Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

DISSERTATION

Meta-Analysis in International Economics

Author: **Tomáš Havránek**

Advisor: **Doc. Roman Horváth, Ph.D.**

Academic Year: **2012/2013**

Thank you!

During my graduate studies I have benefited from the comments of many colleagues. I thank these good people at the beginning of the chapter they helped me with. Two of them, though, I would like to thank here. My wife and classmate, Zuzana, co-authored two out of three papers included in the dissertation (Chapter 3 and Chapter 4). Together we published seven other papers and have three unfinished projects at the time when I am writing this acknowledgment. Zuzana did much of the work on the projects when looking after our daughter, Kristina. Our son Daniel is due in August, and both are kind to let Zuzana work part-time.

My advisor, Roman, encouraged me to submit one of my first papers to an international journal. The editor of the journal was patient enough to allow several rounds of revisions and eventually accepted the paper (Chapter 2 in the dissertation). It was because of the encouragement and other help from Roman that I chose research for my profession.

Typeset in LaTeX 2_ε using the IES Thesis Template.

Bibliographic Record

Havránek, Tomáš: *Meta-Analysis in International Economics*. Dissertation thesis. Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies. May 2013, pages 107. Advisor: Doc. Roman Horváth, Ph.D.

Contents

Abstract **v**

List of Tables **vi**

List of Figures **vii**

1 Three Tasks of Meta-Analysis **1**

2 The Rose Effect and the Euro: Is the Magic Gone? **9**

2.1 Introduction 10

2.2 Combining the Literature 11

2.3 Publication Bias 14

2.4 Explanatory Meta-Regression 20

2.5 Conclusion 24

References 24

2.A Data Description and Additional Results 31

3 Estimating Vertical Spillovers from FDI **37**

3.1 Introduction 38

3.2 The Spillover Estimates Data Set 39

3.3 The Importance of Publication Bias 43

3.4 What Explains Differences in Spillover Estimates 47

3.4.1 Method Heterogeneity 48

3.4.2 Structural Heterogeneity 49

3.5 Results of the Multivariate Meta-Regression 51

3.5.1 Method Heterogeneity 52

3.5.2 Structural Heterogeneity 57

3.6 Conclusion 59

References 60

3.A Data Description 63

4 Determinants of Horizontal Spillovers from FDI **66**
4.1 Introduction . . . 67
4.2 The Data Set . . . 68
4.3 Why Do Spillover Estimates Differ? . . . 73
4.4 Meta-Regression Results . . . 79
4.5 Publication Bias . . . 85
4.6 Conclusion . . . 88
References . . . 88
4.A Studies Used in the Meta-Analysis . . . 97
4.B BMA Diagnostics . . . 97

Abstract

The dissertation consists of three papers presenting applications of meta-analysis in international economics. The first paper examines the effect of common currency on international trade, while the remaining two papers address the relationship between foreign investment and the productivity of domestic firms. An introductory chapter puts these applications into perspective.

In the first application I present a meta-analysis of the effect of currency unions on trade, focusing on the euro area. I find strong publication bias in the literature. The estimated trade-promoting effect of currency unions other than the euro reaches more than 60%. In contrast, the euro's trade-promoting effect is insignificant when I correct for publication bias. The empirical literature on this topic shows signs of the so-called economics research cycle: the relation between the reported t-statistics and publication years has an inverse U-shaped form.

During 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, in the second application I collect 3,626 estimates of spillovers and review the literature quantitatively. The meta-analysis indicates that model misspecifications reduce the reported estimates. Taking these biases into consideration, the average spillover to suppliers is economically significant, whereas the spillover to buyers is statistically significant but small.

In the third application I collect 1,205 estimates of spillovers from the literature and examine which factors influence spillover magnitude. To identify the most important determinants of spillovers among 43 collected variables, I employ Bayesian model averaging. The results suggest that the most important determinants are the technology gap between domestic and foreign firms and the ownership structure in investment projects. Foreign investors who form joint ventures with domestic firms and who come from countries with a modest technology edge create the largest benefits for the domestic economy.

List of Tables

2.1 Tests of publication bias and the true effect, euro area studies 17
2.2 Tests of publication bias and the true effect, non-euro studies 19
2.3 Explanatory meta-regression analysis 22
2.4 Studies used in the meta-analysis 31
2.5 Acronyms of regression variables 34

3.1 Test of publication bias and the corrected spillover effect 45
3.2 Method heterogeneity in backward FDI spillovers 53
3.3 Structural heterogeneity in backward FDI spillovers 58
3.4 Summary statistics of regression variables, backward spillovers 63

4.1 Description and summary statistics of regression variables 75
4.2 Explaining the differences in the estimates of horizontal spillovers . . 81
4.3 The economic significance of potential spillover determinants 85
4.4 Test of publication bias 87
4.5 List of primary studies 97
4.6 Summary of BMA estimation 97

List of Figures

1.1 Task 1. Correct for publication bias 2
1.2 Task 2. Correct for misspecification bias 4
1.3 Task 3. Explain structural heterogeneity 5

2.1 Forest plot of individual estimates of γ, euro area studies 13
2.2 Funnel plot, all studies 15
2.3 Galbraith plot, euro area studies 17
2.4 Forest plot of individual estimates of γ, non-euro studies 35
2.5 Funnel plot corrected for publication bias, all studies 36

3.1 Funnel plots show selection bias in published studies 56

4.1 Country heterogeneity in the estimates of horizontal spillovers 72
4.2 Method heterogeneity in the estimates of horizontal spillovers 72
4.3 Bayesian model averaging, model inclusion 80
4.4 Posterior coefficient distributions for potential spillover determinants . 83
4.5 Funnel plot 86
4.6 Model size and convergence 98

Chapter 1

Three Tasks of Meta-Analysis

The purpose of this introductory chapter is twofold: to explain the main tasks of meta-analysis in economics and to connect them to the papers included in the dissertation. I see three main tasks of meta-analysis, which is why I include three papers, each representing a particular task. The papers also share a topic, broadly defined—international economics. During my graduate studies I have written or co-written 16 papers published in refereed journals, so the dissertation could be much longer. But I prefer to focus on one field and include three of my most favorite papers; only in this introduction I use examples from other projects as well.

Meta-analysis is a general label describing quantitative surveys of empirical literature. That is, if I take estimates of the treatment effect of a particular drug from twenty clinical trials and compute a simple average, I can call my analysis a meta-analysis (an analysis of analyses; the Greek prefix *meta* here means one level of description higher).[1] Such was the beginning of meta-analysis in medical science: aggregation increases statistical power. Medical scientists, though, soon recognized a potential bias in the aggregation. Some results are easier to get published than others. Researchers need to publish, and may therefore hide some results in their file drawers (Rosenthal 1979), intentionally or not. For example, if negative results end in the file drawer, the average taken from the literature will be biased upwards. The simple average is not enough.

Researchers may hide results in their file drawers for two reasons. First, statistically insignificant estimates are hard to sell. As an empirical researcher I sometimes catch myself being unhappy when my results say I cannot reject the null hypothesis. Insignificant estimates seem uninteresting, difficult to publish. Second, most of us empirical researchers have prior beliefs about the effect we are going to estimate. I sometimes catch myself being unhappy when my results contradict the theory I am

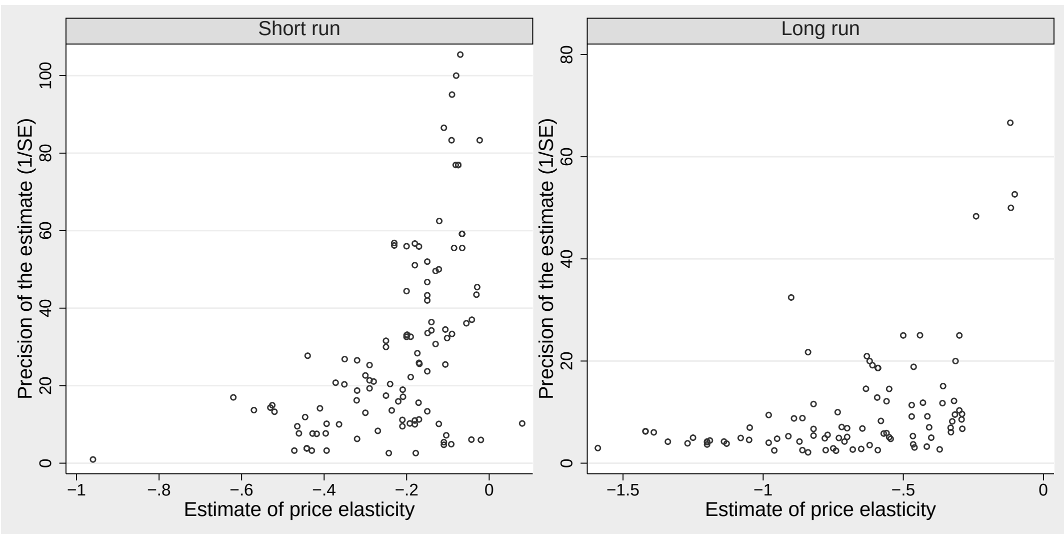
[1] The first medical meta-analysis was Pearson (1904), but the methods of quantitative synthesis were described earlier by astronomers (Airy 1861). My favorite introduction to meta-analysis in economics is Stanley (2001).

working with. The tempting solution is to try a different specification and hope the result will improve. Incentives to publish together with priors about what is right and interesting lead to a distorted public record of empirical results.

An extreme case of the file drawer problem concerns the antidepressant drug Paxil, which was originally reported to be effective by most published clinical trials. When, however, we also take into account unpublished results, the drug does not seem to outperform a sugar pill, and may actually make people more likely to commit suicide (Turner *et al.* 2008). The medical profession recognizes this problem (which is all the more serious that pharmaceutical companies encourage researchers to publish positive results), and the best medical journals now require registration of clinical trials before publication of results, so that we know if results end in file drawers (Krakovsky 2004; Stanley 2005). Similarly the American Economic Association has agreed to establish a registry for randomized control trials “to counter publication bias” (Siegfried 2012, p. 648), with the eventual intention to make registration necessary for submission to the Association’s journals.

The file drawer problem, or *publication bias*, is evident in meta-analysis because of the method’s aggregate point of view. But publication bias distorts inference taken from individual studies and narrative literature surveys as well. Meta-analysis is not a substitute for a good narrative survey, but replaces vague statements like “empirical evidence is mixed” or “there is a consensus in the empirical literature that. . . ” with concrete results. The most important advantage of meta-analysis, I believe, is the possibility to correct for publication bias—when registries of empirical research are missing, we cannot control for the bias in any other way. As the Paxil scandal shows, it may happen that all published studies show a significant effect, but there is none in reality. Thus the first task of meta-analysis: correct for publication bias.

Figure 1.1: Task 1. Correct for publication bias



Medical researchers use funnel plots to detect publication bias, and funnel plots are a useful tool in economics as well. Figure 1.1 shows a funnel plot taken from our paper on the price elasticity of gasoline demand, published in Energy Economics (Havranek *et al.* 2012). The figure is divided into two panels: the left one for the short-run price elasticity and the right one for the long-run elasticity. The horizontal axis shows the magnitude of the estimates of elasticities, while the vertical axis shows the estimates' precision, the inverse of the standard error. The most precise estimates should be close to the underlying average elasticity (the top of the figure), whereas the imprecise estimates at the bottom are more dispersed. The cloud of the estimates should therefore resemble an inverted funnel.

In the absence of publication bias the funnel should be symmetrical. The symmetry follows from the assumptions that researchers make when estimating the elasticity: they assume that elasticities are normally distributed and report t-statistics for their estimates. In fact, however, the funnels we see in Figure 1.1 are not symmetrical. Researchers rarely report positive estimates of the price elasticity, because they do not believe gasoline is a Giffen good (which would imply that demand rises when gasoline becomes more expensive). I share the prior that gasoline is no Giffen good, but I believe we should treat all imprecise estimates, negative and positive, in the same way. Otherwise the simple average, and the conclusion of narrative surveys, gets badly biased toward large negative values. I cannot see the unreported estimates of elasticities, so I concentrate on the most precise reported estimates. The formal corrections of publication bias used in meta-analysis and presented in the dissertation boil down to estimating the top of the inverted funnel.[2] The result is a much smaller elasticity than the simple average would suggest.

I focus on the issue of publication bias in the first paper of the dissertation, Chapter 2, but test for the bias in all chapters. Chapter 2 was published in the Review of World Economics (Havranek 2010). For this paper I received the Olga Radzyner Award by the Austrian National Bank given to selected papers on European economic integration, and the Karel Englis Award by the Czech Economic Society given to selected papers on Czech economic policy.

The two other tasks of meta-analysis, I think, are specific to social sciences, especially economics. Economic research is much more heterogeneous than medical research, and it is therefore not enough to look at the simple average from the literature (even if corrected for publication bias). Since the time meta-analysis was introduced into economics by Stanley & Jarrell (1989), economic meta-analysts has focused on heterogeneity. Why do the results of different empirical papers on the same topic vary so much? Empirical economists must choose among many methods in their estimations; some of the choices seem arbitrary, but most are given by

[2] Stanley *et al.* (2010) show that discarding 90% of the results at the bottom of the funnel can actually improve efficiency. A statistical paradox.

the properties of the data and research topic. For example, ordinary least squares give biased coefficient estimates for endogenous explanatory variables. Using meta-analysis methods we can examine whether method choices have a systematic effect on results and, if they do, compute the average implied by the correct method. Thus the second task of meta-analysis: correct for misspecification bias.

Figure 1.2: Task 2. Correct for misspecification bias

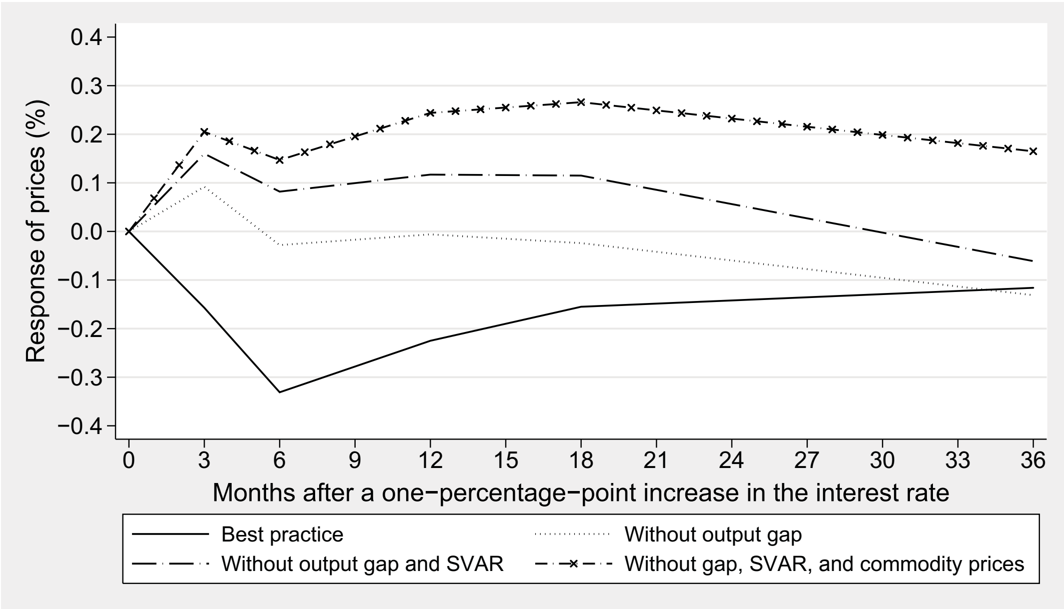
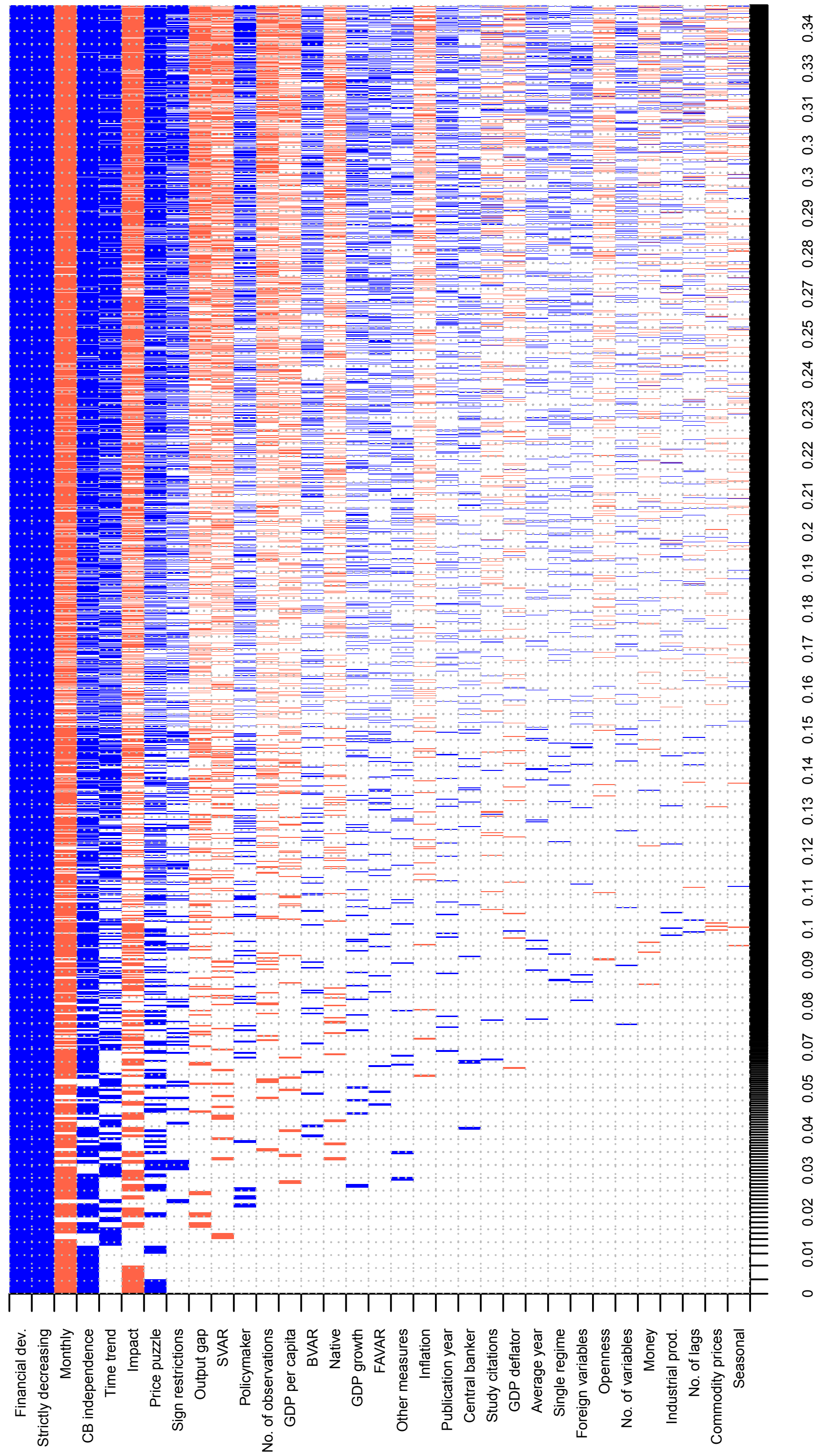


Figure 1.2 is taken from our paper on the effects of monetary policy on the aggregate price level, published in the Journal of Money, Credit and Banking (Rusnak *et al.* 2013). The figure shows average responses of the price level to a monetary contraction. The averages are conditional on different methodologies used in the literature. We define best practice methodology as the inclusion of commodity prices and output gap and use of non-recursive identification in the vector autoregression models employed by primary studies. If best practice is followed, the likely result is that aggregate prices decrease soon after a monetary tightening. In contrast, if a researcher diverts from best practice, she is likely to report the "price puzzle"—the unintuitive finding that prices rise after a monetary contraction.

I focus on misspecification bias in the second paper of the dissertation, Chapter 3, which was published in the Journal of International Economics (Havranek & Irsova 2011). For this paper I received the Medal for Research on Development by the Global Development Network given to selected papers on financing for development written by citizens of emerging countries, the third prize in the Young Economist Competition by the Czech Economic Society given to selected papers written by Czech economists younger than 30 years, and the Economic Research Award by the Czech National Bank given to selected working papers published by the Bank.

Figure 1.3: Task 3. Explain structural heterogeneity



Notes: Response variable: transmission lag. Columns denote individual models; variables are sorted by posterior inclusion probability. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model. The horizontal axis measures cumulative posterior model probabilities. Only the 5,000 models with the highest posterior probabilities are shown.

Meta-analysis can only correct for misspecifications that some studies have overcome. If all studies are misspecified, it gives misspecified results as well—so the result can be interpreted as our best guess based on the existing empirical literature. Meta-analysts often find that the corrected average estimate for many effects in economics is close to zero. One interpretation, advocated by Doucouliagos & Stanley (2013) in a survey of 87 meta-analyses, says that most economic effects are much smaller than commonly thought. Another possibility is that present economic methods are often unable to identify the effect because of misspecification and measurement error (the "iron law of econometrics," Hausman 2001). Innovations in primary studies are essential for meta-analysis. Meta-analysis is essential for robust economic inference.

After I filter out publication bias and misspecifications, I should explain the structural heterogeneity that remains. The effects may differ systematically across industries, countries, and demographic groups; such differences are often economically important. Figure 1.3 is taken from our paper on the determinants of the speed of monetary transmission, published in the International Journal of Central Banking (Havranek & Rusnak 2013) and awarded the first prize in the Young Economist Competition by the Czech Economic Society. The figure displays the results of model averaging, which I believe is especially useful in meta-analysis. Many potential explanatory variables can be included in the model, but I am not sure which ones. The theory does not say much about why, for instance, a particular identification strategy should lead to a systematically different impulse response function. So I estimate models with all combinations of the potential explanatory variables and weight the models by the goodness of fit. In this case we find that a more developed financial system is associated with a slower transmission of monetary policy.

I focus on explaining structural heterogeneity in the third paper of the dissertation, Chapter 4, where I use model averaging, but also discuss structural heterogeneity extensively in Chapter 3. Chapter 4 was published in World Development (Irsova & Havranek 2013) and received the third prize in the Young Economist Competition by the Czech Economic Society.

Meta-analysis, I have noted, is no substitute for good narrative surveys. It complements them with a formal examination of publication bias, misspecifications, and structural heterogeneity. As a graduate student I do not know enough to write a proper narrative survey and do not attempt to do so. Meta-analysis is a formalized method that enables me to focus on one problem in a particular literature, and I usually feel confident about analyzing the problem after several months of study. Although I concentrate on macroeconomics, what I like about meta-analysis is that it allows me to learn from many fields: banking (Irsova & Havranek 2010), monetary policy (Havranek & Rusnak 2013), energy (Havranek *et al.* 2012), development (Havranek & Irsova 2012), and international economics (the remaining three chapters of the dissertation).

References

AIRY, G. B. (1861): *On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations.* Cambridge: MacMillan and co.

DOUCOULIAGOS, H. & T. D. STANLEY (2013): "Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

HAUSMAN, J. (2001): "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives* **15(4)**: pp. 57–67.

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

HAVRANEK, T. & Z. IRSOVA (2011): "Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is." *Journal of International Economics* **85(2)**: pp. 234–244.

HAVRANEK, T. & Z. IRSOVA (2012): "Survey Article: Publication Bias in the Literature on Foreign Direct Investment Spillovers." *Journal of Development Studies* **48(10)**: pp. 1375–1396.

HAVRANEK, T., Z. IRSOVA, & K. JANDA (2012): "Demand for Gasoline is More Price-Inelastic than Commonly Thought." *Energy Economics* **34(1)**: p. 201–207.

HAVRANEK, T. & M. RUSNAK (2013): "Transmission Lags of Monetary Policy: A Meta-Analysis." *International Journal of Central Banking* **(forthcoming)**.

IRSOVA, Z. & T. HAVRANEK (2010): "Measuring Bank Efficiency: A Meta-Regression Analysis." *Prague Economic Papers* **2010(4)**: pp. 307–328.

IRSOVA, Z. & T. HAVRANEK (2013): "Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis." *World Development* **42**: pp. 1–15.

KRAKOVSKY, M. (2004): "Register of Perish." *Scientific American* **291**: pp. 18–20.

PEARSON, K. (1904): "Report on Certain Enteric Fever Inoculation Statistics." *British Medical Journal* **2(2288)**: pp. 1243–1246.

ROSENTHAL, R. (1979): "The 'File Drawer Problem' and Tolerance for Null Results." *Psychological Bulletin* **86**: pp. 638–41.

RUSNAK, M., T. HAVRANEK, & R. HORVATH (2013): "How to Solve the Price Puzzle? A Meta-Analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

SIEGFRIED, J. J. (2012): "Minutes of the Meeting of the Executive Committee: Chicago, IL, January 5, 2012." *American Economic Review* **102(3)**: pp. 645–52.

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. & S. B. JARRELL (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161–70.

