Meta-Regression Analysis in Economics: Lessons Learned[1]

Zuzana Havránková
Institute of Economic Studies, Charles University, Prague

1. Introduction

In this paper I provide a non-technical summary of six papers on macroeconomics, international economics, and energy economics. All the papers are tied together by the use of meta-regression analysis, which is essential for the derivation of robust policy-relevant conclusions from often conflicting results presented in the empirical literature. I use meta-analysis to quantitatively synthesize the reported research results on various topics; the technique has been known and applied in medical science for many decades, but only spread to economics in the 1990s. For each topic I correct the literature for publication selection bias and filter out the effect of various misspecifications present in some primary studies.

My results can be summarized as follows: (1) The elasticity of intertemporal substitution in consumption, a key input to all dynamic models in finance and macroeconomics, varies significantly across countries. The differences can be explained by the level of stock market participation, when countries with higher participation exhibit larger values of the elasticity; the mean reported elasticity is 0.5. (2) The effect of borders on international trade, which most authors find to be surprisingly large, can be explained away by innovations in methodology introduced in the last decade. When these innovations are taken into account jointly, the border effect disappears for developed countries, and is relatively small for developing countries. (3) When all published estimates of the effect of foreign investment on local firms in the same industries are considered and corrected for publication bias, the literature indicates a zero effect. (4) Publication bias is present also in the literature estimating the effect of foreign investment on local firms in different industries, but here the corrected effect is positive and large. (5) The mean reported price elasticity of gasoline demand is exaggerated twofold due to publication bias. (6) Finally, I also find that publication bias distorts the literature estimating the social cost of carbon emissions, because researchers tend to preferentially report large estimates.

The differences between individual chapters of the dissertation, discussed in detail in the next section, demonstrate the development of meta-regression methods in recent years. I discuss the evolution of meta-analysis methods and my take on best-practice approaches in the field, focusing on selection of primary studies, tests of publication bias, selection of variables, robustness checks, and the role of judgement in meta-analysis.

2. Empirical Applications of Meta-Analysis in Economics

2.1 Do Borders Really Slash Trade? A Meta-Analysis

First chapter of the dissertation thesis focuses on the effect of international borders on trade; the paper is co-authored with Tomas Havranek (Havranek & Irsova 2015). The paper is currently at the revise-and-resubmit stage at the IMF Economic Review. The finding that international borders significantly reduce trade, first reported by McCallum (1995), has become a stylized fact of international economics. A high ratio of trade within national borders to trade across borders, after controlling for other trade determinants, implies large unobserved border barriers, an implausibly high elasticity of substitution between domestic and foreign goods, or both. Dozens of researchers have attempted to shrink McCallum's original estimates but the average in the literature is still close to those estimated by McCallum (1995): regions are likely to trade with foreign regions about fifteen times less than with regions in the same country. Using previously reported results we construct a large synthetic study that estimates the border effect, but corrects for potential publication or misspecification biases.

We collect 32 aspects of studies, such as the characteristics of data, estimation, inclusion of control variables, number of citations, and information on the publication outlet. To explore how these characteristics affect the estimates of the border effect, we employ Bayesian model averaging (Raftery et al. 1997). The method addresses model uncertainty inherent in meta-analysis by estimating regressions comprising the potential subsets of the study aspects and weighting them by statistics related to the goodness of fit.

Our results suggest that many innovations in estimating the gravity equation systematically affect the reported border effect: for example, the use of disaggregated data, consistent measure of within and between-country distance, data on actual road or sea distance instead of the great-circle distance, control for multilateral resistance, and the use of the Poisson pseudo-maximum likelihood estimator. When we put these influences together and compute a general equilibrium impact of borders conditional on best practice

1 *This article is a summary of a dissertation entitled "Six Essays on Meta-Regression Analysis" defended at the Institute of Economic Studies, Charles University in Prague. The dissertation was granted first place in the NBS Governor's Award for an outstanding dissertation thesis or diploma thesis in the area of monetary economics, macroeconomics, or financial economics, and is available freely at http://ies.fsv.cuni.cz/default/file/download/id/27497. I use my maiden name, Iršová, in all of my publications.*

methodology, we find that borders reduce international trade by only 28% worldwide. The border effects differ significantly across regions – we obtain large estimates for developing and transition countries, but estimates close to zero for most OECD countries. Finally, we find little evidence of publication bias in the literature: researchers do not preferentially report positive or statistically significant estimates of the border effect.

