- clustered standard errors (A8 and A10).

All columns use inverse variance weights.

<Table 3>

Controlling for within-study dependence, columns A6 and A8 report 10 moderator variables that potentially explain the heterogeneity in the reported spillover estimates. We find that the spillover effects in the primary studies are systematically influenced. The reported spillover effect will ceteris paribus be larger when: spillovers estimated using a two-step process, technological gap and/or lagged spillovers are not included, and the primary study is peer-reviewed. The use of longer data time span, more recent primary studies, publication in higher journal rank, estimations performed in differences and the fixed-effect estimators are associated with smaller reported spillovers.

We compare and test the robustness of the results using estimates with robust standard errors (A7) and clustered standard errors (A8) reported in Table 3. Without due allowance for data dependence, A7 reports similar result as A6 (but the statistical significance reduces for eight variables). Controlling for data dependence, A8 corroborates the findings of A6 (but statistical significance reduces for one-step estimation, fixed-effect estimators, publication in a peer-reviewed journal and scientific journal rank).

4.3.1. Discussion and Implications for Research Design

On average the time span covered by the study affects the reported spillover estimates and significantly reduces their size by 0.003 per annum. This finding suggests the importance of panel data with wider time coverage (as compared to single-period cross-sections). This implies that the use of a longer data time span is an important moderator variable in research on the effect of FDI spillovers. From this perspective, the higher positive spillover estimates based on cross-sections reflect upward bias due to not controlling for unobserved time-invariant firm specific effects. Also, the larger effect from cross-section studies decreases/disappears with time.

Regarding estimation techniques, a one-step estimation of spillovers using output, labour productivity, or value added on average reports a 0.018 lower spillover effect than the two-step estimation. In light of this, given the bias towards more/higher positive spillover in the reported estimates, we prefer the use of the one-step approach. For random effects, GMM or other spillover estimators, effects on average increase by 0.028, reporting more positive spillover estimates compared to the fixed-effects regression.

Regarding the specification characteristics, the inclusion of technological levels of domestic firms and estimates of lagged spillovers appear to affect the estimates. A specification that controls for the existing technological difference between domestic and foreign firms finds lower spillover effects (on average 0.053). This supports Findlay (1978), Wang and Blomström (1992), Castellani and Zanfei (2003) and

Jordaan (2008a) when they argue that spillovers do not occur automatically but depend on technological heterogeneity.

The conceptual debate (see Lapan and Bardhan, 1973; Findlay, 1978; Cohen and Levinthal, 1989; Wang and Blomström, 1992) and the contrasting empirical evidence (see Kokko, 1994; Kathuria, 2001; Castellani and Zanfei, 2003; Kathuria, 2010) over how the domestic firms' technological level influences the outcome of a given spillover estimate indicates a significant opportunity for future research. It would, for example, be important to investigate how the size of the technological gap between domestic and foreign firms influences potential spillovers. The majority of the primary studies either associates high (low) absorptive capacity with low (high) technological difference or excludes these important moderator variables from the specification. We find that the technological gap is statistically significant but that the absorptive capacity is insignificant, pointing out the importance of disentangling the absorptive capacity hypothesis from the technological gap hypothesis. Indeed, both equating the absorptive capacity as the inverse of the technological gap and excluding them from the analysis, is potentially flawed.

Lastly, on publication status, peer-reviewed studies report spillover estimates that are larger (on average higher by 0.035) than unpublished studies. We also find that publication year of the study affects reported estimates, in that recent studies tend to report lower spillover estimates (on average 0.003 lower). Furthermore, a high-ranked journal is likely to report lower spillover estimates.

4.3.2. Further Robustness Checks: Bivariate and Multivariate

A further robustness check relates to interaction variables. When econometric specifications use interaction terms, sample mean and covariance of the interacted variable are required in order to calculate an effect size.[18] However, in the primary empirical studies the mean and covariance of the interaction term is not always available. Rather than omitting these estimates of the primary studies, Havránek and Iršová (2011) evaluate the interaction effect of foreign presence at sample means of the interacted variables. We follow a different approach and add a binary variable in the meta-dataset (1 for estimates coming from interaction term and 0 otherwise).

<Table 4>

Estimates for FAT-PET *excluding* the interaction terms are reported in Table 4. Both the MEM and CDA corroborate our main findings of the corresponding FAT-PET in Table 2 despite the reduction in the number of observations and primary studies as well as statistical significance level of FAT from the CDA.[19] Furthermore, we deal with interaction estimates in the MRA (Table 3) by including a dummy variable for the interacted terms. A9 and A10, respectively, mimic the findings of A6 and A8, except that the number of firms from A9 is now statistically significant. We also found that the effect of including interaction terms in the primary studies from the preferred model on average is associated with higher spillovers effect.

4.4. "Best Practice" Effect

In comparison to the PET (Section 4.2), the inclusion of moderator variables strongly impacts on the size and significance of the genuine spillover effect. This underlines the need for meta-analysts to consider potential moderator variables. However, there are many potential genuine heterogeneity effects than can be related to a single PET. We therefore follow up in order to systematically estimate the underlying genuine effect from the multivariate MRA conditional on the identified sources of heterogeneity that alleviate omitted variable bias, endogeneity problems and controls for publication bias. This analysis is labelled the "best practice" method (Havránek and Iršová, 2011; Stanley and Doucouliagos, 2012; Wang and Shailer, 2015).

First, we set all possible sources of heterogeneity equal to zero in A6 and A8. This yields a statistically significant positive spillover effect of 0.125 (t=5.23) and 0.127 (t=5.11), respectively. Next, we apply the characteristics of the study of Aitken and Harrison (1999) in order to arrive at the "best practice" estimate. We selected this study for three reasons. First, it is published in the *AER*, one of the leading economics journals with a very stringent referee procedure. Second, it is the most cited study in our dataset (3,051 citations as of August 2013). Third, they use firm-level panel data, perform a one-step regression in differences, and their specification controls for productivity differences across industries. For these reasons this study seems to be free from model misspecification. Using the study characteristics of Aitken and Harrison (1999) the predicted effect size is 0.086 and statistically significant at the 99% confidence level. The procedure yields similar result when A8 is used: 0.085 with t=3.41. After correcting for publication bias and misspecification, the magnitude of the underlying genuine effect is about 0.09.

Conversely, as Havránek and Iršová (2011, p.240) put it, the "worst practice" is "… (a) mirror image of the best practice estimate". Studies that use industry level aggregated cross-sectional data, endogenous TFP estimation, OLS, and specifications that do not control for productivity difference across industries among others would fall in this category. This results in a significantly higher positive effect size of 0.155 (similarly 0.159 for A8), suggesting again that estimates reported in those studies are biased due to misspecification.

4.5. Further Investigation for Publication Bias

The inclusion of potential sources of heterogeneity does not remove publication bias. An important issue is whether publication bias could be the result of "best practice" choices regarding research design and methods. So we ask ourselves the question: could it be that studies that comply with "best practice" characteristics have a (stronger) publication bias. We use panel data, firm-level analysis, and controlling for sector fixed-effects (recall our earlier "best practice" research design estimates).[20] Funnel plots using observations from such "best practice" are reported in Figure 5 and Figure 6 for all studies and peer-reviewed studies only, respectively. Figures 5 and 6 mimic the plots of the corresponding Figures 3 and 4. Leaving aside that the plots are now thinner due to substantial reduction in the number of observations, the shape and location are comparable, and therefore "best practice" choices do not appear to be associated with publication bias.