STANLEY, T. D., S. B. JARRELL, & H. DOUCOULIAGOS (2010): "Could It Be Better to Discard 90% of the Data? A Statistical Paradox." *The American Statistician* **64(1)**: pp. 70–77.

TURNER, E. H., A. M. MATTHEWS, E. LINARDATOS, R. A. TELL, & R. ROSENTHAL (2008): "Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy." *New England Journal of Medicine* **358**: pp. 252–260.

Chapter 2

The Rose Effect and the Euro: Is the Magic Gone?

Abstract

In this paper I present an updated meta-analysis of the effect of currency unions on international trade, focusing on the euro area. Using meta-regression methods such as the funnel asymmetry test, I find evidence for strong publication bias. The estimated trade-promoting effect of currency unions other than the euro area reaches more than 60%. In contrast, the euro's trade-promoting effect is insignificant when I filter out publication bias from the literature. The empirical literature on this topic shows signs of the so-called economics research cycle: the relation between the reported t-statistics and publication years has an inverse U-shaped form. Explanatory meta-regression (robust fixed effects and random effects), which is able to explain about 70% of the heterogeneity in the literature, suggests that results published by some authors differ consistently from the mainstream output and that study outcomes depend systematically on study design (the usage of panel data, short- or long-run nature, and number of countries in the data set).

The paper was published in the Review of World Economics [2010, 146(2), pp. 241–261]. I thank Roman Horvath, Zuzana Irsova, Tom Stanley, Katerina Smidkova, and seminar participants at Charles University in Prague for their valuable comments. I am especially grateful to an anonymous referee of the Review of World Economics for very useful suggestions that led to a substantial improvement in the quality and readability of the paper. I acknowledge support of the Grant Agency of Charles University (grant 76810). An earlier version of this paper circulated under a more resolute title "Rose Effect and the Euro: The Magic Is Gone." My advisor persuaded me to add the question mark, and the paper was published like that. The most precise title that I can think of now would be "The Estimates of the Euro's Trade Effect Taken Together Imply No Effect."

> Most of the Rose effect literature treats currency unions as magic wands—one touch and intra-currency-union trade flows rise between 5% and 1400%. The only question is: How big is the magic? (Baldwin 2006, p. 36)

2.1 Introduction

Since the pioneering work of Rose (2000) and his result that currency unions increase trade by more than 200%, a whole new stream of literature has emerged and thrived, focusing especially on the euro area in recent years. How much does the euro boost trade among the euro area members? While some researchers are rather skeptical to search for "the one number" (e.g., Richard Baldwin, as the opening quotation suggests), the others keep seeking: in a narrative literature review, Frankel (2008b) estimates the euro's Rose effect to lie between 10% and 15%. Even Baldwin (2006, p. 48) himself talks about 5%–10% and expects the effect to double as the euro matures. This question is very attractive for welfare economists and policy makers: for instance, Frankel (2008a) uses his estimates to give Central and Eastern European countries advice on the timing of their admission to the euro area; and Masson (2008), employing the result that "currency unions double trade," assesses the welfare effects of creating a monetary union in Africa.

One meta-analysis[1] has been written on the subject. Rose & Stanley (2005), using a combined sample of studies on both the euro area and other currency unions, report the general underlying effect to lie between 30% and 90%. The purpose of this paper is to extend the aforementioned work by including new studies and employing different meta-analysis methods, which enables me to concentrate on the effects of the euro and other currency unions separately. It is shown that the distinction between euro and non-euro studies is important since both sub-samples tell a very different story. Twenty-seven new studies were added to the sample, 21 of which focusing on the euro area. Together, there are 61 studies, 28 on the euro area and 33 on other currency unions (see Table 2.4 in the Appendix). I examine publication bias in the literature (Card & Krueger 1995; Stanley 2005a), using the meta-regression approach (Stanley & Jarrell 1989; Stanley *et al.* 2008) and graphical methods (funnel plots, Galbraith plots); the "true" underlying effect corrected for publication bias is estimated as well. The meta-regression analysis (MRA) by Rose & Stanley (2005) is augmented with multiple different techniques (robust estimators and multilevel methods). Explanatory meta-regression methods, including robust meta-regression (see, for example, Bowland & Beghin 2001) and random effects meta-regression (Abreu *et al.* 2005), are used to examine systematic dependencies of results on study design

[1] For an excellent introduction to the methodology of meta-analysis and its application in economics, see Stanley (2001).

and thus to model heterogeneity present in the sample. Moreover, a test for the "economics research cycle" is conducted (novelty and fashion in economics research, see Goldfarb 1995).

The paper is structured as follows: in Section 2.2, the essence of meta-analysis is briefly described and basic properties of the sample of literature are discussed. Section 2.3 focuses on publication selection and search for the true Rose effect beyond publication bias. In Section 2.4, explanatory MRA is conducted. Section 2.5 concludes.

2.2 Combining the Literature

Meta-analysis has its roots in psychology and epidemiology where it has been employed extensively in the last 3 decades (for an extensive introduction, see Borenstein *et al.* 2009). Originally it was used to increase the number of observations and thus statistical power in those fields of medical research where experiments were extremely costly and scarce, or to estimate the "true" effect when the findings were seemingly mixed. Subsequently, this method spread to social sciences, including economics (beginning with Stanley & Jarrell 1989). The essence of meta-analysis is to use all available studies since even biased and misspecified results may carry useful information that can be decoded by the meta-regression approach. Omitting some empirical papers on the Rose effect *ex ante*, as Baldwin (2006) suggests in his narrative review, is thus the opposite of what a meta-analyst would do.

> He [Richard Baldwin] thinks he knows which of the studies are good and which are bad [...], and wants only to count the good ones. The problem with this is that other authors have other opinions as to what is good and what is bad." (Frankel 2006, p. 83).

Fortunately, the meta-regression methods are able to cope with some degree of misspecification bias (Stanley 2008).

The literature estimating the boost to trade due to the creation of a currency union usually employs a variation of the following regression to estimate the trade effect of currency unions, the so-called gravity equation (for a detailed discussion and criticism, see Baldwin 2006):

$$\log T_{ijt} = \alpha_0 + \gamma CU_{ijt} + \chi_1(\log Y_i \log Y_j)_t + \chi_2 \log D_{ij} + \sum_{k=1}^{K} \eta_k X_{ijt} + \epsilon_{ijt}, \qquad (2.1)$$

where T_{ijt} stands for the trade flow between two countries (i and j) in period t, CU is a dummy which equals one if both countries are engaged in a currency union in period t, Y denotes the real GDP, D is the distance between the two countries, and

X denotes other control variables. The actual percent boost to trade due to the formation of a monetary union is thus given by $\gimel \doteq e^{\gamma} - 1$.

Every meta-analysis starts with the selection of literature to be included in the survey. Some meta-analysts use all point estimates (for instance Abreu *et al.* 2005); sometimes it is advised to use only one estimate from each study since otherwise a single researcher could easily dominate the survey (Stanley 2001; Krueger 2003; Stanley 2005b). Moreover, most researchers report many different specifications starting with benchmarks. If all those estimates were included in the meta-analysis, the influence of benchmark cases would be highly exaggerated (however, this can be partly treated by multilevel data analysis or clustering). Researchers themselves also assign very different weights to the particular specifications. Therefore, while including all estimates would enhance degrees of freedom, for this project I prefer to select representative estimates.[2]

I build on the data set provided by Rose & Stanley (2005) which covers a sample of results taken from 34 papers on currency unions' trade effect. The data set, however, contains only 7 studies on the euro area, which does not make it possible to estimate the euro's effect separately. For this reason, additional search was conducted mainly in the EconLit, RePEc, and Google Scholar databases, concentrating especially on new studies estimating the effect of the euro.[3] All papers on the Rose effect containing a quantitative estimate of γ were included, both published and unpublished, extending the sample to the total of 61 studies, including 28 studies on the euro area. The authors' preferred estimates were selected; in case there was no preference expressed, the model with the best fit was chosen. Nevertheless, most authors in this sample reveal their preferences concerning the "best" estimate directly in the abstract or conclusion.

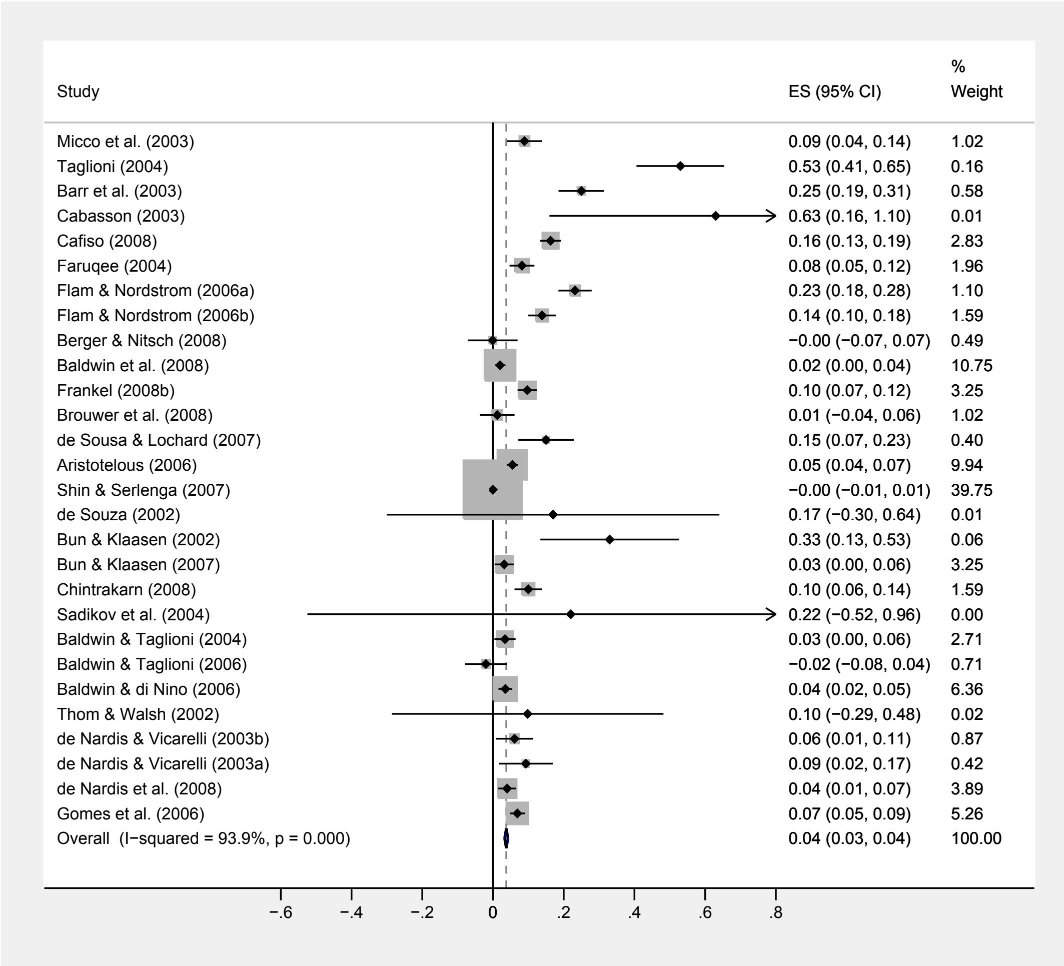
It is generally recognized that the reported Rose effect of the euro is significantly lower than that of other currency unions taken as a whole (Micco *et al.* 2003; Frankel 2008b). Frankel (2008b) tests three possible explanations (euro's youth, bigger size of the euro area economies compared to average members of other monetary unions, and reverse causality unrecognized by the earlier studies), but rejects them one by one. The low estimates of the euro's trade effect thus remain a puzzle. For policy recommendations concerning the euro, in any case, only the estimates derived from the euro area studies should be taken into account. The results of the non-euro papers, however, are useful as well: on the one hand, these studies can serve as a

[2] Two years after this paper was published I believe it would have been better to include all estimates. My student gathered 2,580 estimates of the euro's Rose effect for his bachelor thesis (Polak 2011), but found similar results.

[3] The exact search query used in RePEc was (((currency | monetary) + union) | euro) + trade + (effect | rose) + estimate, abstract search since 2002. The "old" (Rose & Stanley 2005) data were updated—for example, many of the then working papers have been published in a journal since 2005 and their estimates might have slightly changed.

control group; on the other hand, the general underlying Rose effect of other currency unions can be extracted from them.

Figure 2.1: Forest plot of individual estimates of γ, euro area studies



The euro area sample is depicted in Figure 2.1; this type of figure is usually called "forest plot" in medical research. Black dots symbolize individual estimates of γ, horizontal lines show the respective 95% confidence intervals. The traditional method of combining estimates taken from various studies is the standard fixed effects estimator,[4] which weights each observation according to its precision; that is, inverse standard error. The weights constructed on the basis of the inverse-variance method are symbolized by squares with gray fill in the forest plot. The pooled effect estimated by fixed effects is plotted as a vertical dashed line, the solid vertical line symbolizes no effect. Using fixed effects, the pooled estimate of the euro's γ is very low: a mere 0.038 ($\widehat{\Im}$ = 3.87%) with 95% confidence interval CI = (3.36%, 4.39%),

[4]Note that "fixed" and "random" effects estimators in meta-analysis do not correspond to the standard use of these terms in panel data econometrics. For a more detailed explanation, see Abreu *et al.* (2005), Sutton *et al.* (2000), or Chapter 3 in the dissertation.

although it is highly significant (z-stat. = 14.9). These results are not very useful for policy purposes, though, because—among other things starting with heterogeneity and sensitivity to outliers—they do not account for likely publication selection; i.e., preference of editors, referees, or researchers themselves for significant or non-negative results (more on this topic in Section 2.3).

Forest plot of the results of non-euro studies (Figure 2.4 in the Appendix) shows a different picture. The pooled fixed effects estimate is far from zero, namely 0.67 ($\widehat{\mathfrak{I}}$ = 95.42%) with 95% confidence interval CI = (88.89%, 102.18%). Assuming that currency unions double trade, as, e.g., Masson (2008) does when he asseses the welfare effects of forming currency unions in Africa, thus might appear plausible in this respect.

Based on these simple statistics, there is no doubt that the estimates of the Rose effect of the euro and other currency unions are indeed immensely different and that it is not very appropriate to pool them together. Nevertheless, more advanced methods are needed to assess the problem of publication selection and estimate the genuine underlying effect.

2.3 Publication Bias

In his thorough and influential review of the Rose effect literature, Richard Baldwin comments on the meta-analysis of Rose & Stanley (2005):

> The meta-analysis statistical techniques are fascinating, but I don't believe it adds to our knowledge since deep down they are basically a weighted average of all point estimates. (Baldwin 2006, p. 36).

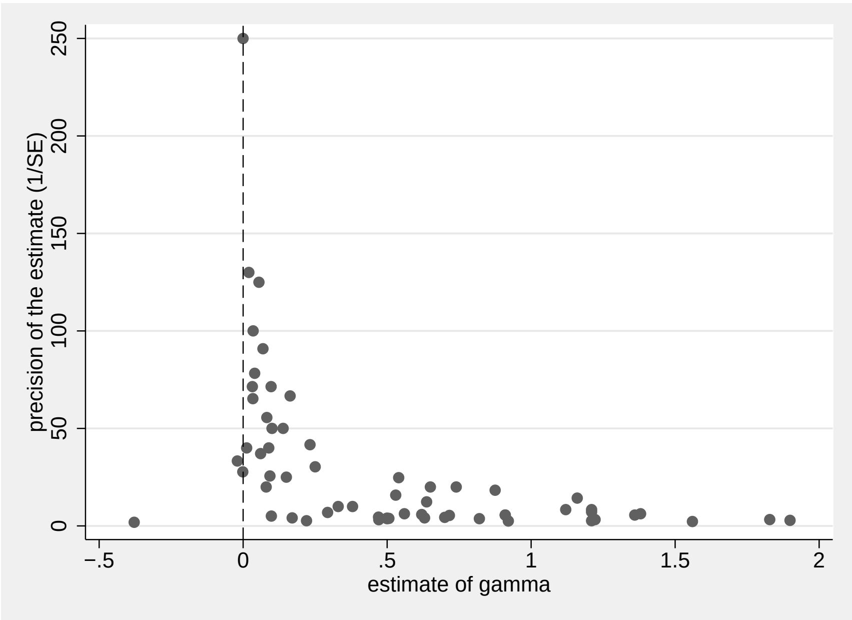
While this statement—or at least its last sentence—may apply to the very simple meta-analysis performed in Section 2.2, it disregards the most important part of Rose & Stanley (2005) as well as of the present study: the MRA, filtering out the publication bias, and modeling the heterogeneity. The search for "the one number" is not the only task of a meta-analyst.

In this section the MRA is employed to test for publication bias and the true underlying Rose effect. Publication selection can take the following two forms (Stanley 2005a):

Type I bias This form of publication bias occurs when editors, referees, or authors prefer a particular direction of results. Negative estimates of γ, for instance, might be disregarded; it would seem quite strange if common currency hampered trade among the monetary union's members. The problem is that even if the true effect was positive, a certain percentage of studies (due to the nature of their data set, methods used, and the laws of probability) should report negative numbers. Otherwise, the average taken from the literature can highly exaggerate the estimated true effect. For instance, Stanley (2005a) shows how the price elasticity of water demand is exaggerated fourfold due to publication bias.

Type II bias The second type of bias arises when statistically significant results are preferred; i.e., when editors choose "good stories" for publication. In this way many questionable effects may be "discovered" and further supported by subsequent research when other authors are trying to produce significant results as well. Intra-industry spillovers from inward foreign direct investment might serve as an example (Görg & Greenaway 2004).

Figure 2.2: Funnel plot, all studies



The presence of type I publication bias is usually investigated employing the so-called funnel plot which shows the estimated effect against its precision (inverse of its standard error, Egger *et al.* 1997). The essence of this visual test is that, in the case of no bias, the shape of the cloud of observations should resemble an inverted funnel; observations with high precision should be concentrated closely to the true effect, while those with lower precision should be more dispersed. Above all, in the absence of type I publication bias, the funnel must be symmetrical.

In Figure 2.2, the funnel plot for all 61 studies is presented. It shows a perfect example of strong publication bias. While positive estimates clearly form one half of a funnel, the left half is almost completely missing as there are only 4 non-positive estimates. The euro area and non-euro studies taken separately resemble an inverted funnel even less. This test can be formalized using a simple MRA (Card & Krueger

1995):

$$\widehat{\gamma_i} = \beta + \beta_0 SE_i + \mu_i, \quad i = 1, \ldots, M, \tag{2.2}$$

where M is the number of studies, β denotes the true effect, and β_0 measures the magnitude of publication bias. However, regression (2.2) is evidently heteroskedastic. The measure of heteroskedasticity is the standard error of the estimate of γ, thus weighted least squares can be performed by running a simple OLS on equation (2.2) divided by the standard error:

$$\frac{\widehat{\gamma_i}}{SE_i} = t_i = \beta_0 + \beta \left(\frac{1}{SE_i} \right) + \vartheta_i. \tag{2.3}$$

The meta-response variable changes to the t-statistic corresponding to the estimate of γ taken from i-th study. A simple t-test on the intercept of (2.3) is then a test for publication bias: the funnel asymmetry test (FAT). Nevertheless, meta-analysis is more vulnerable to data contamination than other fields of empirical economics since it is necessary to choose representative estimates from the literature and collect all data manually. As a robustness check to the basic fixed effects meta-regression, I employ iteratively re-weighted least squares (IRLS) which moreover do not assume normality for hypothesis testing (Hamilton 2006, pp. 239–256). Robust methods in meta-analysis using IRLS are employed, e.g., by Bowland & Beghin (2001) or Krassoi-Peach & Stanley (2009). In the third specification, I allow for a dependence between studies written by the same author; this multilevel approach follows Doucouliagos & Stanley (2009) and uses restricted maximum likelihood estimation. In this case the random intercept model (RIM, only intercept differs across authors) is preferred over the random coefficients model (RCM, both intercept and the coefficient for precision can differ) based on the likelihood ratio (LR) test: corrected p-value of the test is 0.257 in favor of not rejecting the hypothesis that RIM is plausible.[5]

Results of all three tests in the case of the euro area studies are summarized in Table 2.1. In all specifications the intercept is highly significant (t-statistics vary from 2.37 to 4.04). Therefore, the hypothesis of no type I publication bias has to be strongly and robustly rejected, which is all the more remarkable given that these tests are usually believed to have relatively low power (Stanley 2005a). The fact that they all reject the null hypothesis at the 5% level of significance implies that publication bias presents a serious problem for the literature on the euro's Rose effect.

Type II bias can be assessed using the Galbraith plot (Galbraith 1988), which depicts the precision of the estimates of γ against the t-statistics corresponding to those estimates and the (assumed) true effect. If the "true" effect was really true and there was no type II publication bias (selection of papers due to significant results),

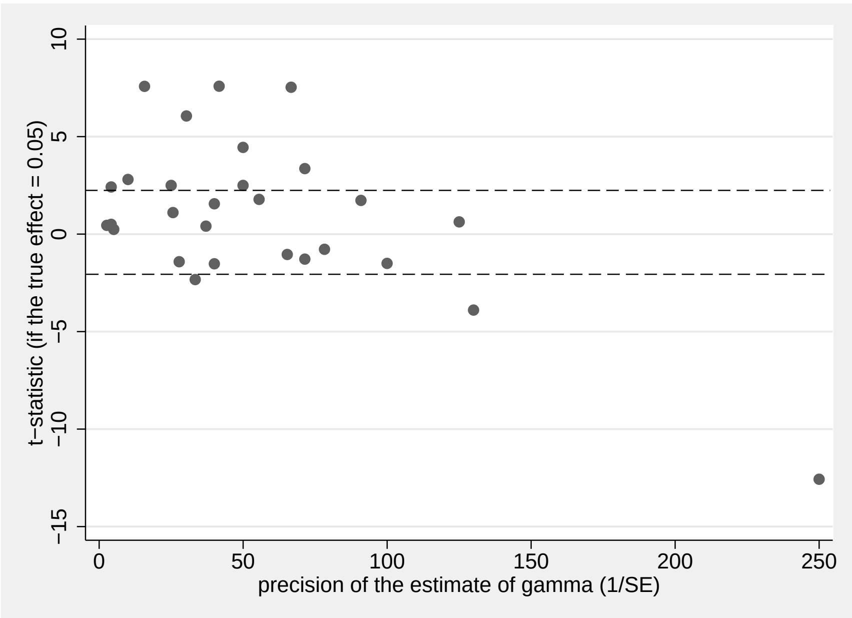
[5] As Rabe-Hesketh & Skrondal (2008, p. 159) note, the LR test is conservative in this case and the correct p-value can be obtained by dividing the original LR p-value by 2.

Table 2.1: Tests of publication bias and the true effect, euro area studies

	FAT-PET	ROBUST	RIM
prec (effect)	0.000667	0.0265	0.00899
	(0.05)	(1.52)	(0.90)
Constant (bias)	3.755***	2.451**	3.517***
	(4.04)	(2.37)	(3.93)
Observations	28	27	28
RMSE	3.169	3.141	

Notes: Meta-response variable: tstat. t-statistics in parentheses (Huber–White heteroskedasticity-robust for FAT-PET). FAT-PET: Funnel assymetry test–precision effect test (fixed effects). ROBUST: Iteratively re-weighted least squares version of fixed effects. RIM: Random intercept model computed using restricted maximum likelihood. *** and ** denote significance at the level of 1 and 5%, respectively.

Figure 2.3: Galbraith plot, euro area studies



only about 5% of the studies' t-statistics should exceed 2 in the absolute value and the cloud of observations should not form any systematic pattern. Figure 2.3 shows the Galbraith plot for the euro area studies (Galbraith plots for all or non-euro studies yield similar results). If the true effect was 0.05, 13 studies out of 28 would report significant results. The goodness of fit test easily rejects the hypothesis of the expected distribution [$\chi^2_{(1)} = 96$, $p < 0.001$]; the null hypothesis is rejected even more powerfully when the true effect is considered to be equal to 0 or 0.1. The t-statistics also show an apparent tendency to decline with rising precision. Therefore, type II bias is clearly present among the euro area studies.

All three methods of detecting type I bias (Table 2.1) can be also used to test for the significance of the true effect beyond publication bias [recall (2.2)]. Specifically, running a t-test on the slope coefficient of (2.3) is denoted as the precision effect test (PET). For euro area studies, the corresponding t-statistic is only 0.05. When robust or random intercept versions of this test of effect are used, the result does not change significantly.[6] This means that, employing the meta-regression methodology, there is not even a slight trace of any true underlying Rose effect of the euro beyond publication bias—compared to the 5%–10% estimate by Baldwin (2006) and 10%–15% estimate by Frankel (2008b). Using meta-regression analysis and the sample of available empirical studies, there is therefore no significant aggregate effect of the euro on trade.