2.2 A Meta-Analysis of Intra-Industry FDI Spillovers

The second chapter of the dissertation presents a meta-analysis of the empirical literature on horizontal productivity spillovers from foreign direct investment (FDI) – the effect of foreign presence on the productivity of domestic competitors. It is a joint work with Tomas Havranek published in the Czech Journal of Economics and Finance (Havranek & Irsova 2010). We gather a sample of 97 models from 67 studies published either in academic journals or as working papers. Using the vote-counting method, we find that the spillover effect does not seem to be statistically significant in general; employing the approach of Djankov & Murrell (2002), on the other hand, we find evidence that positive spillovers from FDI might exist. Nevertheless, this is not the case of the narrower sample of studies that were published in the best economics journals or that use panel and firm-level data – their combined t statistics is insignificant almost in any case. Once publication selection bias is accounted for, the aggregated effect is insignificant. Therefore, we argue that there is no persuasive empirical evidence on intra-industry spillovers. If there are any horizontal spillover effects, their signs and magnitudes vary from country to country and from industry to industry.

We further investigate which study aspects affect the reported significance and polarity of spillovers. We employ robust methods as well as pseudo-panel data methods (Meyer & Sinani 2009) and probability models (Wooster & Diebel 2006). We find that, in general, study results are predictably affected by its design, namely by the usage of cross-sectional or panel data, industry- or firm-level aggregation, and specification of the proxy of foreign presence in the industry. Our results suggest that cross-sectional studies tend to report excessively high spillovers, as well as models with industry-level aggregation and employment as a proxy for foreign presence do. However, this pattern appears to become weaker over time, suggesting that newer studies may suffer from such a bias less. Following Card & Krueger (1995), we test for publication bias in the spillover literature. We do not find evidence of publication bias employing this methodology. When the preferred funnel asymmetry test (Doucouliagos & Stanley 2009) is used, however, moderate publication bias is identified in the literature.

Contrary to Gorg & Strobl (2001), our meta-regression analysis shows that the definition of the proxy for foreign presence is important as well and can also bring predictable results. But many studies do not report sensitivity analysis with respect to the definition of foreign presence. When they do, as for instance Gersl (2008), they often find that the spillover effect is not robust. Such pattern of predictability is widespread in economics research (more in Stanley 2001). It is natural that heterogeneous research brings heterogeneous results. Researchers should, however, be aware of the predictability pattern, best identified by meta-regression analyses, and report thorough robustness checks.

2.3 Demand for Gasoline Is More Price-Inelastic than Commonly Thought

The chapter on the price elasticity of gasoline demand is a joint work with Karel Janda and Tomas Havranek and was published in Energy Economics (Havranek et al. 2012). For the purposes of government policy concerning energy security, optimal taxation, and climate change, precise estimates of the price elasticity of gasoline demand are of principal importance. For example, if gasoline demand is highly price-inelastic, taxes will be ineffective in reducing gasoline consumption and the corresponding emissions of greenhouse gases. During the last 30 years the topic has attracted a lot of attention of economists who produced a plethora of empirical estimates of both short- and long- run price elasticities. Yet the estimates vary broadly.

Two international meta-analyses of the elasticity of gasoline demand have been conducted (Espey 1998; Brons et al. 2008). These meta-analyses examine carefully the causes of heterogeneity observed in the literature. The average short- and long- run elasticities found by these meta-analyses were -0.26 and -0.58 (Espey 1998) and -0.34 and -0.84 (Brons et al. 2008). None of the meta-analyses, however, corrected the estimates for publication selection bias. We employ recently developed meta-analysis methods to test for publication bias and estimate the corrected elasticity beyond. The mixed-effects multilevel meta-regression takes into account heteroscedasticity, which is inevitable in meta-analysis, and between-study heterogeneity, which is likely to occur in most areas of empirical economics. We do not, however, investigate heterogeneity explicitly, as this issue was thoroughly examined by the two previous meta-analyses. Although our estimates of average elasticity are not directly comparable with Espey (1998) and Brons et al. (2008), we argue there is a strong case for the presence of publication bias in favor of larger negative estimates of elasticities in the literature.

We find the publication bias to be strong in this literature; when we correct for the bias, we obtain estimates of short- and long-run elasticities that are approximately half, compared to the results of the previously published meta-analyses and also to the simple mean of all estimates in our sample of literature. If the simple mean reflects our pro-

fession's impression about the magnitude of the price elasticity of gasoline demand, the impression exaggerates the true elasticity twofold. The estimated elasticities corrected for publication bias, -0.09 for the short run and -0.31 for the long run, are average across many countries, methods, and time periods; we report them as reference values. A similar pattern of publication bias, however, is likely to appear in any subset of the literature. Thus large negative estimates of price elasticities should be taken with a grain of salt.