Following Havránek and Iršová (2011), we also provide a more formal test by interacting the variables that define "best practice" with the estimated spillover effect's standard error.[21] Adding these interaction variables to our MRA specification, regression (4), we find that none of the aspects of the "best practice" are statistically significant.[22] We conclude that the reported spillover effects from FDI are biased towards positive and statistically significant results. However, the "best practice" approach does not appear to be causing publication bias.

<Figures 5 and 6>

5. Conclusion

Our main aim is to analyze the intra-sector FDI spillover effect using 1,450 reported estimates from 69 primary studies by 93 researchers covering 31 developing countries published until 2013.

An important implication of this article is that the (systematic) selection of primary studies may lead to bias in a meta-analysis. Meta-analysts often involve judgment in identifying the primary studies. Thanks to our meta-analysis, we know that the selection of studies in previous meta-analyses has had an impact on the findings and led to misleading conclusions (for example the issue of publication bias and the genuine empirical effect). In this study, we try to avoid this bias by gathering all studies for developing countries. We strictly follow the MRA reporting guidelines of MAER-Net recommended by this Journal.

We found a positive and significant uncorrected weighted average spillover effect of 0.16. We investigated whether this effect is genuine or affected by publication (or other) bias using funnel plots, FAT-PET, and GETS meta-analytical approaches. In accordance with the seminal findings by Görg and Strobl (2001) and contrary to the recent studies by Iršová and Havránek (2013) and Mebratie and Bergeijk (2013), our study clearly uncovers that publication bias is a problem for this field and we establish the extent and source of the bias. Reported spillover estimates significantly overstated the true effect, but this does not appear to have been caused by either publication selection pressure from editors and reviewers of journals or authors' prior interest to follow "best practice". Unlike Iršová and Havránek (2013) who find that intra-sector spillovers are statistically indistinguishable from zero, we find that the underlying genuine spillover effect is economically important with a magnitude of 0.09, notably also after taking publication bias and misspecification of the primary studies into account. This became clear in the multivariate approach, showing the need to carefully consider moderator variables which can potentially help to explain the nature of heterogeneity in reported findings.

Primary empirical studies appear to be subjective in the choice of methodology. Application of MRA that investigates the diverse empirical findings of primary studies can help by giving useful insights for theory and stimulate "best practice" to improve research design of future primary studies. The "best practice" from the MRA methodology focuses on controlling publication bias, reducing omitted variable bias and endogeneity issues across the studies under review. Our results for the "best practice" support theories that predict a positive association of spillovers from FDI.

Despite the robust positive underlying spillover effect, there is also substantial heterogeneity in reported spillover estimates. Spillovers' effect size and their sign depend systematically on two major sources of heterogeneity: specification choices and publication bias (but are also affected by data and estimation characteristics). Hence, we suggest that future research needs to carefully consider the selection of explanatory variables in order to avoid omitted variable bias. Indeed understanding publication bias potentially calls for researchers to consider the sources of bias towards positive spillover estimates.

With respect to the technological hypothesis versus the absorptive capacity hypothesis, it is important to disentangle these two hypotheses which may lead to different spillover estimates. Furthermore, not only from the academic point of view, but also from a policy perspective, it would be very important to separate the different transmission mechanisms under which spillovers actually take place. Our review of the empirical findings of 74 studies shows predominance of failure to identify the mechanisms underlying spillovers effect. Thus, future studies should consider not only whether spillovers occur, but also look inside the black box of how spillovers actually emerge. Importantly, empirical research efforts should also expand to cover more developing countries.

Acknowledgements

Two anonymous referees, Hristos Doucouliagos, Anagaw Derseh Mebratie, Elissaios Papyrakis, Henri de Groot, Mansoob Murshed, Sara Lazzaroni, Selwyn J.V. Moons, Tomas Havránek, and Zahra Zarepour provided useful comments. The paper benefited from the discussions and comments received from the participants of the 2014 IGAD Economic conference, the 2014 MAER-Net colloquium and the 2014 German Economic Association conference. The authors would also like to thank all authors of primary studies that provided additional data and/or clarification.

Notes

1. These are: Argentina, Bangladesh, Bolivia, Cambodia, Chile, Colombia, Ecuador, Ethiopia, Ghana, Guatemala, India, Indonesia, Kenya, Malaysia, Mexico, Morocco, Panama, Paraguay, Peru, Philippine, South Africa, Thailand, Taiwan, Tanzania, Turkey, Uganda, Uruguay, Venezuela, Vietnam, Zambia, and Zimbabwe.
2. Iršová and Havránek (2013) analyze 1,205 observations from 52 studies for the period 2002-2010 dealing with both developing and developed countries. The study concludes that spillover effects are insignificant and report absence of publication bias in this literature. The absence of publication bias (using a dummy variable for peer-reviewed studies) is also reported by Mebratie and Bergeijk (2013). They study 156 observations from 30 studies in developing countries and emerging markets using a random selection under the condition of one study per country for the period until 2010.

3. The 74 primary studies we have collected all define foreign-owned firm as a firm that has at least 10% foreign ownership.
4. Inter-industry studies are presumed to be too dissimilar to pool sensibly with intra-sector studies (Wooster and Diebel, 2010).
5. Studies that examine determinants, descriptive and qualitative studies as well as papers that could not be downloaded are excluded. But first we tried to contact authors for inaccessible papers if contact info was found. For example, Sasidharan and Ramanathan (2007) is not accessible online. The study was included after communicating with the authors.
6. The best-set consists of one estimate that the author believes to be the key regression of the study often labelled "preferred equation", whereas the all-set is collected from all relevant estimates that may offer more observations to explain heterogeneity. The average-set is computed from the all-set estimates.
7. Similar approach is used by Feld and Heckemeyer (2011) and Iršová and Havránek (2013). For instance when a double-log form is used, sample mean for spillovers variable needed to convert into a semi-elasticity (i.e., $b/\bar{X}$) as in: $\log Y_i = a + b\log X_i + u_i$
8. These include Blomström and Persson (1983), Kokko et al. (1996), Kokko et al. (2001), Takii (2005), and Kathuria (2010).
9. Mekasha and Trap (2013) re-investigated the aid-growth link based on the 68 primary studies initially meta-analyzed by Doucouliagos and Paldam (2008). In computing the magnitude of the uncorrected weighted average effect, the latter used sample size weights, whereas the former used the inverse of the variance and both result in similar results. This illustrates that the use of either weight may not matters. However, Stanley and Doucouliagos (2012) report that uncorrected weighted averages using the fixed-effects and random-effects are biased in the face of publication bias.
10. That is: $t_{ji} = \beta_1 + \beta_0(\frac{1}{Se_{ji}}) + \mu_{ji} + \varepsilon_{ji}$ where μ_{ji} is study-level random-effects and ε_{ji} is measurement error disturbance term.
11. In the first-step, authors estimate TFP, followed by a second-step estimation of spillovers effect using TFP as the dependent variable.
12. The national Dutch research school for development studies CERES, provides journal quality classification through the impact factor of the Institute of Scientific Information (ISI) http://ceres.fss.uu.nl/rating-lists/rating-list/. A-journals (high quality) ranked from the top one-third cited outlets of the 2013 ISI impact factor. Thus, we create a dummy for high quality journals and use other classifications as reference.
13. Spillovers estimate of about less than -2 and precision of about greater than 4 appear to be outliers. Note that we already dealt with outliers in Section 3.2.
14. Table 2 (B1-B4) and Table 4 (B5-B8) also report the FAT-PET estimates with sample size weights. The FAT from the preferred model displays evidence of positive funnel asymmetry. The PET (meta-average corrected for publication bias) is not statistically significant. After due allowance for publication bias the meta-average of no effect cannot be rejected, with similar result regardless of either weights used.