An obvious objection to this approach arises: if the Rose effect of the euro is growing over time (Bun & Klaassen 2002; Baldwin 2006), it is questionable how one can pool together studies written in 2002, when the euro was still young, and papers published, for example, in 2008. It is a potential problem of any meta-analysis. Nevertheless, as can be seen from Section 2.4, explanatory meta-regression does not find any significant relation between the results of euro area studies and time. Also, for instance, Frankel (2008b) concludes that the euro's trade effect has stabilized after a few starting years.

Table 2.2 summarizes the tests of publication bias and the true effect for non-euro studies. Contrary to the previous case, the random coefficients model is preferred over the random intercept model (p-value of the LR test: 0.0009) and is reported in the table—this basically means that I allow publication bias and the effect to vary across researchers. It is apparent that publication bias is weaker than in the previous case; the intercept is significant according to the basic FAT, but not significant in RCM. Nevertheless, as I have noted, these tests of publication bias are known to have relatively low power. Therefore it seems that there is some evidence of publication bias among non-euro studies, although significantly weaker than among the euro area studies. The difference between euro and non-euro studies is the most important

[6] Other robustness checks are available from the author upon request or in the working paper version of this article.

Table 2.2: Tests of publication bias and the true effect, non-euro studies

	FAT-PET	PEESE	RCM
prec (effect)	0.534***	0.634***	0.583***
	(4.08)	(9.83)	(3.52)
SE (bias)		3.567	
		(1.33)	
Constant (bias)	1.712**		1.167
	(2.21)		(1.33)
Observations	33	33	33
RMSE	3.234	3.320	

Notes: Meta-response variable: tstat. t-statistics in parentheses (Huber–White heteroskedasticity-robust for FAT-PET and PEESE). FAT-PET: Funnel assymetry test–precision effect test (fixed effects). PEESE: Precision effect estimate with standard error. RCM: Random coefficients model computed using restricted maximum likelihood. *** and ** denote significance at the level of 1 and 5%, respectively.

finding in this respect—whereas papers on the euro area are plagued by publication bias, the problem is much less serious for the rest of the literature.

PET rejects the null hypothesis of no underlying effect of currency unions other than euro at the 1% level of significance. There is a caveat, though: Stanley (2005b) uses Monte Carlo simulations to show that PET is reliable only if $\sigma^2_\vartheta \leq 2$. Otherwise, the estimate might be exaggerated by misspecification biases. In this case, $H_0 : \sigma^2_\vartheta \leq 2$ is rejected $[\chi^2_{(32)} = 162,\ p < 0,001]$. For this reason, we should employ caution when interpreting the magnitude of the effect, even though the result of PET is supported by its robust version and the random coefficients model. When the "true effect" passes the test for effect, which is the case here, Stanley & Doucouliagos (2007) recommend employing the so-called precision effect estimate with standard error (PEESE) to estimate the magnitude of the effect in question. Contrary to the precision effect test, PEESE assumes that publication bias is related to the variance of the estimates of γ (not standard error). The weighted least squares version thus yields:

$$\frac{\widehat{\gamma}_i}{SE_i} = t_i = \delta_0 SE_i + \delta \left(\frac{1}{SE_i}\right) + \phi_i. \tag{2.4}$$

PEESE estimates the true Rose effect of currency unions other than euro to lie between 65% and 115% with 95% probability. The result is probably somehow exaggerated by misspecification biases, though. Therefore, I consider this number consistent with the previous meta-analysis by Rose & Stanley (2005) who estimate the effect to lie between 30% and 90% (Rose & Stanley, however, used also a few euro area studies in their predominantly non-euro sample).

Figure 2.5 in the Appendix represents the funnel plot of all studies corrected for publication bias [using the filtered-effect test, details can be found in Stanley (2005a) or the working paper version of this article; observations with corrected $|\gamma| > 1$ are cut from the figure]. In contrast to Figure 2.2, the present funnel plot is clearly symmetrical—this is how the literature *should* look like.

2.4 Explanatory Meta-Regression

MRA can also be employed to determine possible dependencies of study results on its design. In fact, it has been the primary focus of most economic meta-analyses since the pioneering work of Stanley & Jarrell (1989). Economics research is usually much more heterogeneous than epidemiology and psychology, where the meta-analysis approach was originally developed. In this respect, MRA is used to assign pattern to heterogeneity.

We gathered 18 meta-explanatory variables that reflect study design and social and other attributes of the authors (see Table 2.5 in the Appendix); 6 of the regressors are assumed to affect publication bias, the rest 12 are expected to influence the estimates of γ directly. The former include researchers' nationality, ranking, gender, panel nature of the data, and year of publication and its square. The latter cover dummies for specific authors, short or long run nature of the study, euro area data, postwar data, number of countries and years in the data set, and impact factor of the journal that the particular study was published in.

All meta-explanatory variables were chosen *ex ante*. I included the meta-explanatory variables used by Rose & Stanley (2005) and added some commonly used variables which are thought to influence publication selection (gender and nationality, for example; for a list of possible regressors affecting publication bias, see Stanley *et al.* 2008), as well as a few experimental regressors. For instance, impact factor was included to ascertain whether articles published in leading journals produce significantly different results from unpublished papers. Inclusion of variable *topfive* (at least one co-author ranks among top 5% economists listed on RePEc) follows a similar logic.

Contrary to the previous sections, now the focus rests on the whole sample because more degrees of freedom are needed; heterogeneity is not so much problematic since it can be modeled to a large extent. There are 61 observations, which is enough for an explanatory meta-regression since sample size in meta-analysis is substantially more effective in increasing the power of hypothesis testing than sample size of original studies (Koetse *et al.* 2010). I employ the FAT-PET method augmented to the following multivariate version (Stanley *et al.* 2008):

$$\frac{\widehat{\gamma}_i}{SE_i} = t_i = \underbrace{\beta_0 + \sum_{j=1}^{J} \theta_j S_{ji}}_{\text{bias}} + \underbrace{\tilde{\beta}}_{\text{pseudo TE}} \left(\frac{1}{SE_i}\right) + \underbrace{\sum_{k=1}^{K} \frac{\delta_k Z_{ki}}{SE_i}}_{\text{controls}} + \vartheta_i, \tag{2.5}$$

where S_j is a set of variables influencing publication bias and Z_k is a set of variables affecting the estimates of γ directly. I refer to this estimator as fixed effects, even though in the strict sense it is not the traditional fixed effects estimator used in meta-analysis: note that variables S_j are not divided by the standard error.

Fixed effects estimates are summarized in column 1 of Table 2.3. As a robustness check, I employ the IRLS version of the model (column 2). The most insignificant meta-regressors are excluded one by one to get a model which contains only variables significant at least at the 10% level. After insignificant variables were excluded, the "economics research cycle hypothesis"[7] was tested by adding the year of publication and its square value. The hypothesis corresponds to the joint significance of these variables and concave shape of the relationship. In this case, $F_{(2,48)} = 3.84$, $p < 0.05$ and the relationship is indeed concave, hence the economics research cycle hypothesis is supported for this type of literature. This becomes even more apparent when IRLS are used [$F_{(2,48)} = 6.74$, $p < 0.01$]. On the other hand, the research cycle hypothesis is rejected when each group of literature is considered separately: $F_{(2,23)} = 1.56$, $p > 0.05$ for non-euro studies and $F_{(2,20)} = 0.21$, $p > 0.05$ for the euro area studies; there is therefore no apparent dependence on time (recall that I used the result that estimates of the euro's Rose effect do not significantly depend on time in Section 2.3). This might suggest that the research cycle identified in the whole literature emerges also due to a larger proportion of the euro area papers among the new studies.

Regression described in column 1 of Table 2.3 is not very well specified, however. The condition number is high (75) indicating possible multicollinearity, Ramsey's RESET rejects the null hypothesis [$F_{(3,45)} = 4.42$, $p < 0.05$], only normality is not rejected [skewness-kurtosis test: $\chi^2_{(2)} = 1.36$, $p > 0.05$]; nevertheless, the model would pass all specification tests if variables *panel*, *year*, and *year2* were excluded. It is apparent that fixed effects MRA was able to model a significant portion of the heterogeneity inside the sample—note the high R^2s: 0.73 and 0.83 for fixed effects and their robust version, respectively.[8] Nevertheless, a lot of heterogeneity still remains unexplained. Testing $H_0 : \sigma^2_\vartheta = 1$ (fixed effects MRA explains heterogeneity well) yields $\chi^2_{(60)} = 276$, $p < 0.001$; for column 1, therefore, H_0 is rejected—the result is qualitatively the same also for the robust specification.

[7] A predictable pattern of novelty and fashion in economics; initial path-breaking results are confirmed by other highly significant estimates, but as the time passes, skeptical results become preferable (Goldfarb 1995; Stanley *et al.* 2008).

[8] However, because these are weighted least squares versions of the original equation, R^2s have to be recomputed to reflect the actual determination of the estimates of γ. For example, in the case of the robust specification, the corrected R^2 reaches 0.68.

Table 2.3: Explanatory meta-regression analysis

	FIXED	ROBUST	RANDOM
prec	0.780***	0.842***	
	(6.16)	(8.15)	
panel	1.606**	1.864***	2.053***
	(2.07)	(2.88)	(4.67)
rose	0.462***	0.328***	0.452***
	(3.45)	(4.06)	(3.62)
nitsch	−0.145***		
	(−4.11)		
baldwin	−0.0814***	−0.359***	
	(−5.48)	(−2.90)	
denardis	−0.0410**		
	(−2.10)		
taglioni		0.299**	
		(2.42)	
euro	−0.700***	−0.779***	−0.563***
	(−5.99)	(−8.39)	(−5.53)
shortrun	0.0349**	0.0391**	
	(2.22)	(2.61)	
countries	−0.00241***	−0.00209***	−0.00108*
	(−3.21)	(−4.12)	(−1.74)
impact	−0.0590***	−0.0413**	
	(−2.79)	(−2.33)	
year	1.178*	1.822***	0.145**
	(1.77)	(3.61)	(2.07)
year2	−0.0801	−0.183***	−0.0122*
	(−1.08)	(−3.32)	(−1.73)
Constant	−1.497	−2.964***	0.278
	(−1.15)	(−2.71)	(1.45)
Observations	61	60	61
R^2	0.725	0.828	
τ			0.0316

Notes: Meta-response variable: tstat for FIXED and ROBUST, gamma for RANDOM. ROBUST: Iteratively re-weighted least squares version of FIXED. t-statistics in parentheses (Huber–White heteroskedasticity-robust for FIXED). Variables *prec*, *rose*, *nitsch*, *baldwin*, *denardis*, *taglioni*, *euro*, *shortrun*, *countries*, and *impact* are assumed to influence the estimates of γ directly. Variables *panel*, *year*, and *year2* are assumed to influence publication bias. ***, **, and * denote significance at the level of 1, 5, and 10%, respectively.

When this is the case, random effects explanatory MRA might be preferable (see, e.g., Abreu *et al.* 2005):[9]

$$\widehat{\gamma}_i = \iota_0 + \sum_{j=1}^{J} \theta_j S_{ji} SE_i + \sum_{k=1}^{K} \delta_k Z_{ki} + \lambda_i + \rho_i, \tag{2.6}$$

where λ_i stands for a normal disturbance term with standard deviations assumed to be equal to SE_i, and ρ_i is a normal disturbance term with unknown variance τ^2 assumed equal across all studies. This between-study variance is estimated using the restricted maximum likelihood method; t-values are computed employing the Knapp & Hartung (2003) modification. The results of random effects MRA are summarized in the third column of Table 2.3; there are much fewer significant explanatory variables than in the previous two specifications.

It is clear from the conducted tests that explanatory meta-regression is as sensitive to method and specification changes as any other field of empirical research. The most important meta-explanatory variables are those that are found significant by all specifications in both fixed and random effects meta-regression (effect on $\widehat{\gamma}$ is shown in parentheses): studies on the euro area ($-$), Rose's co-authorship ($+$), number of countries in the data set ($-$), and usage of panel data ($+$). Some other variables are significant using fixed effects explanatory MRA and its robust version at the same time: short-run nature of the study ($+$), Baldwin's co-authorship ($-$), and impact factor ($-$).

The negative sign for studies on the euro area was expected and corresponds to the results reported by Rose & Stanley (2005), as well as the influence of the number of countries in the data set and usage of panel data. Nevertheless, contrary to the previous meta-analysis, short-run studies are expected to report larger trade effects. Two dummies for authorship were found consistently significant. It does not mean, though, that those authors would produce anyhow tendentious results. Their results only seem to be significantly different from the "mainstream" output. According to the fixed effects meta-regression and its robust version, articles published in leading journals are likely to report marginally lower Rose effects. The latter finding is provocative but should be treated with caution since it is not confirmed by random effects meta-regression.

[9] Monte Carlo experiments suggest that random effects MRA is preferable if heterogeneity is caused by non-constant effect size variance or differences in the true underlying effect across studies. However, when heterogeneity arises due to omitted variable bias—which is realistic in economics—fixed effects estimators should be relied upon (Koetse *et al.* 2010). For this reason, fixed effects MRA is interpreted here as well along with random effects.

2.5 Conclusion

The empirical literature on the trade effect of currency unions is heterogeneous to a large extent. Studies estimating the trade effect of the euro find on average much smaller effects than studies concentrating on other currency unions. The present meta-analysis shows that it is more appropriate to consider these two groups separately in a search for the underlying "true" average effect.

Evidence for publication selection—that is, preference towards statistically significant and positive results—is robust among the papers on euro area and much stronger than for non-euro studies. Narrative literature reviews discussing the trade effect of the euro, which do not take publication selection into account, are hence vulnerable to a substantial upward bias. Meta-regression methods show that, beyond publication bias, there is a significant and large Rose effect of the currency unions other than euro, more than 60%; but no effect at all for the euro area. The absence of an economically important effect is so robust that even some possible mistakes in the process of choosing the authors' preferred estimates cannot significantly change the outcome.

Employing explanatory meta-regression, about 70% of the heterogeneity in the literature can be modeled. The authorship of a particular study is especially important: papers co-authored by Rose tend to find larger effects, papers co-authored by Baldwin are more likely to report smaller estimates. Papers on the euro area find significantly lower effects as well as do long-run studies and studies with a high number of cross-sectional units in their data sets. When panel data are used, the study tends to report larger effects. Studies published in journals with a high impact factor are likely to find a smaller Rose effect; unpublished manuscripts are likely to report large estimates. The Rose effect literature taken as a whole shows signs of the economics research cycle (Goldfarb 1995; Stanley *et al.* 2008): the reported t-statistics are a quadratic concave function of the publication year. One might take a note that the literature seems to have almost completed the circle, and the results, especially on the euro area, are getting close to those "before Rose" when exchange rate volatility was believed to have low influence on international trade (McKenzie 1999).

I do not argue that the euro has no effect on trade. The effects may indeed vary from country to country and industry to industry, as Baldwin (2006) suggests. At the very least, however, there is something wrong with the present Rose effect literature applied on the euro area. The degree of publication bias is striking and the average trade effect of the euro (at least based on the available empirical studies) is probably much lower than we believed, even if what we believed was already twentyfold less than what Rose reported in his famous article.

References

ABREU, M., H. L. F. DE GROOT, & R. J. G. M. FLORAX (2005): "A Meta-Analysis of β-Convergence: the Legendary 2%." *Journal of Economic Surveys* **19(3)**: pp. 389–420.

ADAM, C. & D. COBHAM (2007): "Exchange Rate Regimes and Trade." *Manchester School* **75(1)**: pp. 44–63.

ALESINA, A., R. J. BARRO, & S. TENREYRO (2003): "Optimal Currency Areas." In "NBER Macroeconomics Annual 2002," MIT Press.

ARISTOTELOUS, K. (2006): "Are There Differences Across Countries Regarding the Effect of Currency Unions on Trade? Evidence from EMU." *Journal of Common Market Studies* **44(1)**: pp. 17–27.

BALDWIN, R., V. DI NINO, L. FONTAGÉ, R. A. DE SANTIS, & D. TAGLIONI (2008): "Study on the Impact of the Euro on Trade and Foreign Direct Investment." *Economic Papers 321*, European Comission.

BALDWIN, R. & D. TAGLIONI (2004): "Positive OCA criteria: Microfoundations for the Rose effect." *Technical report*, Graduate Institute of International Studies, Geneva.

BALDWIN, R. & D. TAGLIONI (2006): "Gravity for Dummies and Dummies for Gravity Equations." *NBER Working Papers 12516*, National Bureau of Economic Research, Inc.

BALDWIN, R. E. (2006): "The Euro's Trade Effect." *Working Paper Series 594*, European Central Bank.

BALDWIN, R. E. & V. DI NINO (2006): "Euros and Zeros: The Common Currency Effect on Trade in New Goods." *NBER Working Papers 12673*, National Bureau of Economic Research, Inc.

BARR, D., F. BREEDON, & D. MILES (2003): "Life on the Outside: Economic Conditions and Prospects Outside Euroland." *Economic Policy* **18(37)**: pp. 573–613.

BARRO, R. & S. TENREYRO (2007): "Economic Effects Of Currency Unions." *Economic Inquiry* **45(1)**: pp. 1–23.

BAXTER, M. & M. A. KOUPARITSAS (2006): "What Determines Bilateral Trade Flows?" *NBER Working Papers 12188*, National Bureau of Economic Research, Inc.

BERGER, H. & V. NITSCH (2008): "Zooming Out: The Trade Effect of the Euro in Historical Perspective." *Journal of International Money and Finance* **27(8)**: pp. 1244–1260.

BOMBERGER, W. (2002): "Decolonization and estimates of the time series effect of currency unions." *Technical report*, University of Florida.

BORENSTEIN, M., H. R. ROTHSTEIN, L. V. HEDGES, & J. P. HIGGINS (2009): *Introduction to Meta-Analysis*. Wiley.

BOWLAND, B. J. & J. C. BEGHIN (2001): "Robust estimates of value of a statistical life for developing economies." *Journal of Policy Modeling* **23(4)**: pp. 385–396.

BROUWER, J., R. PAAP, & J.-M. VIANE (2008): "The trade and FDI effects of EMU enlargement." *Journal of International Money and Finance* **27**: pp. 188–208.

BUN, M. J. & F. J. KLAASSEN (2002): "Has the Euro Increased Trade?" *Tinbergen Institute Discussion Papers 02-108/2*, Tinbergen Institute.

BUN, M. J. & F. J. KLAASSEN (2007): "The Euro Effect on Trade is not as Large as Commonly Thought." *Oxford Bulletin of Economics and Statistics* **69(4)**: pp. 473–496.

CABASSON, D. (2003): "Survey about Monetary Unions." *Technical report*, mimeo.

CAFISO, G. (2008): "The Euro's Influence upon Trade—Rose Effect versus Border Effect." *Working Paper Series 941*, European Central Bank.

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

CHINTRAKARN, P. (2008): "Estimating the Euro Effects on Trade with Propensity Score Matching." *Review of International Economics* **16(1)**: pp. 186–198.

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.

EGGER, M., G. D. SMITH, M. SCHEIDER, & C. MINDER (1997): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **316**: pp. 629–634.

ESTEVADEORDAL, A., B. FRANTZ, & A. M. TAYLOR (2003): "The Rise And Fall Of World Trade, 1870-1939." *The Quarterly Journal of Economics* **118(2)**: pp. 359–407.

FARUQEE, H. (2004): "Measuring the Trade Effects of EMU." *IMF Working Papers 04/154*, International Monetary Fund.

FLAM, H. & H. NORDSTRÖM (2006a): "Euro Effects on the Intensive and Extensive Margins of Trade." *Working Paper 1881*, CESifo.

FLAM, H. & H. NORDSTRÖM (2006b): "Trade Volume Effects of the Euro: Aggregate and Sector Estimates." *Seminar Papers 746*, Stockholm University, Institute for International Economic Studies.

FLANDREAU, M. & M. MAUREL (2005): "Monetary Union, Trade Integration, and Business Cycles in 19th Century Europe." *Open Economies Review* **16(2)**: pp. 135–152.

FRANKEL, J. (2006): "Comments on Baldwin (2006) 'The Euro's Trade Effect'." *Working Paper Series 594*, European Central Bank.

FRANKEL, J. (2008a): "Should Eastern European Countries Join the Euro? A Review and Update of Trade Estimates and Consideration of Endogenous OCA Criteria." *Working Paper Series rwp08-059*, Harvard University, John F. Kennedy School of Government.

FRANKEL, J. & A. ROSE (2002): "An Estimate of the Effect of Common Currencies on Trade and Income." *The Quarterly Journal of Economics* **117(2)**: pp. 437–466.

FRANKEL, J. A. (2008b): "The Estimated Effects of the Euro on Trade: Why Are They Below Historical Effects of Monetary Unions Among Smaller Countries?" *NBER Working Papers 14542*, National Bureau of Economic Research, Inc.

GALBRAITH, R. F. (1988): "A note on graphical presentation of estimated odds ratios from several clinical trials." *Statistics in Medicine* **7**: pp. 889–894.

GLICK, R. & A. K. ROSE (2002): "Does a currency union affect trade? The time-series evidence." *European Economic Review* **46(6)**: pp. 1125–1151.

GOLDFARB, R. S. (1995): "The Economist-as-Audience Needs a Methodology of Plausible Inference." *Journal of Economic Methodology* **2(2)**: pp. 201–22.

GOMES, T., C. GRAHAM, J. HELLIWELL, T. KANO, J. MURRAY, & L. SCHEMBRI (2006): "The Euro and Trade: Is there a Positive Effect?" *Technical report*, Bank of Canada.

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.

HAMILTON, L. C. (2006): *Statistics with STATA*. Duxbury Press.

HONOHAN, P. (2001): "Discussion of Rose (2001) 'Currency unions and trade: The effect is large'." *Economic Policy* **33**: pp. 457–461.

KENEN, D. P. (2002): "Currency Unions and Trade: Variations on Themes by Rose and Persson." *Reserve Bank of New Zealand Discussion Paper Series DP2002/08*, Reserve Bank of New Zealand.

KLEIN, M. W. (2005): "Dollarization and trade." *Journal of International Money and Finance* **24(6)**: pp. 935–943.

KNAPP, G. & J. HARTUNG (2003): "Improved tests for a random effects meta-regression with a single covariate." *Statistics in Medicine* **22**: pp. 2693–2710.

KOETSE, M. J., R. FLORAX, & H. D. GROOT (2010): "Consequences of effect size heterogeneity for meta-analysis: A monte-carlo study." *Statistical Methods and Applications* **19(2)**: pp. 217–236.

KRASSOI-PEACH, E. & T. STANLEY (2009): "Efficiency wages, productivity and simultaneity: A meta-regression analysis." *Journal of Labor Research* **30(3)**: pp. 262–268.

KRUEGER, A. B. (2003): "Economic Considerations and Class Size." *Economic Journal* **113(485)**: pp. F34–F63.

LEVI YEYATI, E. (2003): "On the impact of a common currency on bilateral trade." *Economics Letters* **79(1)**: pp. 125–129.

LOPÉZ-CÓRDOVA, J. E. & C. M. MEISSNER (2003): "Exchange-Rate Regimes and International Trade: Evidence from the Classical Gold Standard Era." *American Economic Review* **93(1)**: pp. 344–353.

MASSON, P. R. (2008): "Currency Unions in Africa: Is the Trade Effect Substantial Enough to Justify their Formation?" *The World Economy* **31(4)**: pp. 533–547.

MCKENZIE, M. D. (1999): "The Impact of Exchange Rate Volatility on International Trade Flows." *Journal of Economic Surveys* **13(1)**: pp. 71–106.

MICCO, A., E. STEIN, & G. ORDONEZ (2003): "The currency union effect on trade: early evidence from EMU." *Economic Policy* **18(37)**: pp. 315–356.

MÉLITZ, J. (2001): "Geography, Trade and Currency Union." *CEPR Discussion Papers 2987*, C.E.P.R. Discussion Papers.

MÉLITZ, J. (2008): "Language and foreign trade." *European Economic Review* **52(4)**: pp. 667–699.

de Nardis, S., R. de Santis, & C. Vicarelli (2008): "The Single Currency's Effects on Eurozone Sectoral Trade: Winners and Losers?" *Economics—The Open-Access, Open-Assessment E-Journal* **2(17)**: pp. 1–36.

de Nardis, S. & C. Vicarelli (2003a): "Currency Union and Trade: the Special Case of EMU." *Review of World Economics (Weltwirtschaftliches Archiv)* **139(4)**: pp. 625–649.

de Nardis, S. & C. Vicarelli (2003b): "The Impact of Euro on Trade: The (Early) Effect is not so Large." *ISAE Working Papers 31*, Institute for Studies and Economic Analyses, Rome.

Nitsch, V. (2002a): "Comparing apples and oranges: The trade effect of multilateral currency unions is small." *Technical report*, Bankgesellschaft Berlin.

Nitsch, V. (2002b): "Honey, I Shrunk the Currency Union Effect on Trade." *The World Economy* **25(4)**: pp. 457–474.

Pakko, M. R. & H. J. Wall (2001): "Reconsidering the trade-creating effects of a currency union." *FRB St. Louis Review* **83**: pp. 37–46.