2.4 Publication Bias in the Literature on FDI Spillovers

Here I focus on spillovers from foreign direct investment to local firms; it is a joint work with Tomas Havranek published in the Journal of Development Studies (Havranek & Irsova 2012). Policy makers, especially in transition and developing countries, usually encourage inward FDI in expectation that domestic firms in the same sectors benefit from know-how brought by foreigner investors. Moreover, many of such policy makers believe those firms in supplier sectors benefit from direct knowledge transfers from foreigners, and perhaps also that firms in customer sectors benefit from higher-quality intermediate inputs produced by foreigners. The search for spillovers has given rise to a burgeoning stream of empirical literature in development economics, and we investigate 57 such papers in this meta-analysis.

In contrast to the earlier meta-analyses on FDI spillovers (Gorg & Strobl 2001; Meyer & Sinani 2009), we examine backward and forward spillovers in addition to horizontal spillovers. Using a large data set, we employ modern meta-analysis methods developed by Stanley (2005; 2008) to estimate the underlying spillover effects and the magnitude of publication bias. We present individual surveys for each country inspected in the literature and construct a unique cross-country data set of estimated spillovers. Furthermore, we retrieve estimates of publication bias for each study and examine how the intensity of publication selection depends on the characteristics of the authors, such as affiliation, experience, and tenure pressure.

Our results suggest that the average effect of foreign affiliates on the productivity of their local competitors (horizontal spillover) is economically insignificant. The effect of foreign affiliates on their local customers (forward spillover) is likewise negligible. On the other hand, we detect a statistically significant and economically meaningful effect of foreign affiliates on their local suppliers (backward spillover). Specifically, a 10-percentage-point increase in foreign presence is associated with a 1.2% boost to the productivity of domestic firms in supplier sectors. Such a spillover effect is consistent with subsidies for FDI. While the average backward spillover is robustly positive, it differs significantly across countries. The degree of economic development plays an important role in explaining the difference, but it is not the only one. In a companion paper (Havranek & Irsova 2011) we examine in detail what causes the differences in the reported FDI spillovers.

2.5 Cross-Country Heterogeneity in Intertemporal Substitution

This chapter focuses on the cross-country heterogeneity in intertemporal substitution; it is a joint work with Roman Horvath, Tomas Havranek, and Marek Rusnak (Havranek et al. 2015a). The paper was published in the Journal of International Economics. The elasticity of intertemporal substitution in consumption (EIS) reflects households' willingness to substitute consumption between time periods in response to changes in the expected real interest rate. Therefore it represents a crucial parameter for a wide range of economic models involving intertemporal choice, from modeling the behavior of aggregate savings and the impact of fiscal policy to computing the social cost of carbon emissions, and has been estimated and used by hundreds of researchers. Most of them would agree with Ai (2010, p. 1357), who starts his discussion of calibration by noting that "empirical evidence on the magnitude of the EIS parameter is mixed."

In this paper we collect 2,735 estimates of the elasticity of intertemporal substitution reported in 169 studies and review the literature quantitatively using meta-analysis methods. While controlling for differences in methodology, we focus on explaining country-level heterogeneity. The studies in our sample provide us with estimates of the EIS for 104 countries, and we show that the mean values reported for the countries vary substantially. We build on the vast amount of literature that explores the heterogeneity in the EIS at the micro level.

The mean estimate of the elasticity of intertemporal substitution reported in empirical studies is 0.5, but we show that cross-country differences are important. Since it is often unclear which aspects of methodology should matter for the magnitude of the estimated EIS, we include all 30 that we collect and employ Bayesian model averaging (Raftery et al. 1997) to deal with the resulting model uncertainty. Our findings suggest that a larger EIS is associated with higher per capita income of the country, and especially with higher stock market participation. According to our baseline model, a 10-percentage-point increase in the rate of stock market participation is associated with an increase in the EIS of 0.24. Moreover, wealth and asset market participation are also important at the micro level: studies estimating the EIS using a sub-sample of rich households or asset holders find on average an EIS larger by 0.21.