15. Note that this leads to a significant reduction in the number of observations of the meta-dataset as well as a loss of additional information available through the full dataset. Although there is no clear cut-off value to separate high precision, we have further checked the robustness of the results by excluding high precision values from the analysis, and this yields similar results. However, high precision values must be retained as they are genuinely informative about the research literature considered (Stanley and Doucouliagos, 2012).
16. Insignificant moderator variables excluded from the reduced model are (ordered from least significance): all firms, exports, panel data, study citations, TFP estimated with OP or LP, data source, foreign presence in equity, firm size, absorptive capacity, linear/log-log, foreign presence in employment, balanced data, year fixed-effects, OLS and firm level data. In support of the removal of these variables, the null hypothesis of a zero joint effect cannot be rejected, $(F_{15,\ 1205}) = 0.49$ (p-value 0.9483): jointly the 15 variables appear to be statistically indistinguishable from zero, and thus not contributing to the explanation of the heterogeneity.
17. We estimated the mixed-effects multilevel model, and thus account for both within-study and across-study variations. Robust methods in meta-analysis using the MEM model are widely employed, (see Doucouliagos and Laroche, 2009; Doucouliagos and Stanley, 2009; Havránek, 2010; Havránek and Iršová, 2011). In the case of the FDI-spillovers, the result from the MEM is very similar to the CDA findings. In our view it remains an issue for future research whether the CDA version or the MEM better reflects meta-analysis of the literature in question.
18. When the regression specification includes interaction of the spillover variables and other control variables (typically, technological gap and/or absorptive capacity) as in the following equation:

$$\log Y_i = a + bX_i + cX_i * Z_i + dZ_i + \cdots + \varepsilon_i$$

the effect size *(e)* and the standard error *(Se)*:

$$e = 100(b + c\bar{X}) \qquad S_e = 100\sqrt{var(b) + 4\,\overline{X_i * Z_i}\ var\ (c) + 4\bar{X}_i cov(b, c)}$$

19. We thank an anonymous referee for this valuable observation.
20. We cannot apply all elements of the "best practice" because too few estimates would remain for analysis (Havránek and Iršová, 2011, chose a similar approach).
21. We have interacted with and tested the time span of the data, publication status, specifications for sector fixed-effect, one-step estimations in differences and specifications that control technological level.
22. Indeed, the joint test of these interaction variables reveals that the null hypothesis of a zero effect cannot be rejected (p-value of 0.5643).

References

(Primary studies in the meta-analysis dataset are marked by *)

Abdullah, A., H. Doucouliagos and E. Manning (2015) Does education reduce income inequality? A meta-regression analysis, *Journal of Economic Surveys*, ***29***(2), 301-316.

Adams, D.C., J. Gurevitch and M.S. Rosenberg (1997) Resampling tests for meta-analysis of ecological data, *Ecology*, ***78***(4), 1277-1283.

Aitken, B.J. and A.E. Harrison (1999) Do domestic firms benefit from direct foreign investment? Evidence from Venezuela, *American Economic Review*, ***89***(3), 605-618.*

Albornoz, F. and M. Kugler (2008) Exporting spillovers: firm-level evidence from Argentina, *Laurier Business and Economics Department of Economics WP Series*, 2008-02.*

Aldaba, R.M. and F.T. Aldaba (2012) Does FDI have positive spillover effects? The case of the Philippine manufacturing industry, *ARTNeT Conference on Empirical and Policy Issues of Integration in Asia and the Pacific*.*

Aslanoğlu, E. (2000) Spillover effects of foreign direct investments on Turkish manufacturing industry, *Journal of International Development*, ***12***(8): 1111-1130.*

Bateman, I.J. and A.P. Jones (2003) Contrasting conventional with multi-level modelling approaches to meta-analysis: expectation consistency in UK woodland recreation values, *Land Economics*, ***79***(2), 235-258.

Björk, I (2005) Spillover effects of FDI in the manufacturing sector in Chile, *School of economics and management, Lund University, Master thesis*.*

Blalock, G. and P.J. Gertler (2008) Welfare gains from foreign direct investment through technology transfer to local suppliers, *Journal of International Economics,****74***(2), 402-421.*

Blalock, G. and P.J. Gertler (2009) How firm capabilities affect who benefits from foreign technology, *Journal of Development Economics*, ***90***(2), 192-199.*

Blalock, G. and D.H. Simon (2009) Do all firms benefit equally from downstream FDI and quest; The moderating effect of local suppliers' capabilities on productivity gains, *Journal of International Business Studies*, ***40***(7), 1095-1112.*

Blomström, M. and H. Persson (1983) Foreign Investment and Spillover Efficiency in an Underdeveloped Economy: Evidence from the Mexican Manufacturing Industry, *World Development*, ***11***(6): 493-501.

Blomström, M. (1986) Foreign investment and productive efficiency: the case of Mexico, *The Journal of Industrial Economics*, ***35***(1), 97-110.*

Blomström, M. and F. Sjöholm (1999) Technology transfer and spillovers: does local participation with multinationals Matter? *European Economic Review*, ***43***(4), 915-923.*

Blomström, M. and E.N. Wolff (1994) Multinational corporations and productivity convergence in Mexico, *National Bureau of Economic Research, WP No. 3141*.*

Blyde, J., M. Kugler and E. Stein (2004) Exporting vs. Outsourcing by MNC subsidiaries: which determines FDI spillovers? *Discussion papers in economics and econometrics-university of Southampton*.*

Bouoiyour, J. and A. Akhawayn (2003) Labour productivity, technological gap and spillovers: evidence from Moroccan manufacturing industries, *WPCATT, University of Pau*.*

Bwalya, S.M. (2006) Foreign direct investment and technology spillovers: evidence from panel data analysis of manufacturing firms in Zambia, *Journal of Development Economics*, ***81***(2), 514-526.*

Castellani, D. and A. Zanfei (2003) Technology gaps, absorptive capacity and the impact of inward investments on productivity of European firms, *Economics of Innovation and New Technology*, ***12***(6):555–76.

Castro, L. (2012) Does licensing induce technological spillovers to domestic firms? *University of Colorado at Boulder, Economics Department Paper*.*

Cheng, S. (2011) Foreign direct investment and productivity spillovers: empirical evidence from Cambodia. *Economics Bulletin*, ***32***(3), 2015-2025.*

Charemza, W. and D. Deadman (1997) *New directions in econometric practice*, Edward Elgar: Cheltenham.