Persson, T. (2001): "Currency unions and trade: how large is the treatment effect?" *Economic Policy* **16(33)**: pp. 433–462.

Polak, P. (2011): *Meta-Analysis in Economics: Application to Measuring the Euro's Trade Effect.* Bachelor thesis, Charles University, Prague.

Rabe-Hesketh, S. & A. Skrondal (2008): *Multilevel and Longitudinal Modeling Using Stata.* Stata Press books. StataCorp LP.

Rose, A. K. (2000): "One money, one market: the effect of common currencies on trade." *Economic Policy* **15(30)**: pp. 7–46.

Rose, A. K. (2001): "Currency unions and trade: the effect is large." *Economic Policy* **16(33)**: pp. 449–461.

Rose, A. K. (2004): "Do We Really Know That the WTO Increases Trade?" *American Economic Review* **94(1)**: pp. 98–114.

Rose, A. K. & C. Engel (2002): "Currency Unions and International Integration." *Journal of Money, Credit and Banking* **34(4)**: pp. 1067–89.

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.

ROSE, A. K. & E. VAN WINCOOP (2001): "National Money as a Barrier to International Trade: The Real Case for Currency Union." *American Economic Review* **91(2)**: pp. 386–390.

SADIKOV, A. M., L. ZENG, P. B. CLARK, S.-J. WEI, & N. T. TAMIRISA (2004): "A New Look at Exchange Rate Volatility and Trade Flows." *IMF Occasional Papers 235*, International Monetary Fund.

SAIKI, A. (2002): "Common currency as an export promoting strategy." *Technical report*, Brandeis University.

SHIN, Y. & L. SERLENGA (2007): "Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors." *Journal of Applied Econometrics* **22(2)**: pp. 361–381.

SHIRONO, K. (2008): "Real effects of common currencies in East Asia." *Journal of Asian Economics* **19(3)**: pp. 199–212.

SMITH, C. (2002): "Currency unions and gravity models revisited." *Reserve Bank of New Zealand Discussion Paper Series DP2002/07*, Reserve Bank of New Zealand.

DE SOUSA, J. & J. LOCHARD (2003): "Do currency unions solve the border effect puzzle?" *Technical report*, University of Paris.

DE SOUSA, J. & J. LOCHARD (2007): "The currency union effect on trade and the FDI channel." *Cahiers de la maison des sciences economiques*, Université Panthéon-Sorbonne (Paris 1).

DE SOUZA, L. V. (2002): "Trade Effects of Monetary Integration in Large, Mature Economies: A Primer on the European Monetary Union." *Kiel Working Papers 1137*, Kiel Institute for the World Economy.

STANLEY, T., C. 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.

STANLEY, T. & H. DOUCOULIAGOS (2007): "Identifying and Correcting Publication Selection Bias in the Efficiency-Wage Literature: Heckman Meta-Regression." *Economics Series 2007/11*, Deakin University.

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

Stanley, T. D. (2005b): "Integrating the Empirical Tests of the Natural Rate Hypothesis: A Meta-Regression Analysis." *Kyklos* **58(4)**: pp. 611–634.

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. & S. B. Jarrell (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161–70.

Subramanian, A. & S.-J. Wei (2007): "The WTO promotes trade, strongly but unevenly." *Journal of International Economics* **72(1)**: pp. 151–175.

Sutton, A. J., K. R. Abrams, D. R. Jones, T. A. Sheldon, & F. Song (2000): *Methods for Meta-analysis in Medical Research.* Chichester: Wiley.

Taglioni, D. (2004): *Monetary union, exchange rate variability and trade.* Ph.D. thesis, University of Geneva.

Tenreyro, S. (2001): "On the causes and consequences of currency unions." *Technical report*, Harvard University.

Thom, R. & B. Walsh (2002): "The effect of a currency union on trade: Lessons from the Irish experience." *European Economic Review* **46(6)**: pp. 1111–1123.

Tsangarides, C. G., P. Ewenczyk, & M. Hulej (2006): "Stylized Facts on Bilateral Trade and Currency Unions: Implications for Africa." *IMF Working Papers 06/31*, International Monetary Fund.

Yamarik, S. & S. Ghosh (2005): "A sensitivity analysis of the gravity model." *The International Trade Journal* **19**: pp. 83–126.

2.A Data Description and Additional Results

Table 2.4: Studies used in the meta-analysis

Study	Euro	Gamma	*t*-stat.	Impact
Rose (2000)	no	1,2100	8,643	1,281
Pakko & Wall (2001)	no	-0,3780	-0,715	0,536
Rose & van Wincoop (2001)	no	0,9100	5,056	2,239
Rose (2001)	no	0,7400	14,800	1,281
Persson (2001)	no	0,5060	1,969	1,281

Continued on next page

Studies used in the meta-analysis (continued)

Study	Euro	Gamma	*t*-stat.	Impact
Honohan (2001)	no	0,9210	2,303	1,281
Mélitz (2001)	no	0,7000	3,043	0,036
Tenreyro (2001)	no	0,4710	1,491	0,018
Nitsch (2002b)	no	0,8200	3,037	0,715
Frankel & Rose (2002)	no	1,3600	7,556	3,688
Thom & Walsh (2002)	yes	0,0980	0,500	0,994
Glick & Rose (2002)	no	0,6500	13,000	0,994
Rose & Engel (2002)	no	1,2100	3,270	0,947
Bun & Klaassen (2002)	yes	0,3300	3,300	0,018
de Souza (2002)	yes	0,1700	0,708	0,018
Nitsch (2002a)	no	0,6200	3,647	0,018
Smith (2002)	no	0,3800	3,800	0,018
Bomberger (2002)	no	0,0800	1,600	0,018
Saiki (2002)	no	0,5600	3,500	0,018
Kenen (2002)	no	1,2219	4,006	0,018
Levi Yeyati (2003)	no	0,5000	2,000	0,302
Estevadeordal *et al.* (2003)	no	0,2930	2,021	3,688
Barr *et al.* (2003)	yes	0,2500	7,576	1,281
Lopéz-Córdova & Meissner (2003)	no	0,7160	3,849	2,239
Micco *et al.* (2003)	yes	0,0890	3,560	1,281
de Nardis & Vicarelli (2003b)	yes	0,0610	2,262	0,018
Cabasson (2003)	yes	0,6300	2,625	0,018
Alesina *et al.* (2003)	no	1,5600	3,545	0,036
de Sousa & Lochard (2003)	no	1,2100	10,083	0,018
de Nardis & Vicarelli (2003a)	yes	0,0930	2,385	0,382
Rose (2004)	no	1,1200	9,333	2,239
Sadikov *et al.* (2004)	yes	0,2200	0,579	0,036
Faruqee (2004)	yes	0,0820	4,556	0,036
Taglioni (2004)	yes	0,5300	8,370	0,018
Baldwin & Taglioni (2004)	yes	0,0340	2,220	0,018
Flandreau & Maurel (2005)	no	1,1600	16,571	0,143
Klein (2005)	no	0,5000	1,852	0,709
Yamarik & Ghosh (2005)	no	1,8285	6,000	0,072
Aristotelous (2006)	yes	0,0550	6,875	0,653
Flam & Nordström (2006a)	yes	0,2320	9,667	0,036

Continued on next page

Studies used in the meta-analysis (continued)

Study	Euro	Gamma	*t*-stat.	Impact
Baldwin & Taglioni (2006)	yes	-0,0200	-0,667	0,036
Baldwin & di Nino (2006)	yes	0,0350	3,500	0,036
Flam & Nordström (2006b)	yes	0,1390	6,950	0,018
Gomes *et al.* (2006)	yes	0,0690	6,273	0,018
Tsangarides *et al.* (2006)	no	0,5400	13,370	0,036
Baxter & Kouparitsas (2006)	no	0,4700	2,136	0,036
Barro & Tenreyro (2007)	no	1,8990	5,410	0,535
Subramanian & Wei (2007)	no	0,6370	7,864	1,541
Adam & Cobham (2007)	no	0,8750	16,010	0,153
Shin & Serlenga (2007)	yes	-0,0003	-0,075	1,094
Bun & Klaassen (2007)	yes	0,0320	2,286	0,732
de Sousa & Lochard (2007)	yes	0,1500	3,750	0,018
Shirono (2008)	no	0,9100	5,056	0,072
Mélitz (2008)	no	1,3800	8,625	0,994
Berger & Nitsch (2008)	yes	-0,0010	-0,028	0,709
Brouwer *et al.* (2008)	yes	0,0120	0,480	0,709
Baldwin *et al.* (2008)	yes	0,0200	2,600	0,036
Cafiso (2008)	yes	0,1630	10,867	0,036
de Nardis *et al.* (2008)	yes	0,0400	3,130	0,072
Frankel (2008b)	yes	0,0970	6,929	0,036
Chintrakarn (2008)	yes	0,1000	5,000	0,072

Notes: Impact factor for the year 2007 obtained from ISI Web of Knowledge.

Table 2.5: Acronyms of regression variables

Variable	Explanation
gamma	Point estimate of common currency's effect on trade.
tstat	t-statistic corresponding to gamma.
SE	Standard error of the estimates of *gamma*.
prec	Inverse of *SE*.
	Moderator variables affecting publication bias
woman	= 1 if there is a woman among co-authors, zero otherwise.
usa	= 1 if all co-authors are Americans (based on current address).
topfive	= 1 if at least one co-author ranks among top 5% in at least 10 categories on RePEc.
panel	= 1 if the study uses panel data with N > T.
year	Publication year−2000.
year2	Variable *year* squared.
	Moderator variables affecting gamma directly
rose	= 1 if Rose is a co-author.
nitsch	= 1 if Nitsch is a co-author.
baldwin	= 1 if Baldwin is a co-author.
denardis	= 1 if de Nardis is a co-author.
taglioni	= 1 if Taglioni is a co-author.
tenreyro	= 1 if Tenreyro is a co-author.
euro	= 1 if the study concentrates on the euro area.
shortrun	= 1 if the study has short-run character.
countries	Number of countries in the data set.
years	Number of years in the data set.
postwar	= 1 if postwar data are used.
impact	Impact factor of the journal where the study was published. Journals without an impact factor obtain weights corresponding to 50% of the lowest impact factor in this sample. Working papers by NBER, ECB, European Commission, CESifo, and CEPR obtain 25%. Other unpublished manuscripts get 12.5%.

Figure 2.4: Forest plot of individual estimates of γ, non-euro studies

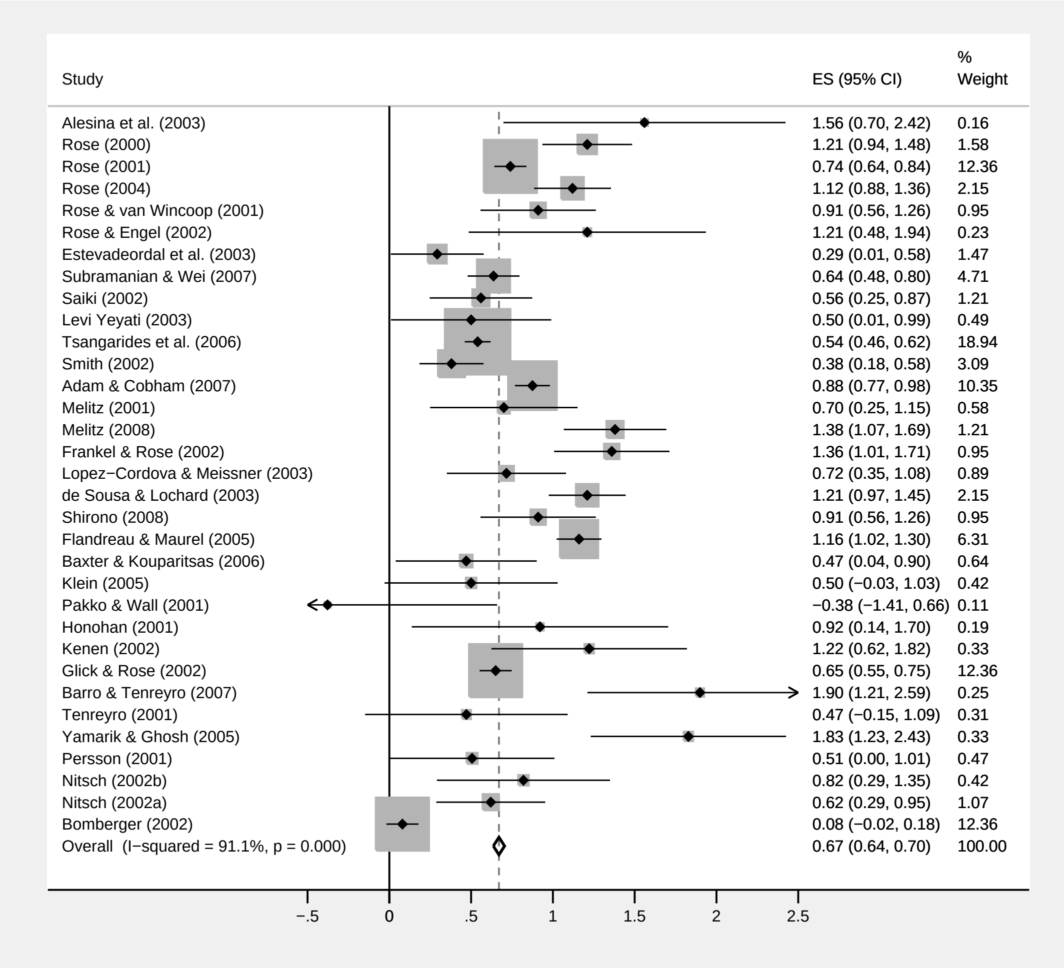
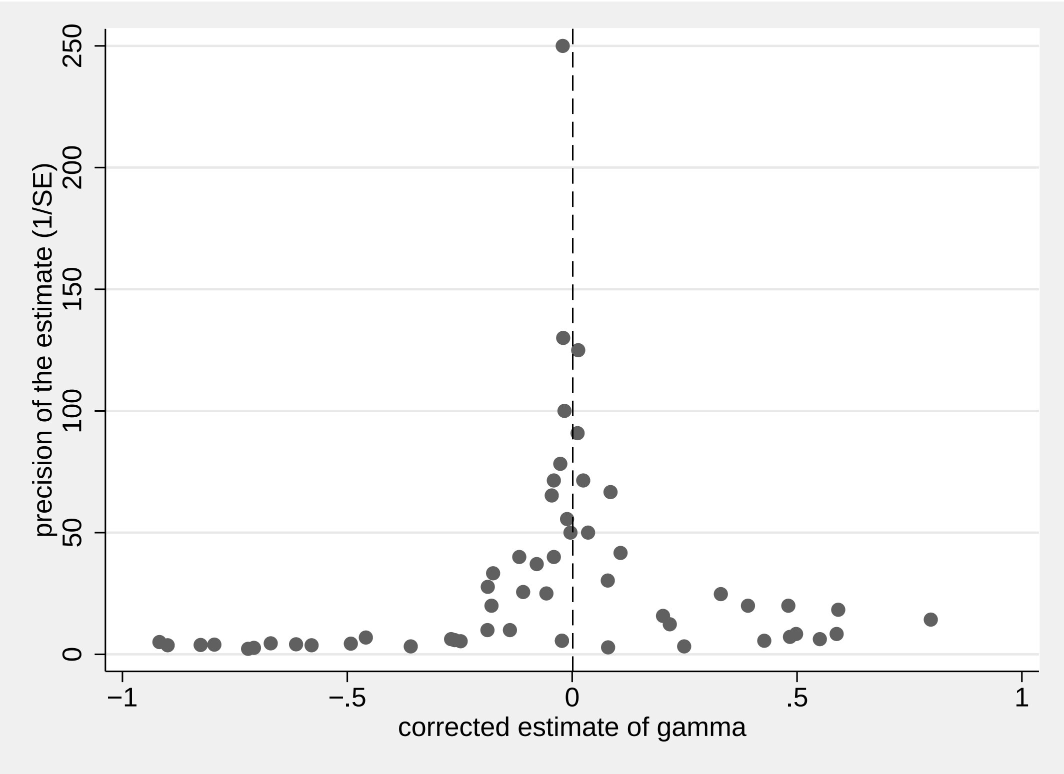


Figure 2.5: Funnel plot corrected for publication bias, all studies



Chapter 3

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

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 economically significant, whereas the spillover to buyers is statistically significant but small. 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.

The paper was co-authored with Zuzana Irsova and published in the Journal of International Economics [2011, 85(2), pp. 234–244]. We thank Bruce Blonigen (the editor), Michal Hlavacek, Roman Horvath, Beata Javorcik, Lubomir Lizal, Tom Stanley, Pavel Vacek, two anonymous reviewers, and seminar participants at Charles University and the Czech National Bank for their helpful comments. The paper benefited from discussions at the Meta-Analysis of Economics Research Colloquium, Conway (Arkansas), 2010, the International Conference on Macroeconomic Analysis and International Finance, Rethymno (Greece), 2011, and the Annual Global Development Conference, Bogota (Colombia), 2011. Tomas Havranek acknowledges support of the Grant Agency of Charles University (grant 76810). Zuzana Irsova acknowledges support of the Grant Agency of the Czech Republic (grant P402/11/0948). An earlier version of this paper circulated under the title "Which Foreigners Are Worth Wooing? A Meta-Analysis of Vertical Spillovers from FDI." An online appendix with data, a Stata program, and a list of excluded studies is available at `meta-analysis.cz/spillovers`.

3.1 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 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.

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. For example, within our meta-analysis framework, we show it is possible to examine the predictions of the theoretical model by Rodriguez-Clare (1996), which 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 source countries in terms of the variety of intermediate goods produced. To test these hypotheses empirically we take the advantage of 57 vertical spillover studies providing estimates for many countries and different types of investors.

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 3.2 briefly describes how spillovers are estimated and explains how we collected the estimates. Section 3.3 examines the extent of publication selection in the literature. Section 3.4 introduces variables assumed to explain heterogeneity in vertical spillovers. Section 3.5 examines how spillover estimates are affected by these variables, and quantifies the underlying effect beyond publication bias and misspecifications. Section 3.6 concludes.

3.2 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:

$$\ln \text{Productivity}_{ijt} = e_0^h \cdot \text{Horizontal}_{jt} + e_0^b \cdot \text{Backward}_{jt} + e_0^f \cdot \text{Forward}_{jt} + \\ + \boldsymbol{\alpha} \cdot \text{Controls}_{ijt} + u_{ijt}, \quad (3.1)$$

where i, j, and t denote firm, sector, and time subscripts; and *Controls* denote a vector of either sector- or firm-specific control variables. 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.

Since the dependent variable of equation (3.1) is in logarithm and the linkage variables are ratios, the estimates of coefficients e_0^h, e_0^b, and e_0^f can be interpreted as semi-elasticities and thus constitute the natural common metric for the economic

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

effect of spillovers. Semi-elasticities approximate the percentage increase in the productivity of domestic firms following an increase in the foreign presence of one percentage point:

$$e_0 \approx (\%\text{ change in productivity})/(\text{change in foreign presence}),$$
$$\text{foreign presence} \in [0, 1]. \quad (3.2)$$

For instance, the estimate of backward spillovers $e^b = 0.1$ would imply that a 10-percentage-point increase in foreign presence is associated with a 1% increase in the productivity of domestic firms in upstream sectors. The estimates are directly comparable across studies that use the log-level specification. Within this basic framework, however, researchers use different methodologies and data sets, which cause substantial differences in results. We address these differences in Section 3.4 by introducing variables that capture method heterogeneity.

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 previous 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 also excluded a few studies that estimated vertical spillovers but did not define linkages as a ratio and thus could not be used to compute the 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 only 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

[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.

specifications we evaluated semi-elasticities at sample means. Other studies use the 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-elasticities 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 who 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-elasticities 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 the 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]

[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

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.

3.3 The Importance of Publication Bias

An arithmetic average of the results reported in the literature will be a biased estimate of the true spillover effect 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 (DeLong & 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.

If the spillover literature is free of publication bias, the reported estimates of semi-elasticities (spillover effects) will be randomly distributed around the true effect. If, in contrast, some estimates fall into the "file drawer" (Rosenthal 1979) because they are insignificant or have an unexpected sign, the reported estimates will be correlated with their standard errors. 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. Hence, publication bias manifests as a systematic relation between the reported effects and the corresponding standard errors (Card & Krueger 1995; Ashenfelter *et al.* 1999):

$$e_i = e_0 + \beta_0 \cdot Se(e_i) + u_i, \tag{3.3}$$

where e_i denotes the reported estimate of a semi-elasticity, e_0 denotes the true spillover effect, β_0 measures the strength of publication bias, $Se(e_i)$ is the standard error of e_i, and u_i is a normal disturbance term. The true spillover (e_0) in this specification is already corrected for publication bias: the bias is "filtered out" since e_0 can be thought of as the average spillover effect conditional on the estimates' standard errors being close to zero. The correction for publication bias is analogical to taking the uncorrected estimate (the arithmetic average of spillover coefficients) and subtracting the estimated publication bias (the estimate of β_0 times the average standard error of spillover coefficients).

Because specification (3.3) is heteroscedastic by definition (the explanatory variable is a sample estimate of the standard deviation of the dependent variable), in practice it is usually estimated by weighted least squares (Stanley 2005; 2008):

$$e_i/Se(e_i) \equiv t_i = \beta_0 + e_0 \cdot 1/Se(e_i) + \xi_i. \tag{3.4}$$

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`).

Note that now the dependent variable changes to the t-statistic of the estimate of a semi-elasticity, the constant measures publication bias, and the slope coefficient measures the true semi-elasticity. Specification (3.4), often called the "meta-regression," has a convenient interpretation: if the true semi-elasticity (e_0) is zero and if only positive and significant estimates of spillovers are reported, the estimated coefficient for publication bias (β_0) will approach two, the most commonly used critical value of the t-statistic. Therefore, values of β_0 close to two would signal extreme publication bias and would be consistent with the case when all studies reported positive and significant estimates of spillovers, but the true spillover was zero. Monte Carlo simulations and many recent meta-analyses suggest that specification (3.4) is effective in filtering out publication bias and estimating the true effect (Stanley 2008).

Since we use more than one estimate of spillovers from each study, it is important to take into account that estimates within one study are likely to be dependent (Disdier & Head 2008). Therefore, (3.4) is likely to be misspecified. A common remedy is to employ the mixed-effects multilevel model, which allows for within-study dependence or, in other words, unobserved between-study heterogeneity (Doucouliagos & Laroche 2009; Doucouliagos & Stanley 2009):

$$t_{ij} = \beta_0 + e_0 \cdot 1/Se(e_{ij}) + \zeta_j + \epsilon_{ij}, \tag{3.5}$$

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}). Regression results are reported in Table 3.1 in three panels, one panel for each type of spillover. In Column 1 estimates collected from all studies, published and unpublished, are included in the regressions. The constants in these regressions are insignificant, which suggests that all types of spillover are free of publication bias if both unpublished and published studies are considered together. 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 might polish even preliminary versions of their papers.

When we only consider estimates from studies published in peer-reviewed journals (Column 2 of Table 3.1), we detect publication bias for backward spillovers, but not for forward and horizontal spillovers. Although the test for publication bias among the estimates of backward spillovers is only significant 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 (Egger *et al.* 1997; Stanley 2008); Egger *et al.* (1997) explicitly recommends employing the more liberal 10% level of significance when using this test.

Table 3.1: Test of publication bias and the corrected spillover effect

Panel A – Backward spillovers	Dependent variable: spillover t-statistic		
	All	Published	Homogeneous
Publication bias			
Constant	-0.0255	1.083^{*}	-1.481
	(0.496)	(0.656)	(0.942)
Spillover effect corrected for bias			
1/(Standard error of the estimate of spillover)	0.168^{***}	0.178^{***}	0.307^{***}
	(0.0241)	(0.0295)	(0.0380)
Observations	1311	370	568
Studies	55	26	39
Panel B – Forward spillovers	Dependent variable: spillover t-statistic		
	All	Published	Homogeneous
Publication bias			
Constant	0.729	-0.437	1.657
	(0.776)	(1.033)	(1.632)
Spillover effect corrected for bias			
1/(Standard error of the estimate of spillover)	0.0872^{***}	0.258^{***}	0.0669^{**}
	(0.0287)	(0.0454)	(0.0288)
Observations	1030	241	591
Studies	44	19	30
Panel C – Horizontal spillovers	Dependent variable: spillover t-statistic		
	All	Published	Homogeneous
Publication bias			
Constant	0.363	0.512	0.818
	(0.295)	(0.498)	(0.500)
Spillover effect corrected for bias			
1/(Standard error of the estimate of spillover)	0.00466	0.0137	0.000549
	(0.00722)	(0.00837)	(0.0127)
Observations	1154	305	471
Studies	52	27	37

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors in parentheses.