2.6 Selective Reporting and the Social Cost of Carbon

The last chapter focuses on social costs of carbon emissions; it is a joint work with Tomas Havranek, Karel Janda, and David Zilberman and

was published in Energy Economics (Havranek et al. 2015b). The social cost of carbon (SCC) is a key parameter for the formulation of climate policy. If the SCC was pinned down precisely, policy makers could use the parameter to set the optimal carbon tax. For this reason, dozens of researchers using different families of models have estimated the SCC – but their findings and the resulting policy implications vary greatly. Several previous studies have offered quantitative surveys of the literature (Tol 2005; 2013), focusing especially on the characteristics of study design that may influence the reported estimates, but no study has discussed nor tested for the potential selective reporting bias in the estimates of the social cost of carbon.

In contrast to most subjects of meta-analysis in economics, the SCC is not estimated in a regression framework. Rather, it is a result of a complex calibration exercise, and the uncertainty surrounding the estimates is usually determined via Monte Carlo simulations. Therefore the literature lacks the usual suspects when it comes to potential selective reporting: specification search across models with different control variables, choice of the estimation technique, and the selection of the data sample. On the other hand, the authors have the liberty to choose among many possible values of the parameters that enter the computation and influence both the estimated magnitude of the SCC and the associated uncertainty. Despite the difficulty in computing the SCC, we believe it is worth trying to pin down this crucial parameter. Testing for the potential selective reporting bias represents a part of this effort.

We examine 809 estimates of the SCC reported in 101 primary studies. We employ meta-regression methods commonly used in economics and other fields to detect potential selective reporting in the literature. Our results suggest that, on average, the authors of primary studies tend to report preferentially estimates for which the 95% confidence interval excludes zero, which creates an upward bias in the literature. In other words, we observe that small estimates of the SCC are associated with less uncertainty (expressed as the approximate standard error used to compute the lower bound of the confidence interval) than large estimates. The finding suggests that some small estimates with large uncertainty – that is, not ruling out negative values of the SCC – might be selectively omitted from the literature. Our results also indicate that selective reporting tends to be stronger in studies published in peer-reviewed journals than in unpublished manuscripts.

3. Conclusion: Lessons Learned

The papers included in the dissertation and summarized on the previous pages use many different meta-analysis methods, which might puzzle the reader. One of the reasons for the differences is the time when these papers where published, which reflects the evolution of meta-regression methods; sometimes, however, the choice of a particular meta-analysis technique depends on the specific data set or research question under examination. Up-to-date guidelines for conducting meta-regression analysis in economics are provided by Stanley et al. (2013), but I still consider it useful to briefly summarize my take on best-practice methods in the field and provide practical details that are missing in other guidelines. I structure the discussion into several paragraphs according to the issues that meta-analysts face.

Selection of primary studies. The first problem that a meta-analyst faces is which studies to include in the meta-analysis. The typical recommendation (Stanley 2001) is to use all studies estimating the parameter in question, if possible. Sometimes, however, such an approach is not feasible because hundreds or even thousands of papers exist on the topic. In this dissertation it is the case of, for example, the literature estimating the elasticity of intertemporal substitution in consumption. Rather than selecting a random sample of studies, I argue it makes sense to focus on published studies only. Published studies can be, ex ante, expected to be of higher quality, are better typeset (which makes data collection easier and reduces the danger of typos), and unlike working papers, for published studies there is only one version available (which makes it easier to date the study). Moreover, several meta-analyses show that there is little difference between published and unpublished studies in the extent of publication bias (Rusnak et al. 2013).

Tests of publication bias. I prefer the funnel asymmetry test discussed by Stanley (2005), because it has been shown to perform well in Monte Carlo simulations and is very intuitive. The test is based on the realization that in the absence of publication bias there should be no systematic relation between estimates and their standard errors. The authors of primary studies usually report t-statistics for their estimates, which means that they assume that the ratio of the estimates to their standard errors have a t-distribution, which in turn implies that estimates and standard errors should be statistically independent quantities. If, on the other hand, researchers prefer to publish estimates with a particular sign or statistical significance, estimates will be correlated with standard errors. The regression is heteroskedastic, so weighted least squares (with inverse of the variance as the weight) should be used. If possible, researchers should use study-level fixed effects and cluster standard errors at the study level. I also recommend using the inverse of the square root of the number of observations as an instrument for the standard error. If the meta-analyst fails to control for a method choice that affects the estimates and their standard errors in the same direction, he or she obtains biased estimates of the extent of publication bias. The inverse of the square root of the number of observations is usually a valid instrument, because it is obviously correlated with the standard error, but not correlated with most method choices.