Chuang, Y.C. and C.M. Lin (1999) Foreign direct investment, R&D and spillover efficiency: evidence from Taiwan's manufacturing firms, *The Journal of Development Studies*, ***35***(4), 117-137.*

Chudnovsky, D., A. Lopez and G. Rossi (2008) Foreign direct investment spillovers and the absorptive capabilities of domestic firms in the Argentine manufacturing sector, *Journal of Development Studies*, ***44***(5), 645-677.*

Cohen, W. and D. Levinthal (1989) Innovation and learning: the two faces of R&D, *Economic Journal*, **99**, 569–596.

Copper, H.M. and L.V. Hedges (Eds.) (1994) *Handbook of research synthesis*, Russell Sage: New York.

Crespo, N. and M.P. Fontoura (2007) Determinant factors of FDI spillovers–what do we really know? *World development*, ***35***(3), 410-425.

Cuyvers, L., R. Soeng, J. Plasmans and D. van den Bulcke (2008) Productivity spillovers from foreign direct investment in the Cambodian manufacturing sector: evidence from establishment-level data, *University of Antwerp, Faculty of Applied Economics*.*

Demena, B.A. (2015) Publication bias in FDI spillovers in developing countries: a meta-regression analysis, *Applied Economics Letters*, ***22***(14), 1170-1174

Disdier, A.C. and K. Head (2008) The puzzling persistence of the distance effect on bilateral trade, *The Review of Economics and Statistics*, ***90***(1), 37-48.

Doucouliagos, H. and M. Paldam (2008) Aid effectiveness on growth: a meta study, *European journal of political economy*, ***24***(1), 1-24.

Doucouliagos, H. and P. Laroche (2009) Unions and Profits: a meta-regression analysis, *Industrial relations: a journal of economy and society*, ***48***(1), 146-184.

Doucouliagos, H. and T.D. Stanley (2009) Publication selection bias in minimum-wage research? A meta-regression analysis, *British Journal of Industrial Relations*, ***47***(2), 406-428.

Doucouliagos, C. and T.D. Stanley (2013) Are all economic facts greatly exaggerated? Theory competition and selectivity, Journal *of economic surveys*, ***27***(2), 316-339.

Erdogan, A.I. (2011) Foreign direct investment and productivity spillovers: evidence from Turkey, *Journal of Applied Finance and Banking*, ***1***(4), 185-199.*

Fan, E.X. (2002) Technological spillovers from foreign direct investment—A survey, *Asian Development Review*, ***20***(1), 34–56.

Feinberg, S.E. and S.K. Majumdar (2001) Technology spillovers from foreign direct investment in the Indian pharmaceutical industry, *Journal of International Business Studies*, ***32***(3): 421-437.*

Feld, L.P. and J.H. Heckemeyer (2011) FDI and Taxation: a meta-study, *Journal of economic surveys*, ***25***(2), 233-272.

Findlay R. (1978) Relative backwardness, DFI, and the transfer of technology: a simple dynamic model, *Quarterly Journal of Economics*, **92**, 1-16.

Gachino, G.G. (2006) Foreign direct investment and firm level productivity in Kenyan manufacturing: a panel data analysis, *United Nations University, Maastricht Economic and social Research and training centre on Innovation and Technology*.*

Görg, H. and E.A. Strobl (2001) Multinational companies and productivity spillovers: a meta-analysis, *The Economic Journal*, ***111***(475), 723-739.

Gujarati, D.N. and D.C. Porter (2009) *Basic Econometrics*, Fifth Edition, New York: McGraw Hill.

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

Hadi, A.S. (1994) A modification of a method for the detection of outliers in multivariate samples, *Journal of the Royal Statistical Society*, ***56***, 393-396.

Hamida, L.B. (2013) Are there regional spillovers from FDI in the Swiss manufacturing industry? *International Business Review*, ***22***(4), 754-769.

Havránek, T. (2010) Rose effect and the euro: is the magic gone? *Review of World Economics*, ***146***(2), 241-261.

Havránek, T. and Z. Iršová (2011) Estimating vertical spillovers from FDI: why results vary and what the true effect is, *Journal of International Economics*, ***85***(2), 234-244.

Henning, H. (2013) Firm-level productivity spillovers from FDI in Latin American countries, *IEE Institute of Development Research and Development Policy, WP No. 196*.*

Hunter, J.E. and F.L. Schmidt (eds.) (2004) *Methods of Meta-Analysis: correcting error and bias in research findings*, Sage: London.

Iršová, Z. and T. Havránek (2013) Determinants of horizontal spillovers from FDI: evidence from a large meta-analysis, *World Development*, ***42***, 1-15.

Javorcik, B.S. (2004) Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages, *The American Economic Review*, *94*(3), 605-627.

Jordaan, J.A. (2005) Determinants of FDI-induced externalities: new empirical evidence for Mexican manufacturing industries, *World Development*, ***33***(12): 2103-2118.*

Jordaan, J.A. (2008a) Intra-and inter-industry externalities from foreign direct investment in the Mexican manufacturing sector: new evidence from Mexican regions, *World Development*, ***36***(12), 2838-2854.*

Jordaan, J.A. (2008b) Regional foreign participation and externalities: new empirical evidence from Mexican regions, *Environment and planning*, ***40***(12): 2948-2969.*

Jordaan, J.A. (2011) Cross-sectional estimation of FDI spillovers when FDI is endogenous: OLS and IV estimates for Mexican manufacturing industries, *Applied Economics*, *43*(19), 2451-2463.*

Jordaan, J.A. (2012) *Foreign Direct Investment, Agglomeration and Externalities: Empirical Evidence from Mexican Manufacturing Industries*, Ashgate: Cheltenham.

Kathuria, V. (2000) Productivity spillovers from technology transfer to Indian manufacturing firms, *Journal of International Development*, ***12***(3), 343-369.*

Kathuria, V. (2001) Foreign firms, technology transfer and knowledge spillovers to Indian manufacturing firms: a stochastic frontier analysis, *Applied Economics*, ***33***(5), 625-642.*

Kathuria, V. (2002) Liberalisation, FDI, and productivity spillovers: an analysis of Indian manufacturing firms, *Oxford Economic Papers*, ***54***(4), 688-718.*

Kathuria, V. (2010) Does the technology gap influence spillovers? A post-liberalization analysis of Indian manufacturing industries, *Oxford Development Studies*, ***38***(2), 145-170.