Notes: The table contains the results of regression (3.5): $t_{ij} = \beta_0 + e_0 \cdot 1/Se(e_{ij}) + \zeta_j + \epsilon_{ij}$. Estimated by the mixed-effects multilevel model using restricted maximum likelihood. "All" denotes all spillover estimates. "Published" denotes published estimates. "Homogeneous estimates" denote estimates taken from specifications that use firm-level panel data, log-level regression, and the standard definition of spillover variables.

The magnitude of the coefficient for publication bias in published results on backward spillovers is approximately 1.1, which signals strong selection efforts: recall that values close to two would be associated with extreme publication bias, and the value found here is considered "substantial" in the survey of economics meta-analyses by Doucouliagos & Stanley (2008). 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). Indeed, since the results concerning backward spillovers determine the main message of the study, they are more likely to be polished.

The importance of publication bias for inference concerning the magnitude of spillovers is best demonstrated by comparing the average uncorrected and corrected spillover effect. The arithmetic average of all published estimates of backward spillovers is 0.88. In contrast, the corrected spillover effect based on estimates from published studies (resulting from the meta-regression reported in Column 2 of Panel A in Table 3.1) is only 0.178. In other words, because of publication bias the average estimate of spillovers reported in peer-reviewed journals is exaggerated fivefold. This simple example shows how dangerous it is to ignore publication bias; therefore, we will correct for the bias throughout the analysis.[4]

The estimated effects corrected for publication bias (the slope coefficients reported in Table 3.1) are consistently positive and significant across all specifications for backward and forward spillovers, but for horizontal spillovers the effect is not significantly different from zero. To get a flavor of the likely magnitude of backward and forward spillovers before turning to more advanced analysis, we prefer to use a more homogeneous subset of data that only consists of estimates which come from firm-level panel-data studies, which use the standard definition of spillover variables, and for which no additional computation was needed (Column 3 of Table 3.1). These preliminary estimates suggest an effect of 0.307 for backward spillovers and 0.067 for forward spillovers. In other words, a 10-percentage-point increase in foreign presence is on average associated with a 3.1% increase in the productivity of domestic firms in upstream sectors. For domestic firms in downstream sectors the increase in productivity is only 0.7%.

In sum, when we account for publication bias and unobserved heterogeneity, the literature suggests that backward spillovers are economically important, forward spillovers are statistically significant but small, and horizontal spillovers are insignificant. Nevertheless, these results are averaged across all countries and methods, and we need multivariate analysis to explain the vast differences in the reported estimates. The estimates may depend systematically on misspecifications or other quality aspects of primary studies. In the following sections, focusing only on back-

[4]More discussion of publication bias, including additional evidence and robustness checks, is available in the working-paper version of this article (Havranek & Irsova 2010).

ward spillovers as the most important spillover channel, we will estimate the effect implied by best-practice methodology and describe spillover determinants.

3.4 What Explains Differences in Spillover Estimates

To investigate whether there is a systematic pattern of heterogeneity in the spillover literature, we augment equation (3.3) with variables that may potentially influence the reported magnitude of spillovers. Again as in Section 3.3 we divide the resulting equation by the standard error of the spillover estimates 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):

$$t_{ij} = \beta_0 + e_0 \cdot 1/Se(e_{ij}) + \boldsymbol{\beta}\mathbf{x}'_{ij}/Se(e_{ij}) + \zeta_j + \epsilon_{ij}, \tag{3.6}$$

where $\mathbf{x}_{ij} = (x_{1ij}, \ldots, x_{pij})$ is the vector of variables potentially influencing spillover estimates, 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 variables $\mathbf{x}$.

As a robustness check of the mixed-effects multilevel model used to estimate (3.6), OLS with standard errors clustered at the study level is usually employed (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, in contrast, gives each study approximately the same weight if the between-study heterogeneity is high (Rabe-Hesketh & Skrondal 2008, p. 75). We report both models, although the mixed-effects model is preferred.

We explore two potential general sources of heterogeneity. First, since previous meta-analyses on horizontal spillovers (and economics meta-analyses in general) often find that reported results are systematically affected by study design, we explore how the use of different methods affects spillover estimates. We label this source of systematic differences in reported estimates *method heterogeneity*. Second, we test the implications of the theoretical model by Rodriguez-Clare (1996) and investigate other potential determinants of spillovers 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 the underlying spillover coefficients *structural heterogeneity*.

Table 3.4 in the Appendix presents the descriptions and summary statistics for all variables assumed to explain method and structural heterogeneity. Variables explaining method heterogeneity are divided into four blocks: data characteristics represent 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. Variables explaining structural heterogeneity are divided into three blocks: host-country characteristics represent aspects of the country for which the particular spillover coefficient was estimated, foreign-firm characteristics are dummy variables representing properties of the firms used to compute linkages, and local-firm characteristics represent the sector of local firms that were included in the spillover regression.

3.4.1 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 with backward spillovers, the proxy for foreign presence (most studies use share in output, others share 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 added, or labor productivity as the dependent 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 peer-reviewed 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.

3.4.2 Structural Heterogeneity

Host-country characteristics The theoretical model of Rodriguez-Clare (1996) 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. As suggested by Rodriguez-Clare (1996), the costs of communication 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 decomposition of inward FDI stocks—but such information is not always directly available. Therefore, as a first step, we use decompositions of outward FDI positions of OECD countries provided by the OECD's International Direct Investment Statistics. (For this and all other host-country characteristics, we select values from 1999, the median year of the data used in primary studies.) In 1999, OECD countries accounted for more than 85% of the world stock of outward FDI. We additionally obtain data 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 geographic decompositions 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. 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 the decompositions of inward FDI 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 host-country characteristics 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).

Foreign-firm characteristics The next structural variables are dummies capturing the degree of foreign ownership used to define foreign presence. 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.

Local-firm characteristics 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.

3.5 Results of the Multivariate Meta-Regression

We begin the multivariate analysis by including all explanatory variables introduced in Section 3.4 into the regression. This general model with 36 method and 8 structural variables is not reported, but is available on request. For method variables, in contrast to structural variables, no theory exists that would determine which of them are important and what sign they should have. Thus, 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. We always include method and structural variables together in the regressions, but for ease of exposition report the results separately: the results for method variables are reported in Table 3.2 and the results for structural variables in Table 3.3.

In the specification reported in Column 1 of both tables all structural variables are included. The specifications in Columns 2 and 3 omit some of them to avoid the relatively high correlations between some host-country characteristics, but the coefficients do not change a lot. Two structural variables are insignificant in Columns

1–3. Excluding these variables yields our preferred model reported in Column 4; that is, the model without redundant variables. This model is then re-estimated using OLS with standard errors clustered at the study level (Column 5). A few method and structural variables become somehow less significant when OLS is used (their new p-values range between 0.1 and 0.2), but many of them would become significant at standard levels when country-level instead of study-level clustering was used for OLS.

3.5.1 Method Heterogeneity

Table 3.2 shows that seventeen variables reflecting the characteristics of the data, specification, estimation, and quality are significant, suggesting that results of spillover regressions 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 dependent variable, and the method of computing TFP matter. Furthermore, 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.

By the definition of FDI spillover regressions most researchers assume that the 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 a 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 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

Table 3.2: Method heterogeneity in backward FDI spillovers

	Dependent variable: t-statistic of the estimate of spillover				
	1-ME	2-ME	3-ME	4-ME	5-WLS
Constant	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 data	1.206^{***}	1.213^{***}	1.224^{***}	1.193^{***}	1.187^{***}
	(0.145)	(0.140)	(0.145)	(0.144)	(0.190)
Average year of data	0.0349^{***}	0.0236^{***}	0.0277^{***}	0.0323^{***}	0.0301^{***}
	(0.00789)	(0.00719)	(0.00754)	(0.00763)	(0.00837)
Amadeus database	-0.686^{***}	-0.489^{***}	-0.861^{***}	-0.680^{***}	-0.603^{***}
	(0.0950)	(0.0855)	(0.0874)	(0.0946)	(0.127)
Specification characteristics					
Foreign presence in employment	-0.168^{*}	-0.149^{*}	-0.131	-0.158^{*}	-0.323^{*}
	(0.0929)	(0.0825)	(0.0930)	(0.0921)	(0.171)
Control for sector competition	-0.315^{***}	-0.353^{***}	-0.368^{***}	-0.333^{***}	-0.306^{***}
	(0.0673)	(0.0664)	(0.0655)	(0.0649)	(0.106)
Control for downstream 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 estimation	-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 effects	0.119^{***}	0.140^{***}	0.135^{***}	0.128^{***}	0.117^{*}
	(0.0401)	(0.0380)	(0.0393)	(0.0393)	(0.0617)
Estimated in 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 co-author	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

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors in parentheses.

Notes: The table contains the results of regression (3.6). Structural variables are included in all specifications and reported in Table 3.3. Column 1: all structural variables are included. Column 2: *Technology gap*, *Financial development*, and *Fully owned subsidiaries* are excluded. Column 3: *Distance to source countries*, *Trade openness*, and *Patent rights* are excluded. Columns 4 and 5: *Patent rights* and *Partially owned subsidiaries* are excluded. All explanatory variables are described in detail in Table 3.4. ME = mixed-effects multilevel model. WLS = weighteed least squares with standard errors clustered at the study level.

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). For comparison, this is about three times less than the average spillover effect reported by Javorcik (2004a), but ten times more than what was found by Blalock & Gertler (2008). The whole procedure yields similar results when outliers are included (1.00) or when OLS is used (0.94).[5]

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 dependent variable in the FDI spillover regression (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, "worst practice" (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.

[5] 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.

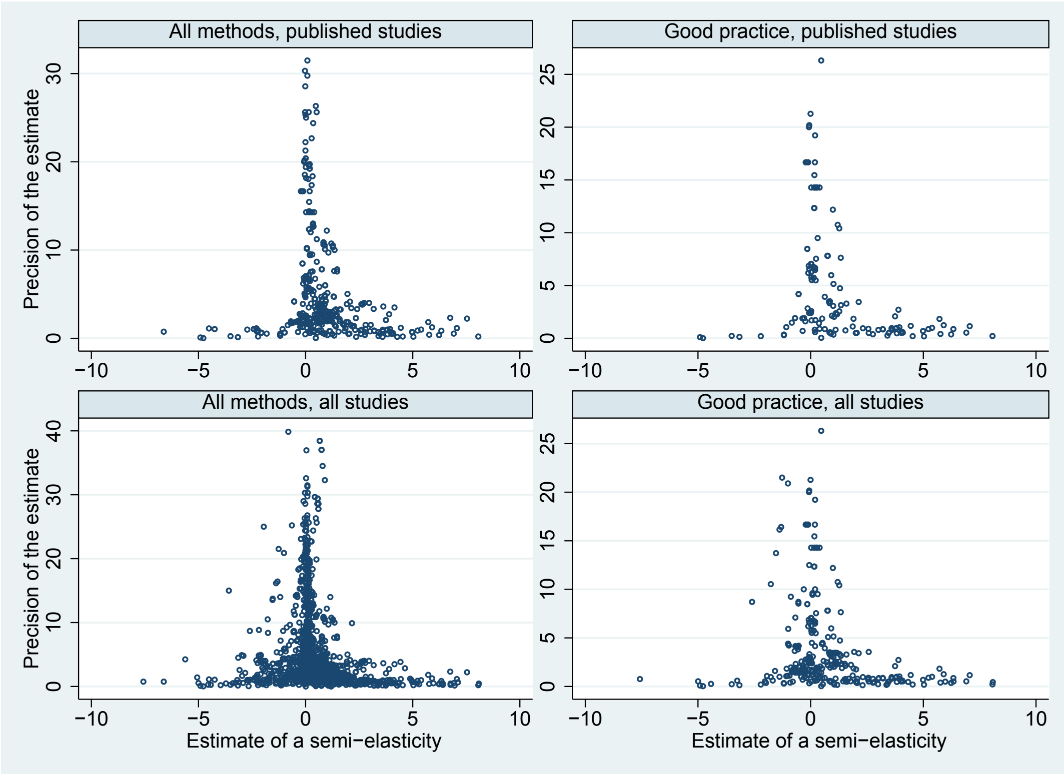
In Section 3.3 we found that estimates published in peer-reviewed journals are exaggerated because of publication selection. Now we have found that, in general, papers using better methods produce larger positive estimates of spillovers. The reader might wonder how the publishing filter works—are some 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 subsection 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 (Egger *et al.* 1997; 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 peer-reviewed journals is depicted in the top-left panel of Figure 3.1. It 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 (3.3), estimated earlier in Table 3.1 of Section 3.3, 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 horizontal axis. Thus both formal and visual tests suggest that only published results exhibit selection bias.

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 3.1. 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

Figure 3.1: Funnel plots show selection bias in published studies



Notes: "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.

input demand (we label theses aspects of methodology "good practice").[6] If journals select these estimates for their good practice and not for positive signs, the funnel plot would be symmetrical. But the new funnel for published estimates (the top-right panel of Figure 3.1) is no less asymmetrical than in the case when coefficients estimated by any method were considered (the top-left panel).

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 (3.3), the standard error, will be significant. When we add these interactions to our full model (3.6), 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 (3.3) 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 peer-reviewed journals

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

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).

3.5.2 Structural Heterogeneity

The meta-regression results for structural variables are reported in Table 3.3. 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 (the test is available on request). 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.

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 par-

Table 3.3: Structural heterogeneity in backward FDI spillovers

	Dependent variable: t-statistic of the estimate of spillover				
	1-ME	2-ME	3-ME	4-ME	5-WLS
Host-country characteristics					
Distance to source countries	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)
Trade 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)			
Foreign-firm characteristics					
Fully owned subsidiaries	-0.203***		-0.209***	-0.216***	-0.281***
	(0.0602)		(0.0603)	(0.0566)	(0.0946)
Partially owned subsidiaries	0.0203	0.0804	0.0227		
	(0.0561)	(0.0535)	(0.0564)		
Local-firm characteristics					
Service sectors	-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

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors in parentheses.

Notes: The table contains the results of regression (3.6). Method variables are included in all specifications and reported in Table 3.2. All explanatory variables are described in detail in Table 3.4. ME = mixed-effects multilevel model. WLS = weighted least squares with standard errors clustered at the study level.

tial ownership (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 3.3) 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 (3.6) reported in Table 3.2 and Table 3.3, 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.

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.

3.6 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 only a small effect on firms in customer sectors (forward spillovers) and no effect on firms 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 estimates 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.

References

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.

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.

DeLong, 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.

Egger, M., G. D. Smith, M. Scheider, & C. Minder (1997): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **316**: pp. 320–246.

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.

HAVRANEK, T. & Z. IRSOVA (2010): "Which Foreigners Are Worth Wooing? A Meta-Analysis of Vertical Spillovers from FDI." *Working Papers 2010/03*, Czech National Bank, Research Department.

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* **96(1)**: pp. 126–138.

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

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.

ROSENTHAL, R. (1979): "The 'file drawer problem' and tolerance for null results." *Psychological Bulletin* **86**: pp. 638–41.

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.

3.A Data Description

Table 3.4: Summary statistics of regression variables, backward spillovers

Variable	Description	Mean	Std. dev.
t-statistic	The t-statistic of the estimate of the spillover semi-elasticity.	0.803	4.997
$1/Se$	The precision of the estimate of the spillover semi-elasticity.	5.465	6.640
Method heterogeneity			
Data characteristics			
Cross-sectional data	=1 if cross-sectional data are used.	0.079	0.269
Aggregated data	=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
No. of firms	The logarithm of [(the number of observations used)/(time span)].	7.598	2.040
Average year of data	The average year of the data used (2000 as a base).	-1.053	3.798

Continued on next page

Summary statistics of regression variables, backward spillovers (continued)

Variable	Description	Mean	Std. dev.
Amadeus database	=1 if the Amadeus database by Bureau van Dijk Electronic Publishing is used.	0.223	0.416
Specification characteristics			
Forward spill. included	=1 if forward spillovers are included in the regression.	0.655	0.475
Horizontal spill. included	=1 if horizontal spillovers are included in the regression.	0.866	0.341
Foreign presence in employment	=1 if employment is the proxy for foreign presence.	0.142	0.349
Foreign presence in equity	=1 if equity is the proxy for foreign presence.	0.060	0.238
Foreign firms included	=1 if both domestic and foreign firms are included in the regression.	0.252	0.435
Control for absorption capacity	=1 if the specification controls for absorption capacity using technology gap or R&D spending.	0.070	0.256
Control for sector competition	=1 if the specification controls for sector competition.	0.272	0.445
Control for downstream demand	=1 if the specification controls for demand in downstream sectors.	0.075	0.263
Regional definition	=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 spillover	=1 if the coefficient represents lagged foreign presence.	0.127	0.334
More estimates	=1 if the coefficient is not the only estimate of backward spillovers in the regression.	0.459	0.499
Combination of estimates	=1 if the coefficient is a marginal effect computed using a combination of reported estimates.	0.072	0.259
Estimation characteristics			
One-step estimation	=1 if spillovers are estimated in one step using output, value added, or labor productivity as the dependent 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 effects	=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 effects	=1 if year fixed effects are included.	0.854	0.353
Sector fixed effects	=1 if sector fixed effects are included.	0.494	0.500
Estimated in 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

Continued on next page

Summary statistics of regression variables, backward spillovers (continued)

Variable	Description	Mean	Std. dev.
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 peer-reviewed journal.	0.288	0.453
Impact factor	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 co-author	=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 co-author	=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
Structural heterogeneity			
Host-country characteristics			
Distance to source countries	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
Trade 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
Foreign-firm characteristics			
Fully owned subsidiaries	=1 if only fully owned foreign investments are considered for linkages.	0.069	0.253
Partially owned subsidiaries	=1 if only investments with joint domestic and foreign ownership are considered for linkages.	0.070	0.256
Local-firm characteristics			
Service sectors	=1 if only firms from service sectors are included in the regression.	0.046	0.209

Notes: For host-country characteristics we select values from 1999, the median year of the data used in primary studies. The data for host-country characteristics are taken from World Development Indicators, Javorcik (2004b), and Walter G. Park's website.

Chapter 4

Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis

Abstract

The voluminous empirical research on horizontal productivity spillovers from foreign investors to domestic firms has yielded mixed results. In this paper, we collect 1,205 estimates of spillovers from the literature and examine which factors influence spillover magnitude. To identify the most important determinants of spillovers among 43 collected variables, we employ Bayesian model averaging. Our results suggest that horizontal spillovers are on average zero, but that their sign and magnitude depend systematically on the characteristics of the domestic economy and foreign investors. The most important determinants are the technology gap between domestic and foreign firms and the ownership structure in investment projects. Foreign investors who form joint ventures with domestic firms and who come from countries with a modest technology edge create the largest benefits for the domestic economy.

The paper was co-authored with Zuzana Irsova and published in World Development [2013, 42(February), pp. 1–15]. 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 us additional data, or explaining the details of their methodology, or both. We thank Martin Feldkircher, Maria Paula Fontoura, Petr Kral, Katerina Smidkova, and three anonymous referees for their helpful comments. The paper benefited from discussions at the Japan Economic Policy Association Conference, Nishinomiya (Japan), 2011, Meta-Analysis of Economics Research Colloquium, Perth (Australia), 2012, and Cities, Open Economies, and Public Policy Conference, Toronto (Canada), 2012. Tomas Havranek acknowledges support from the Czech Science Foundation (grant #P402/11/0948). Zuzana Irsova acknowledges support from the Grant Agency of Charles University (grant #76810). An online appendix with data, a Stata program, and a list of excluded studies is available at `meta-analysis.cz/bma`.

4.1 Introduction

With the rise in global flows of foreign direct investment (FDI) in recent decades, the policy competition for FDI among transition and developing countries has intensified. Consequently, many researchers have focused on the economic rationale of FDI incentives (Blomstrom & Kokko 2003, provide a review). The major hypothesis examined in the literature states that domestic firms may indirectly benefit from FDI: it is assumed that knowledge "spills over" from foreign investors or their acquired firms and helps domestic firms augment their productivity. (There is now solid evidence that FDI directly increases the productivity of the acquired firms; see Arnold & Javorcik 2009, for the case of Indonesia.) Nevertheless, the reported estimates of these "productivity spillovers" differ greatly in terms of both the statistical significance of the effect and its magnitude.

We build on the work of Crespo & Fontoura (2007), who review the literature on the determinants of FDI spillovers and thoroughly discuss the numerous factors that may cause the spillover effects to vary. Whereas the survey of Crespo & Fontoura (2007) is narrative, we examine spillover determinants using a quantitative method of literature surveys: meta-analysis. Meta-analysis was originally developed in medicine to aggregate costly clinical trials, and it has been widely used in economics to investigate the heterogeneity in reported results since the pioneering contribution of Stanley & Jarrell (1989). Recent applications of meta-analysis in economics include, among others, Card *et al.* (2010) on the evaluation of active labor market policies, Rusnak *et al.* (2013) on the effect of monetary policy on prices, and Babecky & Campos (2011) on the relation between structural reforms and economic growth in transition countries. In our case, meta-analysis makes use of evidence reported for many countries and different types of investment projects, enabling us to investigate hypotheses that are difficult to address in single-country case studies.

In the search for spillover determinants we focus on the characteristics of FDI host and source countries, foreign firms, and domestic firms in the host country. Moreover, we collect an extensive set of 34 control variables that may help explain the differences in reported findings, including the aspects of data used by primary studies on FDI spillovers, their methodology, publication quality, and author characteristics. To find the most important determinants we employ Bayesian model averaging. Bayesian model averaging is suitable for meta-analysis because of the inherent model uncertainty: while there is a consensus in the literature that some factors may mediate productivity spillovers (such as the technology gap, trade openness, or financial development), it is not clear which aspects of study design are important. Nevertheless, omission of these control variables may lead to biased estimates of coefficients for the main variables of interest. Bayesian model averaging allows us to concentrate on potential spillover determinants while taking all method variables into account.

In this paper we meta-analyze horizontal spillovers from FDI; that is, the effects of foreign investment on domestic firms in the same sector (as opposed to vertical spillovers, which denote the effect of FDI on domestic firms in supplier or customer sectors). To our knowledge, there have been two meta-analyses of horizontal spillovers: Görg & Strobl (2001) and Meyer & Sinani (2009). The meta-analysis by Görg & Strobl (2001) concentrates on the effect of study design on reported spillover coefficients and additionally tests for publication bias. Meyer & Sinani (2009) examine country heterogeneity in the estimates of spillovers. Compared with the earlier meta-analyses, we gather a more homogeneous sample of estimates so that we are able to examine the economic effect of spillovers. Moreover, we collect ten times more estimates of spillovers and investigate three times more factors that may explain spillover heterogeneity than Meyer & Sinani (2009), the larger of the earlier meta-analyses. We also revisit the issue of publication bias in the literature on horizontal spillovers from FDI employing modern meta-regression methods developed by Stanley (2005) and Stanley (2008).

The paper is structured as follows. Section 4.2 describes the properties of the data set of spillover estimates. Section 4.3 introduces the potential spillover determinants and the methodology of Bayesian model averaging. Section 4.4 presents estimation results. In Section 4.5 we test for publication bias in the literature. Section 4.6 provides a summary and policy implications.

4.2 The Data Set

Our data set comprises evidence on FDI spillovers from 45 countries reported in 52 empirical studies; the list of the studies used in the meta-analysis is available in the Appendix (Table 4.5). To increase the comparability of the estimates in our sample, we only include modern empirical studies that examine horizontal spillovers together with vertical spillovers in the same specification.[1] The first empirical studies on vertical spillovers appeared in the early 2000s, and thus we do not use any studies published before 2000—in contrast with the earlier meta-analyses on horizontal spillovers (Görg & Strobl 2001; Meyer & Sinani 2009), in which the pre-2000 studies account for most of the data. The pre-2000 studies were so heterogeneous in terms of methodology that it was not possible to compare directly the economic effects reported in the studies; instead, the earlier meta-analyses used measures of statistical significance, especially t-statistics. In the modern literature on FDI spillovers, most of the researchers examine how changes in the ratio of foreign presence affect the

[1]This restriction leads to an exclusion of some highly cited papers on FDI spillovers, such as Keller & Yeaple (2009).

productivity of domestic firms, and estimate a variant of the following model:

$$\ln \mathit{Productivity}_{ij} = e_0 \cdot \mathit{Horizontal}_j + e_0^b \cdot \mathit{Backward}_j + e_0^f \cdot \mathit{Forward}_j + \\ + \boldsymbol{\alpha} \cdot \mathit{Controls}_{ij} + u_{ij}, \quad (4.1)$$

where $\mathit{Productivity}_{ij}$ is a measure of the productivity of domestic firm i in sector j, $\mathit{Horizontal}_j$ is the ratio of foreign presence in sector j (the ratio ranges from 0 to 1), $\mathit{Backward}_j$ is the ratio of foreign presence in sectors that buy intermediate products from firms in sector j, and $\mathit{Forward}_j$ is the ratio of foreign presence in sectors that sell intermediate products to firms in sector j. Together, backward and forward spillovers form vertical spillovers. $\mathit{Controls}_{ij}$ denotes control variables included in the regression—for example, the degree of competition in sector j.