Selection of variables. Most applications of meta-analysis involve dozens of variables that may potentially affect the magnitude of the parameter in question. It is not clear which variables should be selected in the baseline model, because for many of them we have little theoretical guidance (for example, the effect of the number of observations), but we still want to control for these aspects of study design. I recommend not to use sequential t-tests and remove the least significant variables one by one; such an approach is not statistically valid. Instead, meta-analysts should use Bayesian model averaging, which is a method that formally addresses model uncertainty in meta-analysis. The methods runs millions of regressions with different combinations of all explanatory variables and makes a weighted average over them (with weights being approximately proportional to the goodness of fit of the individual models).

Robustness checks. It is a matter of taste whether to use weighted least squares in meta-analysis when other explanatory variables than the standard error are included. Tom Stanley argues to always use weighted least squares, because of the heteroskedasticity problem and because weighting always gives priority to more precise results. I prefer not to weight the regression by precision if the regression contains variables defined on the study level, like the number of citations. Because precision differs for each estimate within a study, weighting by precision introduces artificial variation in these variables. Since both approaches often yield very different results, it might be a good idea to report the results of the other approach as a robustness check. Moreover, if the meta-analyst cannot use study-level fixed effects when estimating publication bias (for example, because many studies report only one estimate), it is advisable to report both simple OLS estimates and mixed-effects estimates (which give each study approximately the same weight even though different studies report a different number of estimates).

Judgment in meta-analysis. Although meta-analysis is a formal method of literature surveys, it does not mean that it is judgment-free. I argue that a good meta-analysis should discuss which method choices in the primary studies are preferable and, if possible, it should try to construct an estimate of the mean effect corrected for both publication bias and misspecifications in primary studies. In practical terms, the estimate is derived as a linear combination from the final specification, when the meta-analyst plugs in the preferred values for each variable (for example, "1" for the dummy variable that reflects whether the primary study controls for endogeneity by instrumenting the explanatory variable). Such "best practice" estimation is often controversial, but I believe it is the principal value added of any meta-analysis.

References

Ai, H. (2010): Information Quality and Long-Run Risk: Asset Pricing Implications. Journal of Finance 65(4): pp. 1333-1367.

Brons, M., Nijkamp, P., Pels, E. & Rietveld, P. (2008): A meta-analysis of the price elasticity of gasoline demand. A SUR approach. Energy Economics 30(5): pp. 2105-2122.

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

Djankov, S. & Murrell, P. (2002): Enterprise Restructuring in Transition: A Quantitative Survey. Journal of Economic Literature 40(3): pp. 739-792.

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

Espey, M. (1998): Gasoline demand revisited: an international meta-analysis of elasticities. Energy Economics 20(3): pp. 273-295.

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: pp. 232-247.

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

Havranek, T., Horvath, R., Irsova, Z. & M. Rusnak (2015a): Cross-Country Heterogeneity in Intertemporal Substitution. Journal of International Economics 96(1): pp. 100-118.

Havranek, T. & Irsova, Z. (2010): Meta-Analysis of Intra-Industry FDI Spillovers: Updated Evidence. Czech Journal of Economics and Finance (Finance a uver) 60(2): pp. 151-174.

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

Havranek, T. & Irsova, Z. (2015): Do Borders Really Slash Trade? A Meta-Analysis. CNB Working paper 1/2015, Czech National Bank, Prague.

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

Havranek, T., Irsova, Z., Janda, K. & Zilberman, D. (2015b): Selective Reporting and the Social Cost of Carbon. Energy Economics 51, pp. 394-406.

Jarrell, S. B. & Stanley, T. D. (1990): A meta-analysis of the union-nonunion wage gap. Industrial and Labor Relations Review 44(1): pp. 54-67.

McCallum, J. (1995): National Borders Matter: Canada-U.S. Regional Trade Patterns. American Economic Review 85(3): pp. 615-23.

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

Raftery, A. E., Madigan, D. & Hoeting, J. A. (1997): Bayesian Model Averaging for Linear Regression Models. Journal of the American Statistical Association 92: pp. 179-191.

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

Stanley, T., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., Nelson, J. P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R. S. & Rost, K. (2013): Meta-Analysis of Economics Research Reporting Guidelines. Journal of Economic Surveys 27(2): pp. 390-394.

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.

Tol, R. S. J. (2005): The marginal damage costs of carbon dioxide emissions: an assessment of the uncertainties. Energy Policy 33(16): pp. 2064-2074.

Tol, R. S. J. (2013): Targets for global climate policy: An overview. Journal of Economic Dynamics and Control 37(5): pp. 911-928.

Wooster, R. B. & D. S. Diebel (2006): Productivity Spillovers from Foreign Direct Investment in Developing Countries: A Meta-Regression Analysis. Working paper, Available at SSRN: http://ssrn.com/abstract=898400.