Kee, H.L. (2005) Foreign ownership and firm productivity in Bangladesh garment sector, *Research WP*, The World Bank, Washington DC.*

Kee, H.L. (2013) Local intermediate inputs and the shared supplier spillovers of foreign direct investment, Policy *Research WP*, The World Bank, Washington DC.*

Khalifah, N.A. and R. Adam (2009) Productivity spillovers from FDI in Malaysian manufacturing: evidence from micro-panel data, *Asian Economic Journal*, ***23***(2), 143-167.*

Khawar, M. (2003) Productivity and foreign direct investment: evidence from Mexico, *Journal of Economic Studies*, ***30***(1): 66-76.*
Kinuthia, B.K. (2013) Reversed fortunes in the south: a comparison of the role of FDI in industrial development in Kenya and Malaysia, *African Studies Centre, Leiden*.*
Kohpaiboon, A. (2006) Foreign direct investment and technology spillover: A cross-industry analysis of Thai manufacturing, *World Development*, ***34***(3), 541-556.*
Kokko, A. (1994) Technology, market characteristics, and spillovers, *Journal of Development Economics*, ***43***(2), 279-293.*
Kokko, A. (1996) Productivity spillovers from competition between local firms and foreign affiliates, *Journal of international development*, ***8***(4), 517-530.*
Kosteas, V.D. (2008) Foreign direct investment and productivity spillovers: a quantile analysis, *International Economic Journal*, ***22***(1), 25-41.*
Köymen, S. and S. Sayek (2009) The role of human capital in productivity spillovers from FDI: an empirical analysis on Turkish manufacturing firms. *Workshop paper*. ESRC–ITU and University of Nottingham.*
Lapan, H. and P. Bardhan (1973) Localised technical progress and the transfer of technology and economic development, *Journal of Economic Theory*, **6**, 585–595.
Le, H.Q. and R. Pomfret (2011) Technology spillovers from foreign direct investment in Vietnam: horizontal or vertical spillovers? *Journal of the Asia Pacific Economy*, ***16***(2), 183-201.*
López-Córdova, J.E. (2002) NAFTA and Mexico's manufacturing productivity: An empirical investigation using micro-level data, *Inter-American Development Bank, Washington, DC*.*
Managi, S. and S.M. Bwalya (2010) Foreign direct investment and technology spillovers in sub-Saharan Africa, *Applied Economics Letters*, ***17***(6), 605-608.*
Marin, A. and M. Bell (2006) Technology spillovers from foreign direct investment: The active role of MNC subsidiaries in Argentina in the 1990s, *The Journal of Development Studies*, ***42***(4): 678-697.*
Marin, A. and S. Sasidharan (2010) Heterogeneous MNC subsidiaries and technological spillovers: explaining positive and negative effects in India, *Research Policy*, ***39***(9), 1227-1241.*
Mebratie, A.D. and P.A.G. van Bergeijk (2013) Firm heterogeneity and development: a meta-analysis of FDI productivity spillovers, *The Journal of International Trade and Economic Development*, ***22***(1), 53-74.
Mebratie, A.D. and A.S. Bedi (2013) Foreign direct investment, black economic empowerment and labour productivity in South Africa, *The Journal of International Trade and Economic Development*, ***22***(1), 103-128.*
Mekasha T.J. and F. Trap (2013) Aid and Growth: what meta-analysis reveals, *The Journal of Development Studies*, ***49***(4), 564-583.
Melese, M. and A. Waldkirch (2011) Foreign participation, productivity and transition in a least developed country: firm-level evidence, *Colby College, Department of Economics, Waterville*.*
Meyer, K.E. and E. Sinani (2009) When and where does foreign direct investment generate positive spillovers and quest; a meta-analysis, *Journal of International Business Studies*, ***40***(7), 1075-1094.
Na-Allah, A. and M. Muchie (2009) A firm-level analysis of technological externality of foreign direct investment in South Africa, *African Journal of Science, Technology, Innovation and Development*, ***1***(1), 76-102.*
Narula, R. and A. Marin (2003) FDI spillovers, absorptive capacities and human capital development: evidence from Argentina, *Maastricht Economic Research Institute on Innovation and Technology*.*

Nelson, J., and P. Kennedy (2009) The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment, *Environmental and Resource Economics*, ***42***(3), 345–377.

Nguyen, L. (2008) Productivity spillovers from foreign direct investment: evidence from Vietnamese firm data, *SSRN 1101203*.*

Nguyen, C.D., G. Simpson, D. Saal, A.N. Nguyen and N.Q. Pham (2008) FDI horizontal and vertical effects on local firm technical efficiency, *Development and Policies Research Center WP*, Hanoi.*

Nguyen, A.N., N. Thang, L.D. Trung, N.Q. Pham, C.D. Nguyen and N.D. Nguyen, (2008) Foreign direct investment in Vietnam: Is there any evidence of technological spillover effects, *Development and Policies Research Center WP 18* Hanoi.*

Okot, N. (2013) Are there productivity spillovers from foreign direct investment in Uganda? Evidence from firm level panel data, *UNU-WIDER WP*.*

Rattsø, J. and H.E. Stokke (2003) Learning and foreign technology spillovers in Thailand: empirical evidence on productivity dynamics, *Nordic Journal of Political Economy*, ***29***(1), 47-66.*

Roberts, C. and T.D. Stanley (2005) *Meta-regression analysis: issues of publication bias in economics*, Wiley-Blackwell: London.

Rutaihwa, J. (2013) FDI spillovers in Tanzania's manufacturing sector, *UNU-WIDER WP*.*

Salim, R.A. and H. Bloch (2009) Does foreign direct investment lead to productivity spillovers? Firm level evidence from Indonesia, *World Development*, ***37***(12), 1861-1876.*

Sarkar, S. and Y.C. Lai (2009) Foreign direct investment, spillovers and output dispersion: the case of India, *International Journal of Information and Management Sciences*, ***20***(4), 491-503.*

Sasidharan, S. and A. Ramanathan (2007) Foreign direct investment and spillovers: evidence from Indian manufacturing, *International Journal of Trade and Global Markets*, ***1***(1), 5-22.*

Schulze, R. (2004) *Meta-Analysis: a comparison of approaches*, Hogrefe and Huber: Göttingen.

Sjöholm, F. (1999a) Technology gap, competition and spillovers from direct foreign investment: evidence from establishment data, *Journal of Development Studies* ***36***(1), 53-73.*

Sjöholm, F. (1999b) Productivity growth in Indonesia: the role of regional characteristics and direct foreign investment, *Economic Development and Cultural Change*, ***47***(3), 559-584.*

Stanley, T. D. (2001) Wheat from chaff: meta-analysis as quantitative literature review, *The Journal of Economic Perspectives*, ***15***(3), 131-150.

Stanley, T.D. (2005) Beyond publication bias, *Journal of Economic Surveys*, ***19***(3), 309-345.

Stanley, T.D. and H. Doucouliagos (2010) Picture this: a simple graph that reveals much ado about research, *Journal of Economic Surveys*, ***24***(1), 170-191.

Stanley, T.D. and H. Doucouliagos (2012) *Meta-Regression Analysis in Economics and Business*, Routledge, Oxford.

Stanley, T.D., H. Doucouliagos, M. Giles, J.H. Heckemeyer, R.J. Johnston, P. Laroche and K. Rost (2013) Meta-Analysis of Economics Research Reporting Guidelines, *Journal of Economic Surveys*, ***27***(2), 390-394.

Takii, S. (2009) Multinationals, technology upgrading, and wages in urban and rural Indonesia, *Review of Development Economics*, ***13***(1), 151-163.*

Takii, S. (2011) Do FDI spillovers vary among home economies? Evidence from Indonesian manufacturing, *Journal of Asian Economics*, ***22***(2), 152-163.*

Taymaz, E. and K. Yilmaz (2008) Foreign direct investment and productivity spillovers: identifying linkages through product-based measures, *35th EARIE Conference, Toulouse*.*

Thuy, L.T. (2005) Technological spillovers from foreign direct investment: the case of Vietnam. *School of Economics, University of Tokyo*.*

Todo, Y. and K. Miyamoto (2006) Knowledge spillovers from foreign direct investment and the role of local R&D activities: evidence from Indonesia, *Economic Development and Cultural Change*, ***55***(1), 173-200.*

United Nation Conference on Trade and Development (UNCTAD 2013) *Global value chains: investment and trade for development*. *World Investment Report*, New York and Geneva.