These "FDI spillover regressions" are usually run on firm-level panel data, but some primary studies still use cross-sectional data or data aggregated at the sectoral level (for example when examining countries for which better data are not available). Total factor productivity (TFP) is usually employed as the left-hand-side variable, but some studies use output, value added, or labor productivity. Foreign presence is most commonly measured as the share of output of foreign firms on the total output of all firms in the sector, but sometimes researchers use the share of employment or equity. In some specifications researchers control for other firm-level characteristics (such as, for instance, R&D spending) or sector-level characteristics (Herfindahl-Hirschman Index of competition among firms in the sector).

Some of the methods used in these papers are considered obsolete by the majority of researchers; for example, Görg & Strobl (2001) show that the use of cross-sectional instead of panel data often results in biased estimates of the spillover effect. Nevertheless, different researchers have different opinions on what constitutes the best practice in FDI spillover regressions (for example, whether the Olley-Pakes or Levinsohn-Petrin method should be used to compute TFP), and we thus follow the advice of Stanley (2001) and "better err on the side of inclusion" in our meta-analysis. If we excluded studies that do not correspond to a particular definition of best practice, we would greatly increase the subjectivity of our analysis and decrease the number of observations available. The general method of moments (GMM), for instance, is only used by a few studies in our data set. Therefore, we include all these studies in our analysis but control for the differences in data and methodology.

The regression coefficients from equation (4.1) represent the economic effect of FDI on the productivity of domestic firms. For instance, the coefficient for horizontal spillovers (e_0) expresses the percentage change in domestic productivity associated with an increase in foreign presence in the same sector of one percentage point, or, in other words, the semi-elasticity of domestic productivity with respect to foreign presence.

It is worth noting that the term "spillover" has become overused in the literature; the semi-elasticities in equation (4.1) may also capture effects other than knowledge externalities. As for horizontal effects, 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 thus captures the net effect of knowledge spillovers and competition on productivity. For the sake of simplicity, we follow the convention of calling productivity semi-elasticities "spillovers." The takeaway from this discussion is that even positive and economically significant estimates of semi-elasticities do not necessarily call for governments to subsidize FDI.

We searched for empirical studies on FDI spillovers in the EconLit, Scopus, and Google Scholar databases; and extracted results from all studies, published and unpublished, that report an estimate of e_0 with a measure of precision (standard error or t-statistic) and that control for vertical spillovers in the regression. In some cases we had to re-compute the estimates of spillovers so that they represented semi-elasticities—for example, if the regression was not estimated in the log-level form. For the computation we required sample means of the spillover variables, but this information is usually not reported in the studies. Therefore, we had to write to the authors of primary studies and ask for additional data or clarifications; the sample of the estimates available for meta-analysis would be much smaller without the help from the authors. The data, a Stata program, and a list of excluded studies with reasons for exclusion are available in the online appendix at `meta-analysis.cz/bma`.

Most studies report various estimates of spillovers: estimates for different countries, different types of investment projects, or estimates computed using a different methodology. To avoid arbitrary decisions on what the "best" estimate of each study could be, we extract all reported estimates. In sum, our data set contains 1,205 estimates of horizontal spillovers. We also codify 43 variables that may explain the differences among spillover estimates. For comparison, Nelson & Kennedy (2009) survey 140 meta-analyses conducted in economics since 1989; they find that an average meta-analysis uses 92 estimates and 12 explanatory variables. Therefore, our data set is large compared with that of conventional economics meta-analyses. (The largest meta-analysis in the sample of Nelson & Kennedy 2009, includes 1,592 estimates and employs 41 variables to explain heterogeneity.)

How big must the semi-elasticity be for spillovers to gain practical importance? Suppose, for instance, that e (an estimate of e_0) equals 0.1. Then, a ten-percentage-point increase in foreign presence is associated with an increase in domestic productivity in the same sector of 1%. This is not a great effect; nevertheless, Blalock & Gertler (2008) find similar magnitudes of spillover coefficients for Indonesia and note that such spillovers are important, because in the case of Indonesia there are large changes in foreign presence (large inflows of FDI): often in tens of percentage points within a few years.

The threshold determining the economic importance of FDI spillovers is of course subjective, and, unfortunately, economic importance is rarely discussed in primary studies. One of the exceptions is Haskel *et al.* (2007), who find the spillover semi-elasticity for the UK of about 0.05. They calculate the per-job value of spillovers implied by four well-known FDI projects in the UK and USA and compare them to per-job government subsidies granted to the investors. The Motorola plant established in Scotland in the early 1990s, for example, is predicted by the authors to generate a present-value spillover benefit of GBP 18,841 (compared to the per-job subsidy of GBP 14,356). In contrast, the Siemens plant established in 1996 in Tyneside, England, generated only GBP 3,430 in spillover benefits, much less than the per-job government subsidy of GBP 35,417. For the sake of simplicity, in this paper we consider spillover effects economically unimportant if they are lower than 0.1, irrespective of their statistical significance. On the other hand, the estimates that are statistically significant and larger than 0.1 we consider economically important.

Out of the 1,205 estimates that we collected, six are larger than 10 in absolute value. These observations are also more than three standard deviations away from the mean of all estimates. When we exclude these outliers, the mean hardly changes, but the standard deviation drops by two thirds. We thus continue in the analysis with a narrower set consisting of 1,199 estimates of horizontal spillovers, without the outliers. The simple mean of the remaining estimates is -0.002, not significantly different from zero at any conventional level. In meta-analysis it is common to weight the estimates by their precision (the inverse of the standard error); the procedure is commonly called fixed-effects meta-analysis (see, for example, Borenstein *et al.* 2009). In our case the fixed-effects meta-analysis provides a result broadly similar to the simple arithmetic average: 0.017, which is far from values at which the spillover effect could be considered important.

The fixed-effects meta-analysis assumes that there is no heterogeneity in the spillover effects across countries and estimation methods. In practice, however, heterogeneity is likely to be substantial. This is confirmed formally in our case by the Q test of heterogeneity, which is significant at any conventional level. An alternative method for estimating the average effect from the literature is called random-effects meta-analysis. Random-effects meta-analysis assumes that the true estimated effect is randomly distributed in the literature and, thus, can vary across countries and methods. Even with this approach the estimate of the average effect is close to zero and equals -0.011. These results, based on a broad sample of modern literature with a study of median age published only in 2008, corroborate the common impression that the evidence on horizontal spillovers is mixed (Görg & Greenaway 2004; Crespo & Fontoura 2007; Smeets 2008). In contrast, a recent meta-analysis of vertical

Figure 4.1: Country heterogeneity in the estimates of horizontal spill-overs for Europe

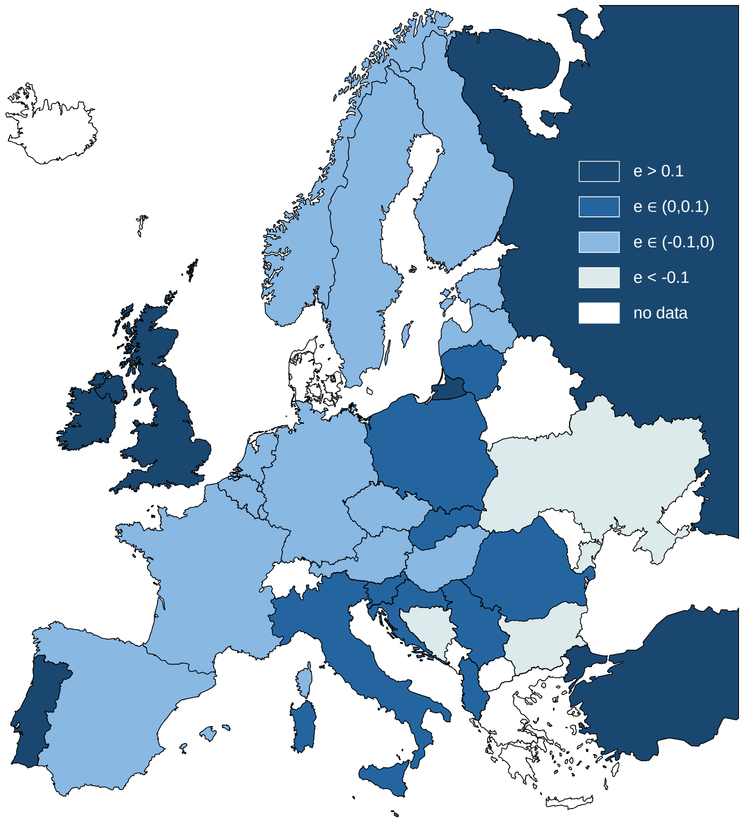
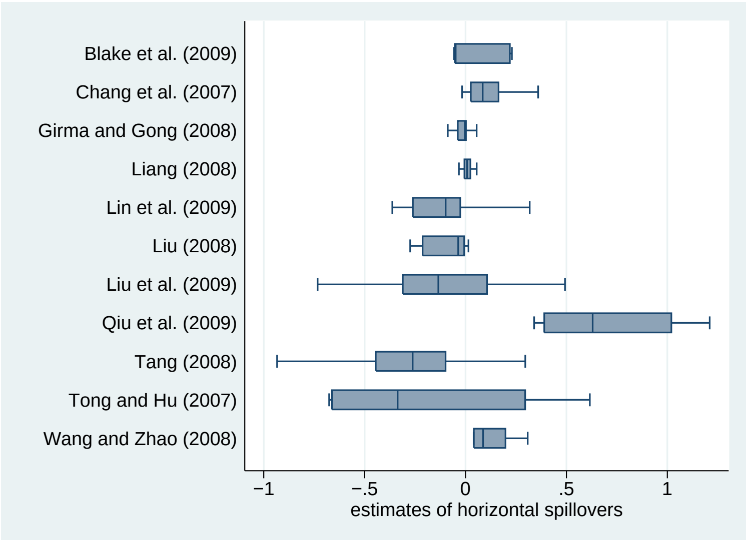


Figure 4.2: Method heterogeneity in the estimates of horizontal spill-overs for China



spillovers shows that they are on average important, in both statistical and economic terms (Havranek & Irsova 2011, Chapter 3 in the dissertation).

Horizontal spillovers are zero on average, but this does not have to mean that they are negligible in general. Perhaps host countries differ in their ability to benefit from FDI, as Lipsey & Sjöholm (2005) suggest; for some countries the effect may well be positive, whereas for others the negative effects of foreign competition on domestic firms (crowding out of the domestic market or draining of skilled labor force) may prevail. Since in the sample we have estimates of horizontal spillovers for almost all European countries, we illustrate in Figure 4.1 how spillovers differ from one European country to another. The values for individual countries are computed using random-effects meta-analysis and range from negative and economically important ($e < -0.1$) to positive and economically important ($e > 0.1$): horizontal spillovers are highly heterogeneous across countries. From the figure it is difficult to infer any clear relationship between the degree of economic development and the magnitude of spillovers. Clearly, the host-country characteristics are important for the benefits from FDI, but the relationship seems to have more than one dimension.

Another factor that may influence the reported spillover coefficients is the methodology used in the estimation. Though most researchers nowadays follow the general approach introduced earlier [equation (4.1)], they still have to make many method choices concerning data, specification, and estimation. Figure 4.2 shows how the results vary across studies with different methodologies for the country that is most frequently examined in the FDI spillover literature, China. The results are all over the place: from negative to positive, from negligible to economically significant. Therefore, if we want to discover what makes countries benefit from FDI, it is also important to control for the method choices employed in the studies.

4.3 Why Do Spillover Estimates Differ?

Building on the narrative surveys of the FDI spillover literature (Crespo & Fontoura 2007; Smeets 2008) and on the recent research concerning the factors that may determine the magnitude of horizontal spillovers, we compile a list of the potential spillover determinants that can be examined in a meta-analysis framework. Because spillovers are usually estimated for individual countries, and our database contains estimates of spillovers for 45 countries, it is convenient to express most of the determinants at the country level (Meyer & Sinani 2009, choose a similar approach).

On the other hand, in the meta-analysis framework it is not possible to investigate the influence of most microeconomic and regional factors on the magnitude of FDI spillovers. For example, Crespo *et al.* (2009) highlight the importance of the proximity between domestic and foreign firms and the existence of agglomeration externalities at the regional level. Since the authors of primary studies usually report spillover estimates for entire countries, meta-analysis unfortunately cannot shed further light on these important determinants. We can, however, still include a few important microeconomic factors: researchers often estimate separately productivity spillovers flowing from fully foreign-owned firms and from joint ventures of domestic and foreign firms, so we add a dummy for one of these cases and investigate whether this distinction is important for the reported magnitude of spillovers. Many researchers also estimate spillovers separately for the subsamples of manufacturing and services firms, and we can examine whether spillovers differ across these sectors.

As documented by Crespo & Fontoura (2007), the theory as well as empirical evidence gives mixed guidance on what the exact influence of the individual mediating factors on spillovers should be. Since the empirical results often vary from country to country, a meta-analysis for 45 countries could give us a more general picture. Here we provide a brief intuition for the inclusion of each of the nine potential determinants of horizontal spillovers:

Technology gap If the difference in the level of technology between domestic firms and foreign investors is too large, domestic firms are less likely to be able to imitate technology and adopt know-how brought by foreign investors. On the other hand, a small technology gap may mean that there is too little to learn from foreign investors (for more discussion on the role of the technology gap in mediating spillovers, see, for example, Blalock & Gertler 2009; Sawada 2010).

Similarity When the source country of FDI is closer to the host country in terms of culture, domestic firms are likely to adopt foreign technology more easily (as noted by Crespo & Fontoura 2007, p. 414). A common language or a similar legal system may represent an important mediating factor of horizontal spillovers. Moreover, a common language and historical colonial links are associated with migration patterns, and Javorcik *et al.* (2011) find that migration networks significantly affect FDI flows.

Trade openness In countries open to international trade, domestic firms are likely to have more experience with foreign firms and, hence, also with foreign technology. This may increase the domestic firms' absorptive capacity for spillovers (Lesher & Miroudot 2008), but it may also mean that there is less potential to learn because the firms are already exposed to foreign technology.

Financial development To benefit from the exposure to foreign technology, domestic firms should have access to financing so that they are able to implement the new technology in their production processes. In consequence, countries with a less developed financial system are likely to enjoy smaller horizontal spillovers (Alfaro *et al.* 2004).

Patent rights If the protection of intellectual property rights in the country is poor, the country is likely to attract relatively less sophisticated foreign investors (with only a modest technology edge over domestic firms). In addition, better protection of intellectual property rights makes it more difficult for domestic firms to copy technology from foreigners, and may lead to less positive horizontal spillovers (Smeets 2011).

Human capital With a more skilled labor force, domestic firms are likely to exhibit a greater capacity to absorb spillovers from foreign firms in the same sectors (Narula & Marin 2003).

FDI penetration If the country is already saturated with inward FDI, new foreign investment may have quite a small impact on domestic firms. In other words, the spillover semi-elasticity could be larger for an increase in foreign presence in the industry from 0 to 10% than, for example, from 50 to 60% (Gersl 2008).

Fully owned The degree of foreign ownership of investment projects is likely to matter for spillovers. Domestic firms can be expected to have harder access to the technology of fully foreign-owned affiliates than to the technology of joint ventures of foreign firms and other domestic firms (Abraham *et al.* 2010; Javorcik & Spatareanu 2008).

Service sectors Domestic firms in the service and manufacturing sectors may differ in their ability to benefit from foreign presence (Lesher & Miroudot 2008). For example, firms in service sectors are usually less export-intensive, and hence are likely to have less ex-ante experience with foreign firms. Less experience with foreign technology may lead to either a lower absorptive capacity or a higher potential to learn from FDI because of a larger technology gap.

Table 4.1: Description and summary statistics of regression variables

Variable	Description	Mean	Std. dev.
e	The estimate of the semi-elasticity for horizontal spillovers	-0.002	0.905
Potential spillover determinants			
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).	9.771	0.538
Similarity	The country's FDI-stock-weighted proxy for cultural and language similarity with respect to the source countries of FDI (=1 if countries share either a common language or a colonial link, =2 if both, =0 if neither).	0.628	0.616

Continued on next page

Description and summary statistics of regression variables (continued)

Variable	Description	Mean	Std. dev.
Trade openness	The trade openness of the country: (exports + imports)/GDP.	0.709	0.323
Financial dev.	The development of the financial system of the country: (domestic credit to private sector)/GDP.	0.600	0.432
Patent rights	The Ginarte-Park index of patent rights of the country.	3.052	0.793
Human capital	The tertiary school enrollment rate in the country.	0.269	0.186
FDI penetration	The ratio of inward FDI stock to GDP in the country.	0.267	0.186
Fully owned	=1 if only fully foreign-owned investments are considered for linkages.	0.078	0.269
Service sectors	=1 if only firms from service sectors are included in the regression.	0.062	0.241
Control Variables			
Data characteristics			
Cross-sectional	=1 if cross-sectional data are used.	0.088	0.284
Aggregated	=1 if sector-level data for productivity are used.	0.034	0.182
Time span	The number of years of the data used.	7.080	3.832
No. of firms	The logarithm of [(the number of observations used)/(time span)].	7.884	2.003
Average year	The average year of the data used (2000 as a base).	-1.120	3.953
Amadeus	=1 if the Amadeus database by Bureau van Dijk Electronic Publishing is used.	0.215	0.411
Specification characteristics			
Forward	=1 if forward vertical spillovers are included in the regression.	0.704	0.457
Employment	=1 if employment is the proxy for foreign presence.	0.139	0.346
Equity	=1 if equity is the proxy for foreign presence.	0.066	0.248
All firms	=1 if both domestic and foreign firms are included in the regression.	0.280	0.449
Absorption cap.	=1 if the specification controls for firms' absorption capacity using the technology gap or R&D spending.	0.057	0.231
Competition	=1 if the specification controls for sector competition.	0.297	0.457
Regional	=1 if vertical spillovers are measured using the ratio of foreign firms in the region as a proxy for foreign presence.	0.048	0.213
Lagged	=1 if the coefficient represents lagged foreign presence.	0.075	0.264
More estimates	=1 if the coefficient is not the only estimate of horizontal spillovers in the regression.	0.488	0.500
Combination	=1 if the coefficient is a marginal effect computed using a combination of reported estimates.	0.068	0.253
Estimation characteristics			
One step	=1 if spillovers are estimated in one step using output, value added, or labor productivity as the response variable.	0.461	0.499
Olley-Pakes	=1 if the Olley-Pakes method is used for the estimation of total factor productivity.	0.224	0.417
OLS	=1 if ordinary least squares (OLS) are used for the estimation of total factor productivity.	0.092	0.289
GMM	=1 if the system general-method-of-moments estimator is used for the estimation of spillovers.	0.028	0.164

Continued on next page

Description and summary statistics of regression variables (continued)

Variable	Description	Mean	Std. dev.
Random eff.	=1 if the random-effects estimator is used for the estimation of spillovers.	0.035	0.184
Pooled OLS	=1 if pooled OLS is used for the estimation of spillovers.	0.162	0.368
Year fixed	=1 if year fixed effects are included.	0.837	0.369
Sector fixed	=1 if sector fixed effects are included.	0.566	0.496
Differences	=1 if the regression is estimated in differences.	0.517	0.500
Translog	=1 if the translog production function is used.	0.048	0.213
Log-log	=1 if the coefficient is taken from a specification different from log-level.	0.018	0.134
Publication characteristics			
Published	=1 if the study was published in a peer-reviewed journal.	0.289	0.454
Impact	The recursive RePEc impact factor of the outlet. Collected in April 2010.	0.222	0.455
Study citations	The logarithm of [(Google Scholar citations of the study)/(age of the study) + 1]. Collected in April 2010.	1.180	1.026
Native co-author	=1 if at least one co-author is native to the investigated country.	0.714	0.452
Author citations	The logarithm of (the number of RePEc citations of the most-cited co-author + 1). Collected in April 2010.	2.956	2.508
US-based	=1 if at least one co-author is affiliated with a US-based institution (usually highly ranked institutions in our sample).	0.292	0.455
Publication date	The year and month of publication (January 2000 as a base).	7.827	1.418

Source of the data: UNCTAD, World Development Indicators, www.cepii.org, OECD, and Walter Park's website. For country-level variables we use values for 1999, the median year of the data used in the primary studies.

The first seven potential spillover determinants are computed at the country level. Out of these seven variables, *Technology gap* and *Similarity* show average bilateral values with respect to the source countries of FDI. The remaining two variables, *Fully owned* and *Service sectors*, are dummy variables, and their values are determined by the manner of estimation of spillovers in the primary studies (researchers often estimate separately the effects of fully foreign-owned investment projects and joint ventures and also examine separately the effects on domestic firms in manufacturing and in service sectors). Details on the construction of all variables and their summary statistics are provided in Table 4.1. The table also lists all 34 control variables that we use in our estimation: the characteristics of the data, specification, estimation, and publication of the primary studies on horizontal spillovers from FDI.

Our intention is to examine how the nine potential determinants influence the reported estimates of horizontal spillovers. As documented by the intuition outlined on the previous pages, all of the potential determinants may play a role in explaining spillover heterogeneity. On the other hand, it is far from clear which control variables from our extensive set should be included in the regression, or what signs their regression coefficients should have. A regression with all 43 explanatory variables would certainly contain many redundant control variables and would unnecessarily inflate the standard errors. The general model, a so-called "meta-regression," can be described in the following way:[2]

$$e_k = a + \boldsymbol{\beta} \cdot \mathit{Determinants}_k + \boldsymbol{\gamma} \cdot \mathit{Controls}_k + \epsilon_k, \quad k = 1, \ldots, 1199, \tag{4.2}$$

where e is an estimate of horizontal spillovers, *Determinants* denotes the nine potential spillover determinants, which should be included in the regression, and *Controls* denotes control variables, some of which may be included in the regression. This is a typical example of model uncertainty that can be addressed by a method called Bayesian model averaging (BMA; for example, Fernandez *et al.* 2001a; Sala-i-Martin *et al.* 2004; Ciccone & Jarocinski 2010; Moral-Benito 2012). BMA has been applied in meta-analysis, for instance, by Moeltner & Woodward (2009).

BMA estimates many models comprising the possible subsets of explanatory variables and constructs a weighted average over these models. In a way, BMA can be thought of as a meta-analysis of meta-analyses, because it aggregates many possible meta-regression models. The weights in this methodology are the so-called *posterior model probabilities*. Simply put, posterior model probability can be imagined as a measure of the fit of the model, analogous to information criteria or adjusted R-squared: the models that fit the data best get the highest posterior model probability, and vice versa. Next, for each explanatory variable we can compute the *posterior inclusion probability*, which represents the sum of the posterior model probabilities of all models that contain this particular variable. In other words, the posterior inclusion probability expresses how likely it is that the variable should be included in the "true" regression. Finally, for each explanatory variable we are able to extract the *posterior coefficient distribution* across all the regressions. From the posterior coefficient distribution we can infer the posterior mean (analogous to the estimate of the regression coefficient in a standard regression) and the posterior standard deviation (analogous to the standard error of the regression coefficient in a standard regression).

Because we have to consider 43 explanatory variables, it is not technically feasible to enumerate all 2^{43} of their possible combinations; on a standard personal computer this would take several years. In such cases, Markov chain Monte Carlo methods are used to go through the most important models (those with high posterior model probabilities). For the computation we use the `bms` package in R (Feldkircher & Zeugner 2009), which employs the Metropolis-Hastings algorithm. Following Fernandez *et al.*

[2]Ideally, nonlinear functions and interactions of the variables should be included as well. Nevertheless, with so many potential explanatory variables this would greatly increase the complexity of the model and introduce problems with multicollinearity.

(2001b), we run the estimation with 200 million iterations, which ensures a good degree of convergence. We apply conservative priors on both the regression coefficients and the model size to let the data speak. More details on the BMA procedure employed in this paper are available in Section 4.B; more details on BMA in general can be found, for example, in Feldkircher & Zeugner (2009).