Van Thanh, H. and P.T. Hoang (2010) Productivity spillovers from foreign direct investment: the case of Vietnam, In Hahn, C.H. and D. Narjoko (eds.) Causes and consequences of globalization in East Asia: what do the micro data analyses show? *ERIA Research Project Report*: Jakarta.*

Villegas-Sanchez, C. (2008) FDI spillovers and the role of local financial markets: evidence from Mexico, *Manuscript*, University of Houston: Texas.*

Waldkirch, A. and A. Ofosu (2010) Foreign presence, spillovers, and productivity: evidence from Ghana, *World Development*, ***38***(8):1114-1126.*

Wang, J.Y. and M. Blomström (1992) Foreign investment and technology transfer: a simple model, *European economic review*, ***36***(1), 137-155.

Wang, K. and G. Shailer (2015) Ownership concentration and firm performance in emerging markets: a meta-analysis, *Journal of Economic Surveys*, ***29***(2)1467-6419.

Wooster, R.B. and D.S. Diebel (2010) Productivity spillovers from foreign direct investment in developing countries: a meta-regression analysis. *Review of Development Economics*, ***14***(3), 640-655.

World Bank (2012) Enterprise Surveys (http://www.enterprisesurveys.org).

Tables

Table 1. Definition and descriptive statistics of explanatory variables

Moderator Variables	Definition	Mean	Standard deviation
1/se	Precision of estimated spillover	30.458	56.276
	Data Characteristics		
Panel data	=1 if panel-data are used (cross-sectional data is the base)	0.726	0.446
Firm level	=1 if firm-level data are used (industry-level data is the base)	0.929	0.256
Data source	=1 if the data come from the World Bank (national statistics bureaus data as a base)	0.815	0.388
Time-span	The number of years of the data used	5.957	3.702
No. of firms [a]	Sample size/time-span	2.130	3.246
Balanced data	=1 if balanced dataset is used	0.145	0.352
	Estimation Characteristics		
Linear/Log-log	=1 if the coefficient is taken from a specification different from log-level	0.120	0.326
Differences	=1 if the regression is estimated in differences	0.131	0.338
Year fixed effects	=1 if year fixed effects are included	0.564	0.496
Sector fixed effects	=1 if sector fixed effects are included	0.580	0.494
OLS	=1 if OLS used for spillover estimations (random-effects, GMM, WLS and others as a base)	0.398	0.490
Fixed-effects	=1 if fixed-effects used for the estimation of spillovers	0.269	0.444
One-step estimations	=1 if spillovers are estimated in one-step	0.554	0.497
OP-LP	=1 if TFP estimated using the OP-LP method to account for input endogeneity	0.297	0.457
	Specification Characteristics		
Foreign presence in employment	=1 if proxy for foreign presence is employment (base output and others)	0.352	0.478
Foreign presence in equity	=1 if proxy for foreign presence is equity (base output and others)	0.178	0.383
Technological gap	=1 if specification controls for technology gap.	0.097	0.295
Absorptive capacity	=1 if the specification controls for absorption capacity using R&D expenditure or percentage of a firm's workers with college or higher degrees	0.202	0.402
Firm size	=1 if the specification controls for firm size (sector competition)	0.333	0.472
Exports	=1 if the specification control for exports	0.205	0.404
All firms	=1 if both domestic and foreign firms are included in the regression	0.337	0.473
Lagged spillover	=1 if the coefficient represents lagged foreign presence	0.105	0.306
Interaction terms	=1 if the coefficient comes from interaction variables	0.161	0.368
	Publication Characteristics		
Publication date	The publication year of the study (base, 1986)	21.883	3.985
Published	=1 if the study was published in a peer-reviewed journal	0.563	0.496
Study citations	Study citations in Google Scholar per age of the study, as of August 2013	8.174	25.026
Journal rank	=1 if the study published in high journal rank, 2013 ISI impact factor	0.325	0.468

[a] Mean and standard deviation are divided by a thousand to make the figures easier to read.

Table 2. Bivariate MRA for FAT-PET

Panel 1:MEM				
	All-studies		**Peer-reviewed studies**	
Variables	*(A1)* t-value	*(B1)* Effect-size	*(A2)* t-value	*(B2)* Effect-size
Bias/FAT	1.054** (2.20)	0.419*** (3.42)	1.325* (1.76)	0.513* (1.84)
Genuine effect/ PET	-0.006 (-1.16)	0.044 (0.60)	-0.011 (-1.11)	-0.018 (-0.19)
Observations	1233	1233	694	694
Studies	65	65	41	41
Panel 2:CDA				
	All-studies		**Peer-reviewed studies**	
Variables	*(A3)* t-value	*(B3)* Effect-size	*(A4)* t-value	*(B4)* Effect-size
Bias/FAT	0.505** (2.07)	0.454*** (3.27)	0.572* (1.77)	0.461 (1.47)
Genuine effect/ PET	0.0004 (0.06)	-0.008 (-0.17)	-0.0005 (-0.02)	-0.039 (-0.44)
Observations	1233	1233	694	694
Studies	65	65	41	41

Notes: ***, **, * stands for 1, 5 and 10% level of significance. Figures in brackets are *t*-values. Test for between-study heterogeneity (Q-test) is 61705.15***on1232 degrees of freedom with p-value less than 0.001 and I^2 statistics (variation in spillover estimates attributable to heterogeneity) is 98.0%.The total number of observations is 1233 instead of 1237 since four spillover estimates have zero standard errors. The number of all studies is 65 (41 peer-reviewed studies) as two peer-reviewed and two unpublished studies are identified as outliers. t-value columns estimated from equation (3) that uses inverse variance as weights and effect size columns from equation (2) using sample size weights. Panel 1 (MEM: mixed-effects multilevel model) estimated through the restricted maximum likelihood, whereas Panel 2 (CDA: clustered data analysis) from study level clustered standard errors.