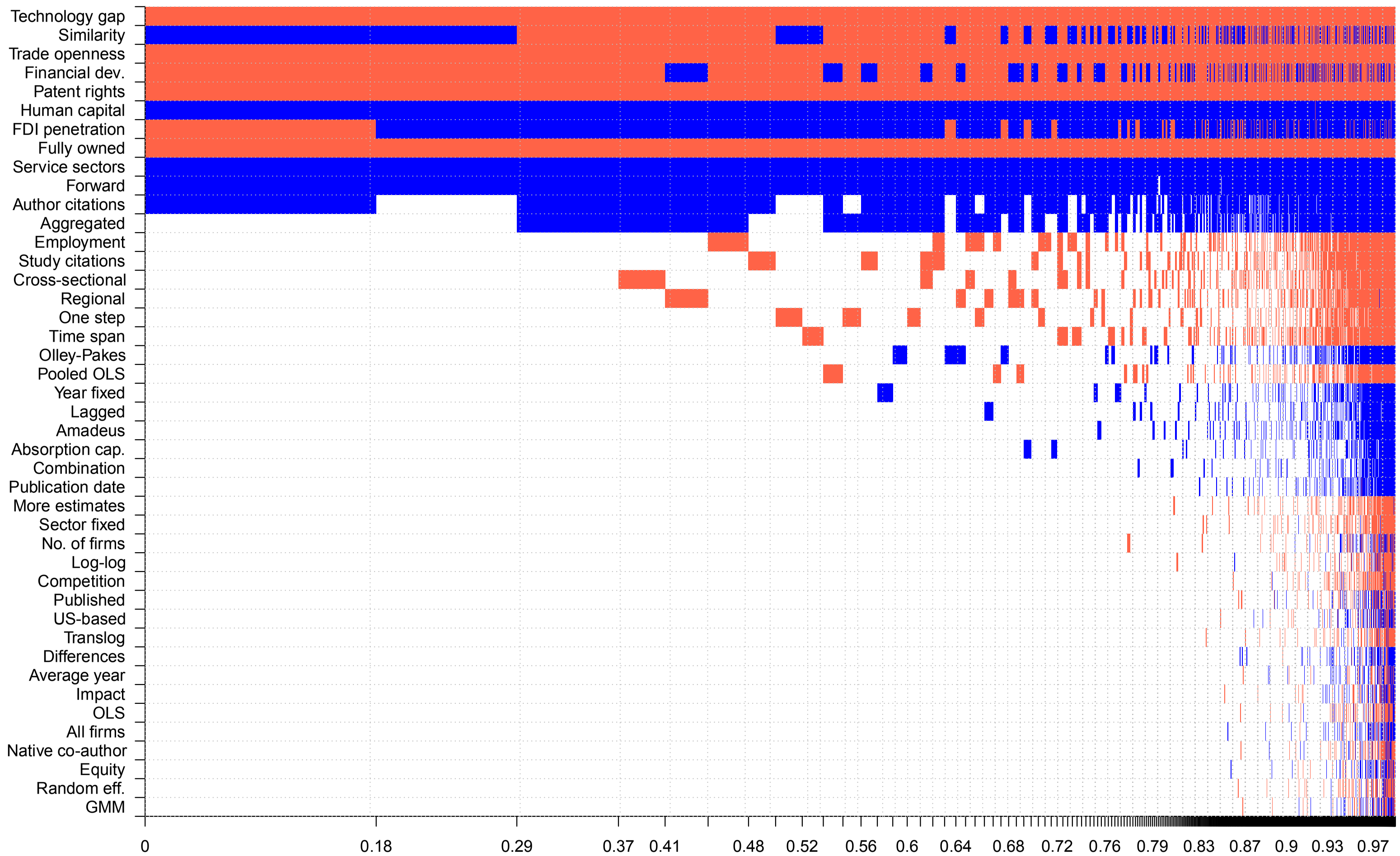
4.4 Meta-Regression Results

A graphical representation of the results of the BMA estimation is depicted in Figure 4.3. Columns denote individual models; these models include the explanatory variables for which the corresponding cells are not blank. Blue color (darker in grayscale) of the cell means that the variable is included in the model and that the estimated sign of the regression coefficient is positive. Red color (lighter in grayscale) means that the variable is included and that the estimated sign is negative. On the horizontal axis the figure depicts the posterior model probabilities: the wider the column, the better the fit of the model. For example, the best model, the first one from the left, includes only two control variables—*Forward* (a dummy variable that equals one if the primary study controls for both backward and forward vertical spillovers when estimating horizontal spillovers) and *Author citations* (the number of citations of the most frequently cited co-author of the primary study). The posterior probability of the best model, however, is only 18%, and we have to take a look at the rest of the model mass as well.

The posterior inclusion probability, computed as the sum of the posterior model probabilities for the models that include the corresponding variable, also exceeds 50% for variable *Aggregated* (a dummy variable that equals one if the data in the primary study are aggregated at the sector level; that is, if firm-level data are not available). A few other control variables seem to be important in many models, but especially in the worse ones to the right. From Figure 4.3 we can infer how stable the regression coefficients are for potential spillover determinants. The sign of the coefficient is consistently negative for *Technology gap*, *Trade openness*, *Patent rights*, and *Fully owned*. On the other hand, the figure shows mixed results for *Similarity*, *Financial development*, and *FDI penetration*: the coefficients for these variables are unstable and depend on which control variables are included in the regression. Finally, the sign seems to be clearly positive for variables *Human capital* and *Service sectors*.

Table 4.2 reports numerical details on the results of the BMA estimation. Because for one country a few variables are not available, we can only use 1,195 out of all 1,199 spillover estimates in the BMA. Most control variables have a posterior inclusion probability lower than 0.1; therefore they do not seem to be important. A few control variables have a posterior inclusion probability between 0.1 and 0.5, which suggests that they may play a role in influencing the magnitude of the reported spillover

Figure 4.3: Bayesian model averaging, model inclusion



Notes: Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. Blue color (darker in grayscale) = the variable is included and the estimated sign is positive. Red color (lighter in grayscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model. The horizontal axis measures the cumulative posterior model probabilities.

Table 4.2: Explaining the differences in the estimates of horizontal spillovers

Response variable:	Bayesian model averaging			Frequentist check (OLS)		
Estimate of spillovers	Post. mean	Post. std. dev.	PIP	Coef.	Std. er.	p-value
Potential spillover determinants						
Technology gap	**-0.294**	**0.088**	**1.000**	**-0.260**	**0.145**	**0.080**
Similarity	-0.006	0.097	1.000	-0.086	0.108	0.430
Trade openness	**-0.246**	**0.138**	**1.000**	**-0.367**	**0.176**	**0.044**
Financial dev.	-0.083	0.162	1.000	0.020	0.178	0.909
Patent rights	**-0.144**	**0.076**	**1.000**	**-0.183**	**0.119**	**0.131**
Human capital	**0.437**	**0.316**	**1.000**	**0.710**	**0.499**	**0.162**
FDI penetration	0.085	0.232	1.000	0.218	0.276	0.435
Fully owned	**-0.144**	**0.103**	**1.000**	**-0.104**	**0.057**	**0.077**
Service sectors	0.092	0.118	1.000	0.150	0.144	0.303
Data characteristics						
Cross-sectional	-0.043	0.123	0.124	-0.290	0.091	0.003
Aggregated	0.352	0.378	0.524	0.965	0.210	3.E-07
Time span	-0.003	0.010	0.093			
No. of firms	-1.E-04	0.003	0.007			
Average year	9.E-06	0.001	0.003			
Amadeus	0.005	0.034	0.026			
Specification characteristics						
Forward	0.313	0.068	0.997	0.281	0.074	0.001
Employment	-0.036	0.093	0.146	-0.178	0.104	0.094
Equity	8.E-05	0.007	0.003			
All firms	7.E-05	0.004	0.003			
Absorption cap.	0.005	0.041	0.022			
Competition	-4.E-04	0.008	0.005			
Regional	-0.065	0.194	0.115	-0.309	0.278	0.274
Lagged	0.008	0.050	0.029			
More estimates	-0.001	0.009	0.008			
Combination	0.002	0.024	0.012			
Estimation characteristics						
One step	-0.017	0.058	0.095			
Olley-Pakes	0.012	0.049	0.068			
OLS	-9.E-05	0.007	0.003			
GMM	3.E-06	0.009	0.003			
Random eff.	-1.E-04	0.008	0.003			
Pooled OLS	-0.014	0.057	0.062			
Year fixed	0.008	0.041	0.040			
Sector fixed	-0.001	0.010	0.007			
Differences	2.E-04	0.005	0.004			
Translog	-4.E-04	0.011	0.004			
Log-log	-0.001	0.031	0.006			
Publication characteristics						
Published	3.E-07	0.008	0.005			
Impact	4.E-06	0.004	0.003			
Study citations	-0.012	0.033	0.127	-0.093	0.075	0.222
Native co-author	-5.E-05	0.005	0.003			
Author citations	0.042	0.029	0.745	0.088	0.037	0.024
US-based	8.E-05	0.007	0.004			
Publication date	4.E-04	0.005	0.010			
Observations	1,195			1,195		

Notes: For variables in bold the BMA estimates that the posterior means of the regression coefficients are larger than the corresponding posterior standard deviations. PIP = posterior inclusion probability. Potential spillover determinants are always included. In the frequentist check we only include control variables with PIP > 0.1. Standard errors in the frequentist check are clustered at the country level.

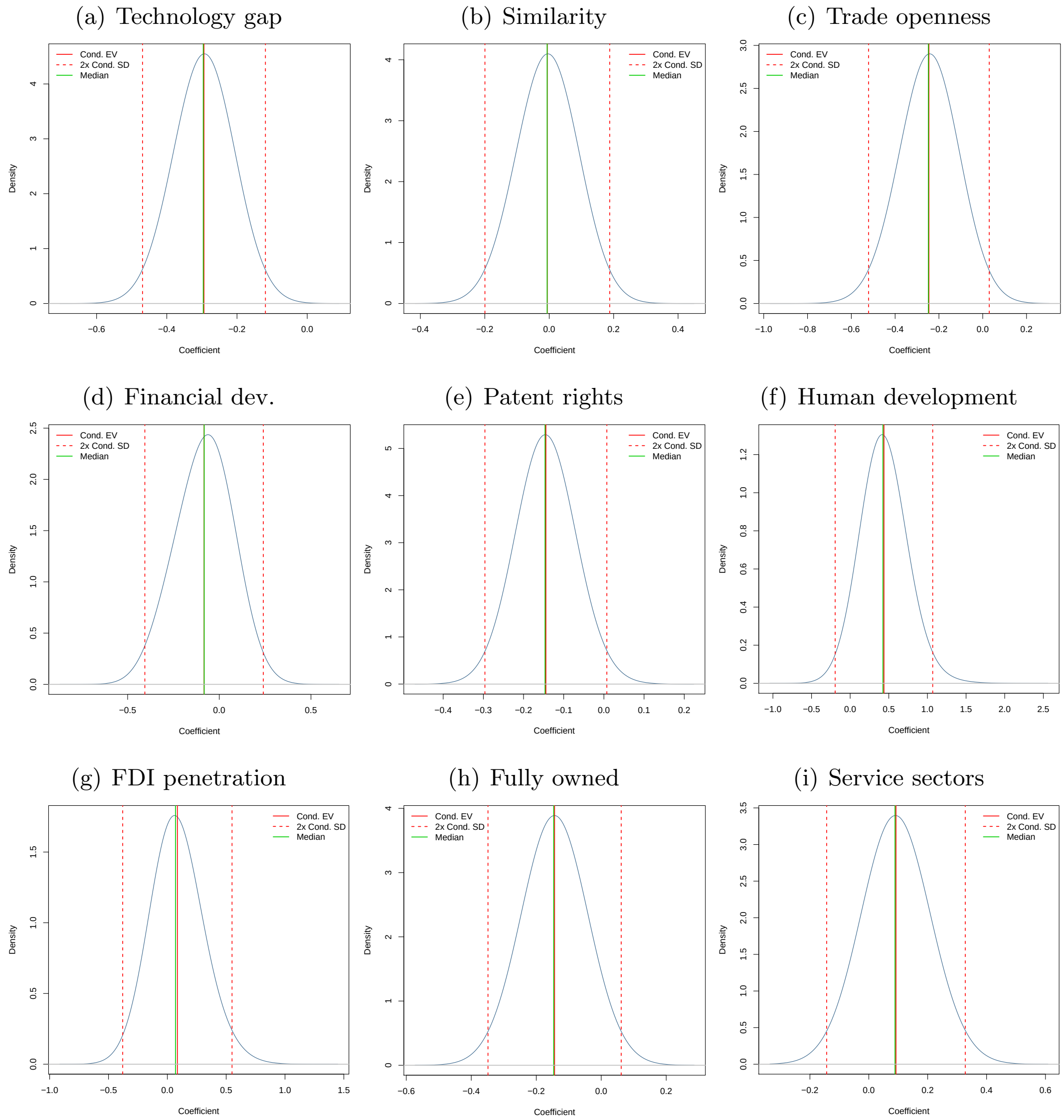
coefficients. The variables with such a moderate posterior inclusion probability are the following: *Cross-sectional* (a dummy variable that equals one if cross-sectional data instead of panel data are used in the primary study), *Employment* (a dummy variable that equals one if the share of foreign firms in the sector's employment is used as the proxy for foreign presence), *Regional* (a dummy variable that equals one if vertical spillovers in the regression are measured using the ratio of foreign firms in the region), and *Study citations* (the number of citations of the study divided by the age of the study).

As a "frequentist" check of the BMA estimation, we run a simple OLS regression with all potential spillover determinants and the control variables with posterior inclusion probabilities higher than 0.1 (that is, the control variables that the BMA estimation finds to be relatively important). In other words, using OLS we run one of the many models shown in Figure 4.3. Because we are interested in the potential spillover determinants, most of them being defined at the country level, we use country-level clustered standard errors in the regression (the potential spillover determinants would be a bit more significant if study-level clustering was used instead). The results are reported in the last three columns of Table 4.2 and are broadly in line with the BMA estimation in terms of the predicted coefficient values and their standard errors. The potential spillover determinants that seem to be important based on the BMA estimation are typeset in bold; we highlight variables for which the posterior mean of the regression coefficient exceeds the posterior standard deviation. Apart from variables with clearly unstable signs as was seen from Figure 4.3, additionally the variable *Service sectors* does not seem to be important; its regression coefficient is also highly insignificant in the frequentist check.

Table 4.2 only shows the summary statistics of the posterior distribution of the regression coefficients; for a closer look at the posterior distributions for potential spillover determinants, we need to advance to Figure 4.4. The solid line in the graphs denotes the posterior mean of the regression coefficients, which was already reported in Table 4.2. The dotted lines denote coefficient values that are two posterior standard deviations away from the posterior mean; if zero lies outside these intervals, the interpretation of the result is broadly similar to statistical significance at the 5% level in the frequentist case.

Figure 4.4 suggests that the coefficient for *Technology gap* is negative with a high probability. Therefore, our results suggest that a high technology gap between domestic firms and foreign investors results in smaller horizontal spillovers. In contrast, the coefficient for *Similarity* is almost precisely zero: it seems that neither a common language nor a historical colonial link between the host and source country from FDI helps increase the benefits of FDI. (The results would hold even if we considered only a common language or only a colonial link in the definition of *Similarity*.) Next, we find that the coefficient for *Trade openness* is robustly negative, which is consistent

Figure 4.4: Posterior coefficient distributions for potential spillover determinants



Notes: The figure depicts the densities of the regression parameters for the corresponding spillover determinant encountered in different regressions (with subsets of all control variables on the right-hand side). For example, the regression coefficient for *Technology gap* is negative in almost all models, irrespective of the control variables included. The most common value of the coefficient is approximately −0.3. On the other hand, the coefficient for *Similarity* is negative in one half of the models and positive in the other half, depending on which control variables are included. The most common value is 0.

with the hypothesis that companies with ex-ante experience from international trade have little to learn from foreign investors coming to their country. The degree of *Financial development* does not seem to be important for horizontal spillovers. In contrast, the degree of protection of intellectual property rights matters: the coefficient for *Patent rights* is robustly negative. With stronger protection of intellectual property, the host country can expect less horizontal spillovers from incoming FDI since it becomes more difficult for domestic firms to copy technology from foreign firms.

The estimated coefficient corresponding to *Human development* is positive, which suggests that to benefit from FDI, host countries need a skilled labor force; skilled employees increase the absorptive capacity of domestic firms. *FDI penetration* does not seem to matter for the size of horizontal spillovers. This result is consistent with the implicit hypothesis behind most regressions in primary studies: the researchers usually assume that the effect of FDI on domestic firms is linear, or, in other words, that the spillover semi-elasticity is constant for different values of foreign presence. The coefficient for *Fully owned* is negative, which means that joint ventures are more likely to bring positive spillovers for domestic firms than fully foreign-owned investment projects. Finally, the mean of the coefficient for *Service sectors* is positive, but for many models negative coefficients are reported.

The results discussed on the previous pages give us some idea about the direction with which the various mediating factors influence horizontal spillovers from FDI. For practical purposes, however, we need to determine the economic importance of the individual spillover determinants. In Table 4.3 we consider two measures of economic importance. First, we examine how the BMA estimation would predict the horizontal spillovers to change if the value of the spillover determinant increased from the minimum value in our sample to the maximum value. The results suggest that *Technology gap* is by far the most important determinant: extreme changes in the difference between the technological level of domestic firms and foreign investors can increase or decrease the spillover coefficient by 1.321. If we consider values above 0.1 to be economically important, as discussed in Section 4.2, a value of 1.321 represents a huge difference.

Nevertheless, such large changes in spillover determinants are not realistic, and in the next column of Table 4.3 we thus report the changes in spillovers associated with a one-standard-deviation increase in the spillover determinants. Even according to this measure the most important determinant is *Technology gap*, but the predicted effect on the spillover coefficient is much lower than in the previous case: 0.158. Other important determinants are *Patent rights* (the one-standard-deviation effect equals 0.115), *Human capital* (0.081), and *Trade openness* (0.079). Note that a one-standard-deviation effect is not suitable for dummy variables such as *Fully owned*, because the value of *Fully owned* is either 0 or 1. The spillover effect of fully

Table 4.3: The economic significance of potential spillover determinants

Variable	Maximum effect	Std. dev. effect
Technology gap	-1.321	-0.158
Similarity	-0.012	-0.004
Trade openness	-0.341	-0.079
Financial dev.	-0.097	-0.036
Patent rights	-0.478	-0.115
Human capital	0.282	0.081
FDI penetration	0.102	0.016
Fully owned	-0.144	-0.039
Service sectors	0.092	0.022

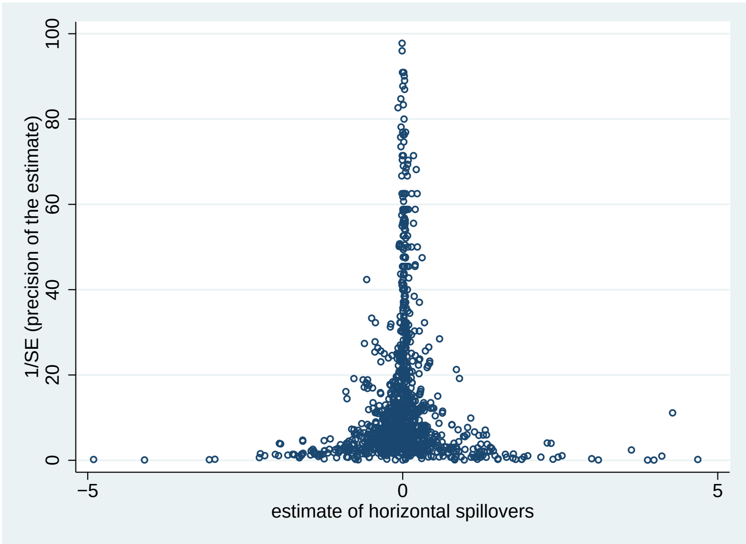
Notes: The table depicts the predicted effects of increases in the variables on the spillover estimates based on BMA. Maximum effect = an increase from sample minimum to sample maximum. Std. dev. effect = a one-standard-deviation increase.

foreign-owned investment projects is 0.144 smaller compared with the case when all investments are considered. Therefore, if the host country encourages foreign investment projects involving joint ventures with a somewhat smaller technology advantage with respect to domestic firms, it may increase the average spillovers by $0.144 + 0.158 = 0.302$, an economically significant value.

4.5 Publication Bias

An important concern in meta-analysis is publication selection bias (see, for example, Stanley 2001; 2005; Havranek 2010; Havranek *et al.* 2012): some estimates of spillovers may be more likely to be selected for publication than others. The presence of publication selection would probably not affect the analysis of spillover determinants in the previous two sections, but it could seriously bias our estimate of the average spillover reported in Section 4.2. Publication selection in the spillover literature has two potential sources. First, researchers may treat statistically significant results more favorably, as seems to be the case in many areas of empirical economics (see, for example, the surveys of DeLong & Lang 1992; Card & Krueger 1995). Second, researchers may prefer a particular direction of the estimate of spillovers. Some researchers may be tempted to report "good news" (positive estimates) for developing countries in contrast to skeptical results. Moreover, until the 1990s there was a relatively strong consensus in the literature that horizontal spillovers were truly positive, so researchers could use this intuition as a specification check. Indeed, publication selection bias was found in the first meta-analysis of horizontal spillovers by Görg & Strobl (2001).

Figure 4.5: Funnel plot



The presence of publication bias is usually tested both graphically and formally. The graphical test uses the so-called funnel plot (Egger *et al.* 1997; Stanley & Doucouliagos 2010), a scatter plot of the estimates of spillovers (on the horizontal axis) against their precision (the inverse of the standard error; on the vertical axis). In the absence of publication bias the funnel plot is symmetrical: the most precise estimates are close to the true spillover, while the imprecise estimates are dispersed widely. In consequence, the scatter plot should resemble an inverted funnel. On the other hand, if some estimates of spillovers are discarded because of their unintuitive sign, the funnel will become asymmetrical. If insignificant estimates are not reported, the funnel will become hollow (results yielding small coefficients with large standard errors will be discarded).

The funnel plot for our sample of horizontal spillovers is reported in Figure 4.5. The funnel seems to be full and symmetrical, although the left portion of the funnel might be a little heavier than the right one. In any case, most funnels reported in economics meta-analyses show much stronger asymmetry than what we see in Figure 4.5 (Stanley 2008; Stanley & Doucouliagos 2010). Because the interpretation of the funnel plot is rather subjective, more formal methods are needed to assess the presence of publication bias in the spillover literature.

The most commonly employed test for publication bias reformulates the funnel plot as a regression relationship: the funnel asymmetry test. If we switch the axes in the funnel plot and invert the values on the new horizontal axis, we get a relation between the estimate of spillovers and its standard error. In the absence of publication bias, the estimated size of the coefficient should not be correlated with its standard error (Card & Krueger 1995; Egger *et al.* 1997). If, however, some estimates are selected for publication because of their significance or an intuitive sign, the relation will become significant. The following regression, used already by Card & Krueger (1995), formalizes the idea:

$$e_k = e_0 + \beta_0 \cdot Se(e_k) + u_k, \quad k = 1, \ldots, 1199, \tag{4.3}$$

where e denotes the estimate of spillovers, e_0 is the average underlying spillover, $Se(e)$ is the standard error of e, and β_0 measures the magnitude of publication bias. Because specification (4.3) is likely heteroscedastic (the explanatory variable is a sample estimate of the standard deviation of the response variable), in practice it is usually estimated by weighted least squares to ensure efficiency (Stanley 2005; 2008). Since we have many estimates from different studies, we add study fixed effects and cluster the standard errors at the study level (country-level clustering would yield similar results).

Table 4.4: Test of publication bias

	Study fixed effects			Study and country fixed effects		
Response variable: e	Coef.	Std. er.	p-value	Coef.	Std. er.	p-value
Constant	0.021	0.015	0.150	0.021	0.015	0.183
Se (publication bias)	-0.325	0.262	0.220	-0.284	0.305	0.357
Observations	1,199			1,199		

Notes: Standard errors are clustered at the study level. Estimated by weighted least squares with the precision (the inverse of standard error) taken as the weight.

The results reported in Table 4.4 confirm the intuition based on the funnel plot: the coefficients for publication bias are small and insignificant. In a quantitative survey of economic meta-analyses, Doucouliagos & Stanley (2013) state that values of the coefficient for publication bias in the funnel asymmetry test are important if they are statistically significant and larger than one in absolute value. With coefficients for publication bias reported in Table 4.4 to be around -0.3, we can conclude that publication selection in the spillover literature is negligible. The result contrasts with the findings of Görg & Strobl (2001), who find strong publication selection. Nevertheless, in this meta-analysis we use the estimates of horizontal spillovers published after 2000, and in the following decade the focus of many studies shifted to vertical spillovers, so that the selection pressure could have moved to those estimates. Indeed, Havranek & Irsova (2012) show that publication bias in the literature on vertical spillovers is strong.

4.6 Conclusion

In a large meta-analysis of horizontal spillovers from FDI estimated for 45 countries, we examine which factors determine the magnitude of spillovers. On average, horizontal spillovers are negligible, but the estimates are distributed unevenly across countries and estimation methods. Building on the previous literature we investigate nine potential spillover determinants that capture the characteristics of the FDI source countries, host countries, domestic firms, and investment projects. Additionally we assemble a list of 34 aspects of methodology that may affect the estimates of spillovers. Using Bayesian model averaging we investigate the importance of individual spillover determinants and control for the aspects of methodology. We also test for possible publication selection bias.

Our results suggest that the nationality of foreign investors is important: when the technology gap of domestic firms with respect to foreign investors is too large, horizontal spillovers are small. Moreover, spillovers are likely to be smaller with higher trade openness and better protection of intellectual property rights in the host country. On the other hand, higher levels of human capital in the host country are associated with larger spillovers. Finally, investment projects in the form of joint ventures with domestic firms bring more positive spillovers than fully foreign-owned projects. We found no evidence of publication bias in the literature on horizontal spillovers.