Table 3. Multivariate MRA for source of heterogeneity: reduced model

Moderator Variables	***(A5)*** **Specific**	***(A6)*** **MEM**	***(A7)*** ***Robust se***	***(A8)*** **CDA**	***(A9)*** **MEM**	***(A10)*** **CDA**
Genuine effect	0.127***	0.125***	0.127*	0.127***	0.138***	0.142***
(Precision, $\boldsymbol{\beta_0}$)	(0.023)	(0.024)	(0.067)	(0.025)	(0.024)	(0.039)
Bias coefficient	0.407*	0.456*	0.407***	0.407*	0.588**	0.505**
(Intercept, $\boldsymbol{\beta_1}$)	(0.215)	(0.248)	(0.109)	(0.226)	(0.255)	(0.206)
			Data			
Time span	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
No. of firms [a]	0.004**	0.003	0.004	0.004	0.004**	0.003
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)
			Estimation			
One-step	-0.019**	-0.018**	-0.019	-0.019*	-0.021**	-0.021*
estimations	(0.008)	(0.008)	(0.016)	(0.010)	(0.008)	(0.012)
Fixed-effects	-0.028***	-0.028***	-0.028	-0.028*	-0.025***	-0.025*
estimators	(0.008)	(0.008)	(0.023)	(0.015)	(0.008)	(0.014)
Difference	-0.027**	-0.026**	-0.027	-0.027**	-0.022**	-0.022*
	(0.011)	(0.011)	(0.019)	(0.012)	(0.011)	(0.011)
Sector fixed effects	-0.024***	-0.023***	-0.024	-0.024***	-0.024***	-0.026***
	(0.008)	(0.008)	(0.021)	(0.006)	(0.008)	(0.010)
			Specification			
Technological gap	-0.054***	-0.053***	-0.054	-0.054***	-0.059***	-0.059***
	(0.014)	(0.014)	(0.045)	(0.016)	(0.014)	(0.017)
Lagged spillover	-0.167***	-0.164***	-0.167***	-0.167***	-0.165***	-0.168***
	(0.018)	(0.018)	(0.064)	(0.020)	(0.018)	(0.020)
					0.035***	0.034
Interaction terms	-	-	-	-	(0.009)	(0.023)
			Publication			
Published	0.035**	0.035**	0.035*	0.035	0.035**	0.035
	(0.016)	(0.017)	(0.020)	(0.022)	(0.017)	(0.023)
Publication date	-0.003***	-0.003***	-0.003*	-0.003***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Journal rank	-0.035**	-0.035**	-0.035*	-0.035*	-0.053***	-0.051***
	(0.016)	(0.016)	(0.021)	(0.019)	(0.017)	(0.020)
Observations	1233	1233	1233	1233	1233	1233
Studies	65	65	65	65	65	65

Notes: The dependent variable is the t-value of the spillover estimates. Figures in parenthesis are standard errors (SE). Columns A5 report the specific model without adjusting SE, A6 and A9 (mixed-effects multilevel model) using the restricted maximum likelihood, A7 robust SE, and A8 and A10 (clustered data analysis) study-level clustered SE. All columns use inverse variance weights. [a] Coefficients and standard errors are multiplied by a thousand to make the figures easier to read.

Table 4. FAT-PET: regressions excluding observations from the interaction variables

Panel 1:MEM				
	All-studies		**Peer-reviewed studies**	
Variables	*(A8)* t-value	*(B5)* Effect-size	*(A10)* t-value	*(B6)* Effect-size
Bias/FAT	1.232** (2.22)	0.396*** (2.98)	2.151**(2.22)	0.508* (1.88)
Genuine effect/ PET	-0.014** (-2.09)	0.059 (0.59)	-0.051* (-3.15)	-0.074 (-0.67)
Observations	1034	1034	591	591
Studies	64	64	40	40
Panel 2:CDA				
	All-studies		**Peer-reviewed studies**	
Variables	*(A11)* t-value	*(B7)* Effect-size	*(A12)* t-value	*(B8)* Effect-size
Bias/FAT	0.505 (1.64)	0.446** (2.55)	0.547 (1.08)	0.378 (1.00)
Genuine effect/ PET	-0.002 (-0.21)	0.003 (0.06)	-0.007 (-0.13)	-0.039 (-0.43)
Observations	1034	1034	591	591
Studies	64	64	40	40

Notes: See Table 2.

Figures

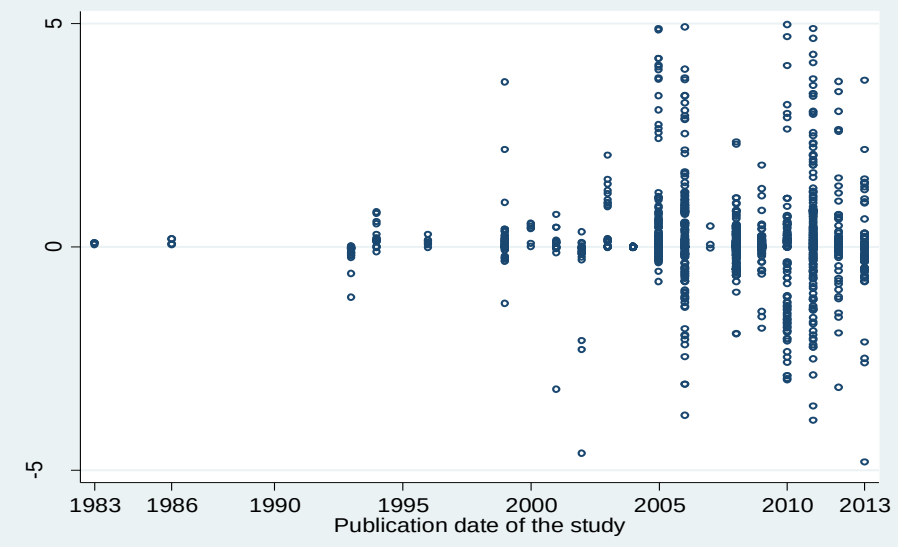

Figure 1. FDI spillover effects by the year of the publication (N=1545).

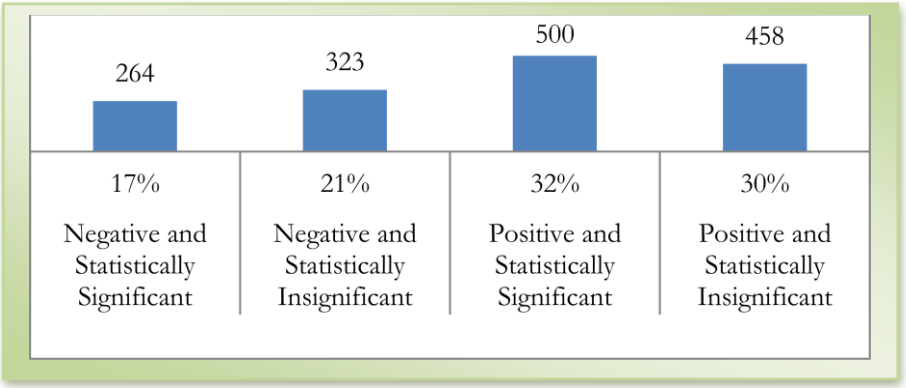

Figure 2. FDI spillover effects reported in 74 studies published in 1983-2013 (N=1545).

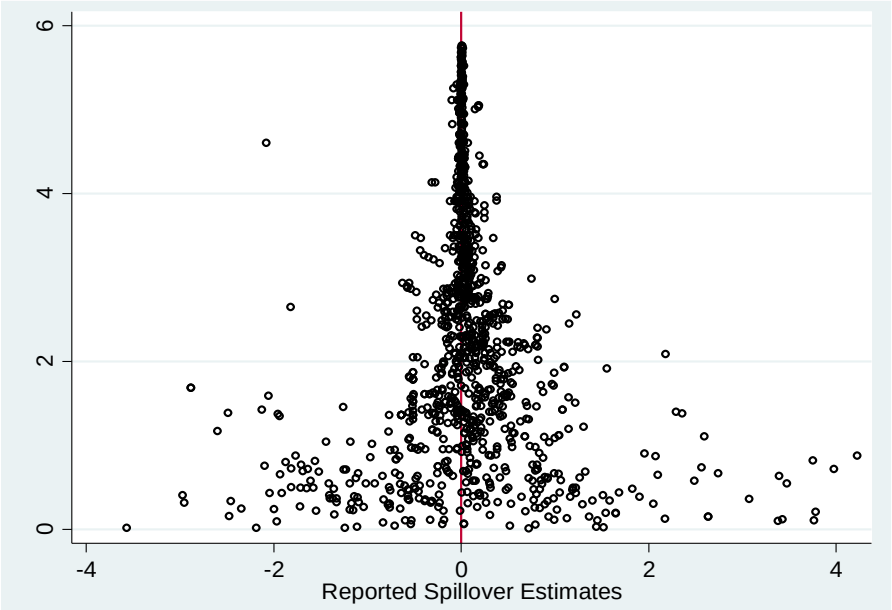

Figure 3. Funnel plot for all-studies (N=1233 from 65 studies).
Note: Instead of excluding extremely high precision values, we use the logarithm of the precision derived from the inverse of the standard error of the reported spillover estimates to allow better eye-o-metrics, i.e., visualization of the graphic images illustrating the relationship between the underlying treatment effects size and their measure of precisions.

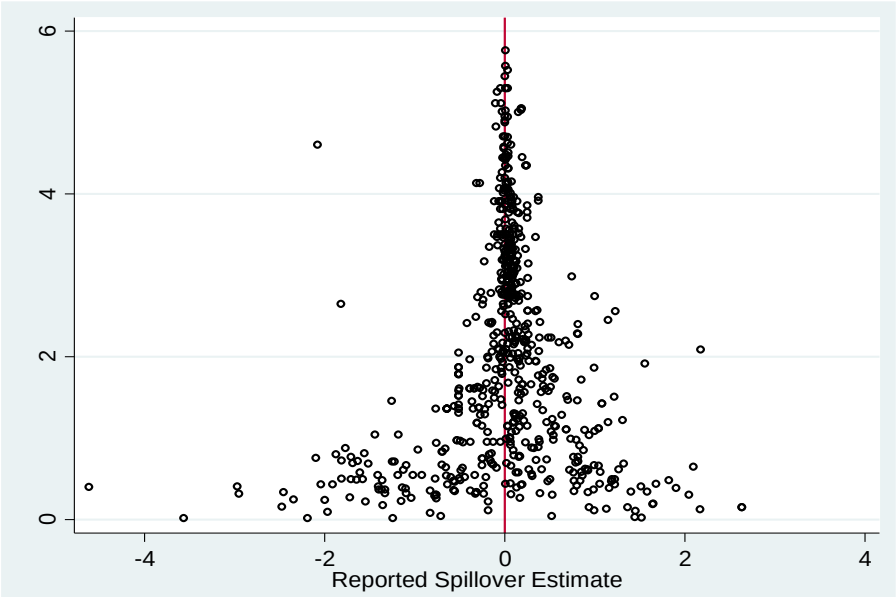

Figure 4. Funnel plot for peer-reviewed studies (N=694 from 41 studies).
Note: See Figure 3

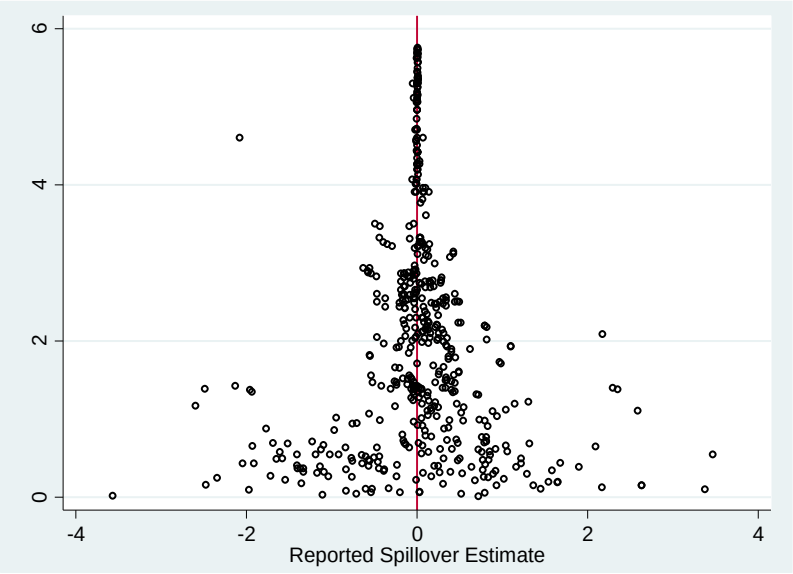

Figure 5. Best practice funnel: all-studies (N=593 from 29 studies).
Note: See Figure 3

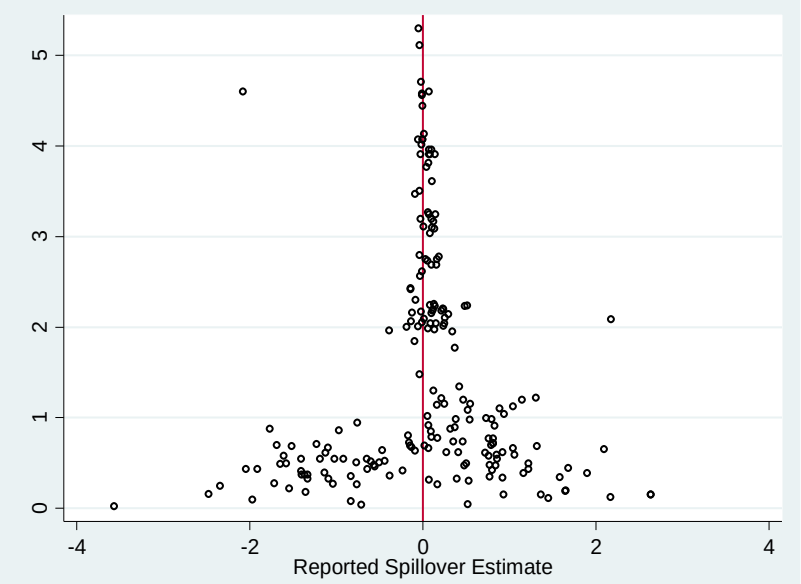

Figure 6. Best practice funnel: peer-reviewed studies (N=234 from 15 studies).
Note: See Figure 3

Annex

Table A.1. Bivariate MRA for FAT-PET: Sensitivity analysis with the inclusion of outliers.

Panel 1:MEM				
	All-studies		**Peer-reviewed studies**	
Variables	t-value	Effect-size	t-value	Effect-size
Bias/FAT	0.877** (2.19)	0.196* (1.83)	1.018 (1.63)	0.181 (1.62)
Genuine effect/ PET	-0.000 (-0.28)	0.129 (1.14)	-0.000 (-0.61)	0.107 (0.67)
Observations	1446	1450	778	782
Studies	69	69	43	43
Panel 2:CDA				
	All-studies		**Peer-reviewed studies**	
Variables	t-value	Effect-size	t-value	Effect-size
Bias/FAT	0.525*(1.76)	0.190* (1.95)	0.565 (1.10)	0.172 (1.66)
Genuine effect/ PET	0.000 (0.35)	0.109 (1.12)	-0.000 (-1.23)	0.076 (0.48)
Observations	1446	1450	778	782
Studies	69	69	43	43

Notes: See Table 2. Test for between-study heterogeneity (*Q*-test) is 64256.49***on 1445 degrees of freedom with p-value less than 0.001and I^2 statistics is 97.8%.