Productivity spillovers from FDI are often cited as the most important reason for promoting inward FDI (Blomstrom & Kokko 2003). Therefore, if horizontal spillovers were the only effect of inward FDI on the domestic economy, our meta-analysis would suggest that promotion of FDI brings no benefits on average. Although we found that changes in some country characteristics can be expected to have positive effects on FDI spillovers, some of these changes are also likely to have serious detrimental side effects. For example, changing the degree of protection of intellectual property or the degree of trade openness, difficult as it is, would certainly affect many other aspects of the economy, the volume of FDI attracted among them, and is thus not suitable for policy purposes.

Nevertheless, there are tools that may, with caution, be used to increase the benefits from FDI without obvious side effects. If the country already spends money on promoting foreign investment, it could benefit from focusing the resources on investors who are most likely to generate positive spillovers. Our meta-analysis indicates that these are investors coming from countries with a modest technology edge who are willing to form joint ventures with domestic firms. Such investment projects would help foster not only horizontal, but also vertical spillovers, as documented by the meta-analysis of Havranek & Irsova (2011, Chapter 3 of the dissertation).

References

Abraham, F., J. Konings, & V. Slootmaekers (2010): "FDI spillovers in the Chinese manufacturing sector." *The Economics of Transition* **18(1)**: pp. 143–182.

Alfaro, L., A. Chanda, S. Kalemli-Ozcan, & S. Sayek (2004): "FDI and economic growth: The role of local financial markets." *Journal of International Economics* **64(1)**: pp. 89–112.

Arnold, M. J. & B. S. Javorcik (2009): "Gifted kids or pushy parents? Foreign direct investment and plant productivity in Indonesia." *Journal of International Economics* **79(1)**: pp. 42–53.

Atallah Murra, S. (2006): "Revaluando la transmision de spillovers de la IED: Un estudio de productividad para Colombia." *Revista Desarrollo y Sociedad* **57(1)**: pp. 163–213.

Babecky, J. & N. F. Campos (2011): "Does reform work? An econometric survey of the reform–growth puzzle." *Journal of Comparative Economics* **39(2)**: pp. 140–158.

Barrios, S., H. Görg, & E. Strobl (2009): "Spillovers Through Backward Linkages from Multinationals: Measurement Matters!" *IZA Discussion Papers 4477*, Institute for the Study of Labor.

Békés, G., J. Kleinert, & F. Toubal (2009): "Spillovers from Multinationals to Heterogeneous Domestic Firms: Evidence from Hungary." *The World Economy* **32(10)**: pp. 1408–1433.

Blake, A., Z. Deng, & R. Falvey (2009): "How does the productivity of foreign direct investment spill over to local firms in Chinese manufacturing?" *Journal of Chinese Economic and Business Studies* **7(2)**: pp. 183–197.

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.

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

Blalock, G. & D. H. Simon (2009): "Do all firms benefit equally from downstream FDI? The moderating effect of local suppliers' capabilities on productivity gains." *Journal of International Business Studies* **40(7)**: pp. 1095–1112.

Blomstrom, M. & A. Kokko (2003): "The economics of foreign direct investment incentives." *NBER Working Papers 9489*, National Bureau of Economic Research.

Blyde, J., M. Kugler, & E. Stein (2004): "Exporting vs. Outsourcing by MNC Subsidiaries: Which Determines FDI Spillovers?" *Discussion Paper Series In Economics And Econometrics 0411*, Economics Division, School of Social Sciences, University of Southampton.

Borenstein, M., H. R. Rothstein, L. V. Hedges, & J. P. Higgins (2009): *Introduction to Meta-Analysis.* London: Wiley.

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)**: pp. 514–526.

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.

Chang, S. J., J. Chung, & D. Xu (2007): "FDI and Technology Spillovers in China." *CEI Working Paper Series 2007-7*, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University.

Ciccone, A. & M. Jarocinski (2010): "Determinants of Economic Growth: Will Data Tell?" *American Economic Journal: Macroeconomics* **2(4)**: pp. 222–46.

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

Crespo, N., M. P. Fontoura, & I. Proença (2009): "FDI spillovers at regional level: Evidence from Portugal." *Papers in Regional Science* **88(3)**: pp. 591–607.

Damijan, J. P., M. Knell, B. Majcen, & M. Rojec (2003): "Technology Transfer through FDI in Top-10 Transition Countries: How Important are Direct Effects, Horizontal and Vertical Spillovers?" *William Davidson Institute Working Papers Series 549*, William Davidson Institute at the University of Michigan.

Damijan, J. P., M. Rojec, B. Majcen, & M. Knell (2008): "Impact of Firm Heterogeneity on Direct and Spillover Effects of FDI: Micro Evidence from Ten Transition Countries." *LICOS Discussion Papers 21808*, LICOS—Centre for Institutions and Economic Performance, K. U. Leuven.

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

DOUCOULIAGOS, H. & T. D. STANLEY (2013): "Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

EGGER, M., G. D. SMITH, M. SCHEIDER, & C. MINDER (1997): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **316**: pp. 320–246.

FELDKIRCHER, M. & S. ZEUGNER (2009): "Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging." *IMF Working Papers 09/202*, International Monetary Fund.

FERNANDEZ, C., E. LEY, & M. F. J. STEEL (2001a): "Benchmark priors for Bayesian model averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

FERNANDEZ, C., E. LEY, & M. F. J. STEEL (2001b): "Model uncertainty in cross-country growth regressions." *Journal of Applied Econometrics* **16(5)**: pp. 563–576.

GERSL, A. (2008): "Productivity, Export Performance, and Financing of the Czech Corporate Sector: The Effects of Foreign Direct Investment." *Czech Journal of Economics and Finance* **58(05-06)**: pp. 232–247.

GERSL, A., I. RUBENE, & T. ZUMER (2007): "Foreign Direct Investment and Productivity Spillovers: Updated Evidence from Central and Eastern Europe." *Working Paper 2007/08*, Czech National Bank.

GIRMA, S. & Y. GONG (2008): "FDI, Linkages and the Efficiency of State-Owned Enterprises in China." *Journal of Development Studies* **44(5)**: pp. 728–749.

GIRMA, S., H. GÖRG, & M. PISU (2008): "Exporting, linkages and productivity spillovers from foreign direct investment." *Canadian Journal of Economics* **41(1)**: pp. 320–340.

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.

GORODNICHENKO, Y., J. SVEJNAR, & K. TERRELL (2007): "When Does FDI Have Positive Spillovers? Evidence from 17 Emerging Market Economies." *CEPR Discussion Papers 6546*, Centre for Economic Policy Research.

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.

Hagemejer, J. & M. Kolasa (2008): "Internationalization and economic performance of enterprises: Evidence from firm-level data." *National Bank of Poland Working Papers 51*, National Bank of Poland.

Halpern, L. & B. Muraközy (2007): "Does distance matter in spillover?" *Economics of Transition* **15**: pp. 781–805.

Haskel, J. E., S. C. Pereira, & M. J. Slaughter (2007): "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?" *The Review of Economics and Statistics* **89(3)**: pp. 482–496.

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

Havranek, T. & Z. Irsova (2011): "Estimating vertical spillovers from FDI: Why results vary and what the true effect is." *Journal of International Economics* **85(2)**: pp. 234–244.

Havranek, T. & Z. Irsova (2012): "Survey Article: Publication Bias in the Literature on Foreign Direct Investment Spillovers." *Journal of Development Studies* **48(10)**: pp. 1375–1396.

Havranek, T., Z. Irsova, & K. Janda (2012): "Demand for Gasoline is More Price-Inelastic than Commonly Thought." *Energy Economics* **34(1)**: p. 201–207.

Jabbour, L. & J. L. Mucchielli (2007): "Technology transfer through vertical linkages: The case of the Spanish manufacturing industry." *Journal of Applied Economics* **10(1)**: pp. 115–136.

Javorcik, B. S. (2004): "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. & 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 (2011): "Does it matter where you come from? Vertical spillovers from foreign direct investment and the origin of investors." *Journal of Development Economics* **96(1)**: pp. 126–138.

Javorcik, B. S., a. Özden, M. Spatareanu, & C. Neagu (2011): "Migrant networks and foreign direct investment." *Journal of Development Economics* **94(2)**: pp. 231–241.

Jordaan, J. A. (2008): "Regional foreign participation and externalities: New empirical evidence from Mexican regions." *Environment and Planning A* **40(12)**: pp. 2948–2969.

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)**: pp. 821–831.

Kolasa, M. (2008): "How does FDI inflow affect productivity of domestic firms? The role of horizontal and vertical spillovers, absorptive capacity and competition." *Journal of International Trade & Economic Development* **17(1)**: pp. 155–173.

Le, Q. H. & R. Pomfret (2008): "Technology Spillovers from Foreign Direct Investment in Vietnam: Horizontal or Vertical Spillovers?" *Working Paper 85*, Vietnam Development Forum.

Lesher, M. & S. Miroudot (2008): "FDI Spillovers and their Interrelationships with Trade." *OECD Trade Policy Working Papers 80*, OECD, Trade Directorate.

Ley, E. & M. F. Steel (2009): "On the effect of prior assumptions in Bayesian model averaging with applications to growth regressions." *Journal of Applied Econometrics* **24(4)**: pp. 651–674.

Liang, F. H. (2008): "Does Foreign Direct Investment Improve the Productivity of Domestic Firms? Technology Spillovers, Industry Linkages, and Firm Capabilities." *Working paper*, Haas School of Business, University of California, Berkeley.

Lileeva, A. (2006): "Global Links: The Benefits to Domestically-controlled Plants from Inward Direct Investment—The Role of Vertical Linkages." *The Canadian Economy in Transition Research Paper 010*, Statistics Canada, Economic Analysis Division.

Lin, P., Z. Liu, & Y. Zhang (2009): "Do Chinese domestic firms benefit from FDI inflow? Evidence of horizontal and vertical spillovers." *China Economic Review* **20(4)**: pp. 677–691.

Lipsey, R. E. & F. Sjöholm (2005): "The Impact of Inward FDI on Host Countries: Why Such Different Answers?" In T. H. Moran, E. Graham, & M. Blomström (editors), "Does Foreign Direct Investment Promote Development?", pp. 23–43. Institute for International Economics and Center for Global Development.

Liu, X., C. Wang, & Y. Wei (2009): "Do local manufacturing firms benefit from transactional linkages with multinational enterprises in China?" *Journal of International Business Studies* **40(7)**: pp. 1113–1130.

Liu, Z. (2008): "Foreign direct investment and technology spillovers: Theory and evidence." *Journal of Development Economics* **85(1-2)**: pp. 176–193.

Managi, S. & S. M. Bwalya (2010): "Foreign direct investment and technology spillovers in sub-Saharan Africa." *Applied Economics Letters* **17(6)**: pp. 605–608.

Merlevede, B. & K. Schoors (2005): "Conditional Spillovers from FDI Within and Between Sectors: Evidence from Romania." *Working papers*, Örebro University, Sweden.

Merlevede, B. & K. Schoors (2007): "FDI and the Consequences: Towards more complete capture of spillover effects." *William Davidson Institute Working Papers Series 886*, William Davidson Institute at the University of Michigan.

Merlevede, B. & K. Schoors (2009): "Openness, competition, technology and FDI spillovers: Evidence from Romania." *Working Paper 42*, Forum for Research in Empirical International Trade.

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.

Moeltner, K. & R. Woodward (2009): "Meta-Functional Benefit Transfer for Wetland Valuation: Making the Most of Small Samples." *Environmental & Resource Economics* **42(1)**: pp. 89–108.

Moral-Benito, E. (2012): "Determinants of Economic Growth: A Bayesian Panel Data Approach." *The Review of Economics and Statistics* **94(2)**: pp. 566–579.

Narula, R. & A. Marin (2003): "FDI spillovers, absorptive capacities and human capital development: Evidence from Argentina." *Research Memoranda 018*, Maastricht: MERIT, Maastricht Economic Research Institute on Innovation and Technology.

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.

Nguyen, A. N., N. Thang, L. D. Trung, N. Q. Pham, C. D. Nguyen, & N. D. Nguyen (2008a): "Foreign Direct Investment in Vietnam: Is There Any Evidence Of Technological Spillover Effects." *Working Papers 18*, Development and Policies Research Center, Vietnam.

Nguyen, C. D., G. Simpson, D. Saal, A. N. Nguyen, & N. Q. Pham (2008b): "FDI Horizontal and Vertical Effects on Local Firm Technical Efficiency." *Working Papers 17*, Development and Policies Research Center, Vietnam.

Qiu, B., S. Yang, P. Xin, & B. Kirkulak (2009): "FDI technology spillover and the productivity growth of China's manufacturing sector." *Frontiers of Economics in China* **4(2)**: pp. 209–227.

Reganati, F. & E. Sica (2007): "Horizontal and Vertical Spillovers from FDI: Evidence from Panel Data for the Italian Manufacturing Sector." *Journal of Business Economics and Management* **8(4)**: pp. 259–266.

Rusnak, M., T. Havranek, & R. Horvath (2013): "How to Solve the Price Puzzle? A Meta-Analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

Sala-i-Martin, X., G. Doppelhofer, & R. Miller (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *American Economic Review* **94(4)**: pp. 813–835.

Sasidharan, S. & A. Ramanathan (2007): "Foreign Direct Investment and Spillovers: Evidence from Indian Manufacturing." *International Journal of Trade and Global Markets* **1(1)**: pp. 5–22.

Sawada, N. (2010): "Technology Gap Matters on Spillover." *Review of Development Economics* **14(1)**: pp. 103–120.

Schoors, K. & B. van der Tol (2002): "Foreign direct investment spillovers within and between sectors: Evidence from Hungarian data." *Working Papers 02/157*, Ghent University, Faculty of Economics and Business Administration, Belgium.

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

Smeets, R. (2011): "Knowledge diffusion from FDI and Intellectual Property Rights." *CPB Discussion Paper 168*, CPB Netherlands Bureau for Economic Policy Analysis.

Stancik, J. (2007): "Horizontal and Vertical FDI Spillovers: Recent Evidence from the Czech Republic." *CERGE-EI Working Papers 340*, CERGE-EI, Prague.

Stancik, J. (2009): "FDI Spillovers in the Czech Republic: Takeovers vs. Greenfields." *Economic Papers 369*, The European Comission.

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. & S. B. JARRELL (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161–70.

TANG, H. (2008): *Essays on international trade and investment.* Ph.D. thesis, Massachusetts Institute of Technology.

TAYMAZ, E. & K. YŁLMAZ (2008): "Foreign Direct Investment and Productivity Spillovers: Identifying Linkages Through Product-Based Measures." *Working papers*, Koc University.

TONG, S. Y. & Y. HU (2007): "Productivity Spillovers from FDI: Detrimental or Beneficial? A Study of Chinese Manufacturing." In B. M. FLEISHER, H. LI, & S. SONG (editors), "Market Development In China: Spillovers, Growth and Inequality," pp. 190–208. Cheltenham, England: Edward Elgar.

VACEK, P. (2007): "Productivity Spillovers from Foreign Direct Investment: Industry-level Analysis." In "Essays on International Productivity Spillovers," Ph.D. thesis, Cornell University.

WANG, C. & Z. ZHAO (2008): "Horizontal and vertical spillover effects of foreign direct investment in Chinese manufacturing." *Journal of Chinese Economic and Foreign Trade Studies* **1(1)**: pp. 8–20.

YUDAEVA, K., K. KONSTANTIN, N. MELENTIEVA, & N. PONOMAREVA (2003): "Does Foreign Ownership Matter? Russian Experience." *Economics of Transition* **11(3)**: pp. 383–409.

ZAJC KEJZAR, K. & A. KUMAR (2006): "Inward Foreign Direct Investment and Industrial Restructuring: Micro Evidence—The Slovenian Firms' Growth Model." *Proceedings of Rijeka School of Economics: Journal of Economics and Business* **24(2)**: pp. 185–210.

4.A Studies Used in the Meta-Analysis

Table 4.5: List of primary studies

Atallah Murra (2006)	Hagemejer & Kolasa (2008)	Merlevede & Schoors (2007)
Barrios *et al.* (2009)	Halpern & Muraközy (2007)	Merlevede & Schoors (2009)
Békés *et al.* (2009)	Jabbour & Mucchielli (2007)	Nguyen *et al.* (2008a)
Blake *et al.* (2009)	Javorcik (2004)	Nguyen *et al.* (2008b)
Blalock & Gertler (2008)	Javorcik & Spatareanu (2011)	Qiu *et al.* (2009)
Blalock & Simon (2009)	Javorcik & Spatareanu (2008)	Reganati & Sica (2007)
Blyde *et al.* (2004)	Jordaan (2008)	Sasidharan & Ramanathan (2007)
Bwalya (2006)	Kolasa (2008)	Schoors & van der Tol (2002)
Chang *et al.* (2007)	Le & Pomfret (2008)	Stancik (2007)
Crespo *et al.* (2009)	Lesher & Miroudot (2008)	Stancik (2009)
Damijan *et al.* (2003)	Liang (2008)	Tang (2008)
Damijan *et al.* (2008)	Lileeva (2006)	Taymaz & Yłlmaz (2008)
Gersl (2008)	Lin *et al.* (2009)	Tong & Hu (2007)
Gersl *et al.* (2007)	Liu (2008)	Vacek (2007)
Girma & Gong (2008)	Liu *et al.* (2009)	Wang & Zhao (2008)
Girma *et al.* (2008)	Managi & Bwalya (2010)	Yudaeva *et al.* (2003)
Girma & Wakelin (2007)	Merlevede & Schoors (2005)	Zajc Kejzar & Kumar (2006)
Gorodnichenko *et al.* (2007)		

Notes: Both published and unpublished studies are included if they control for vertical spillovers. We use all comparable estimates reported in the studies. The search for primary studies was terminated on March 31, 2010. A list of excluded studies, with reasons for exclusion, is available in the online appendix.

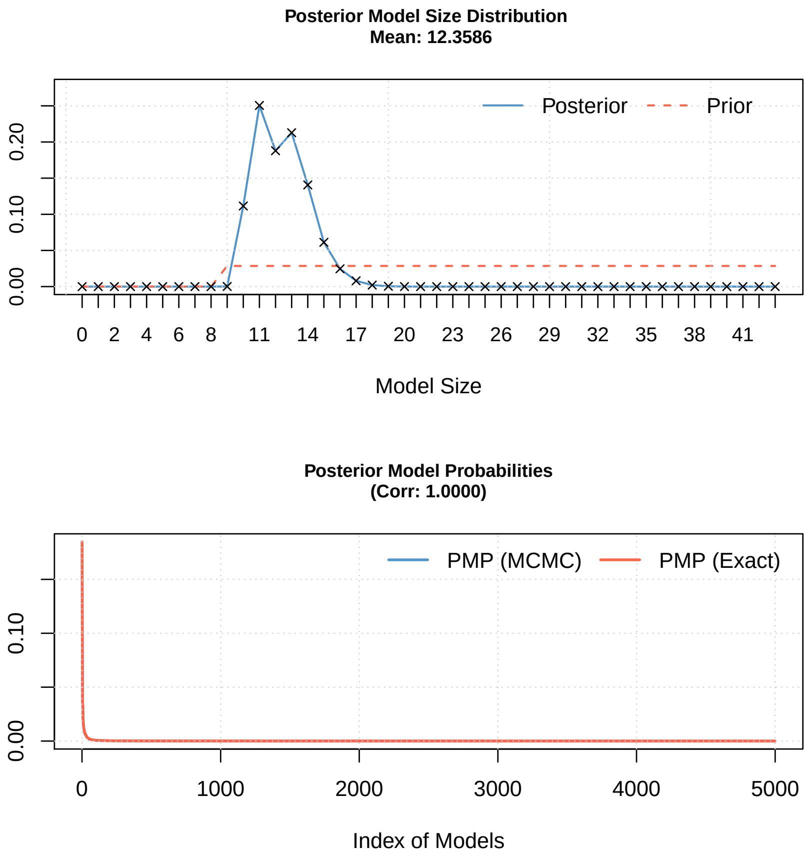
4.B BMA Diagnostics

Table 4.6: Summary of BMA estimation

Mean no. regressors	*Draws*	*Burn-ins*	*Time*
12.359	$2 \cdot 10^8$	$1 \cdot 10^8$	13.679 hours
No. models visited	*Modelspace*	*Visited*	*Topmodels*
$23,619,112$	$8.8 \cdot 10^{12}$	0.00027%	99%
Corr PMP	*No. Obs.*	*Model Prior*	*g-Prior*
1.0000	1195	random	BRIC
Shrinkage-Stats			
Av= 0.9995			

Notes: The "random" model prior refers to the beta-binomial prior advocated by Ley & Steel (2009): prior model probabilities are the same for all possible models; in other words, we do not a priori prefer any particular model size. We set the Zellner's g prior following Fernandez *et al.* (2001a).

Figure 4.6: Model size and convergence



Response to Reviewers

for the *Dissertation Defense*

June 3, 2013

on manuscript
Meta-Analysis in International Economics
by Tomas Havranek

I thank the reviewers for insightful comments on the pre-defense version of my dissertation. Since the reviewers suggest that the dissertation can be submitted without major changes, I only make minor adjustments in the text.

Response to Comments from Julia Wörz

I am grateful to Prof. Wörz for her kind words on my dissertation. Prof. Wörz has one comment concerning possible revision and suggests that I rethink the ordering of the chapters. In line with her recommendation the revised version of the dissertation starts with the chapter on the Rose Effect, which addresses publication bias.

Response to Comments from Tom Stanley

I thank Prof. Stanley for his kind assessment of my research and would like to respond to some of his comments (the comments are typeset in italics in this letter).

- *I do not entirely agree with Mr. Havranek's choice of explanatory (or moderator) variables. It is rather standard to use dummy variables to represent obvious omitted variables in the original study because omitted variables bias has been shown to often be the principal source of research heterogeneity. I do not care for the use of various citation statistics as moderator variables. At some point, coding all of these tangential factors can cause the acceptance of a spurious MRA, inadvertently. But then there is room for a healthy difference of opinion about what variables should be included as moderator variables, and the research record itself should ultimately be what decides.*

  I prefer to include variables describing citation statistics or journal prestige to control for quality characteristics not captured by method variables (moreover, I have found that journal referees who are not meta-analysts like this practice). I agree that this practice can lead to multicolinearity problems. To alleviate this concern, I typically drop variables that are jointly insignificant at some level, which leaves me with a more parsimonious model. In any case Prof. Stanley is right that in this paper I use too many explanatory variables—in my more recent meta-analyses I ignore many of these "publication characteristics." As Prof. Stanley notes in his textbook on meta-analysis, it is more appropriate to use fixed effects to control for study quality; but in this paper I was interested in study-level variables, so fixed effects were not an option.

- *The column in Tables 2.2 and 2.3 labeled 'OLS' should be called 'WLS,' because using the t-value as the dependent variable with precision gives a WLS regression.*

  This is a good point; I have corrected the label.

- *pp. 23-24: I agree with Mr. Havranek's use of "best practice" methods to make a 'prediction' of spillover elasticity using his multiple MRA. This is what we do (Doucouliagos & Stanley, 2009), but few follow. However, I sense that there might be a sizable mistake here somewhere. The best practice estimates are reported on p. 25 to be around 1, but the corrected estimates are .17-.31 as reported in Table 2.1. This difference raises questions. Was the publication bias term driven to zero in the "best practice" prediction? In my research, these two approaches usually give the same approximate estimate, and I would regard it as a robustness check to do so. It is my opinion that the magnitude of spillover elasticity may be overestimated. Especially when we further consider the appropriate values of some of the publication variables.*

  The "best practice" estimate in the paper is conditional on the publication bias term being close to zero. I agree that the result is a little surprising (publication bias drives the estimates up, but misspecifications drive them down), and this was one of the major issues raised by the referees of the Journal of International Economics. I tried to explain it in the last section of the paper (now page 55 in the dissertation). I can imagine a situation in which misspecifications and publication bias have opposite effects on the results. For example, if one part of the literature uses instrumental variables to address endogeneity and the other ignores endogeneity and uses OLS, the other will yield more precise, but biased estimates. If the bias is downward and researchers preferentially report positive estimates, I would get a result similar to what is presented in this paper.

- *Why isn't the 'best practice' approach used to predict the trade effect of the euro from the multiple MRA as was done in Chapter 2? In that way, various values could have been used to explore the robustness of this small trade effect from the euro.*

  My intention in the paper on the Rose Effect was to focus on publication bias, so I did not collect much information on methodology (which would have taken much more time). I have some multiple MRA analysis in the paper, but I do not think it is enough for a proper best practice analysis.

- *Why is panel data selected to influence only publication selection—a K-variable in our multiple MRA approach (Stanley* et al.*, 2008; Doucouliagos & Stanley, 2009)? I would have thought the use of panel data to have a genuine effect on the estimated trade effect by better controlling for omitted-variable bias.*

  This is again a good point; the dummy variable for the use of panel data should have been included among the variables influencing the true effect as well. The hypothesis used in the paper was that studies employing panel data are more likely to get published.

References

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.

Stanley, T., C. 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.