TECHNISCHE UNIVERSITÄT
CHEMNITZ

Faculty of Economics and
Business Administration

Conventional Wisdom, Meta-Analysis, and Research Revision in Economics

Sebastian Gechert, Bianka Mey, Matej Opatrny,
Tomas Havranek, T. D. Stanley, Pedro R. D. Bom,
Hristos Doucouliagos, Philipp Heimberger, Zuzana Irsova,
Heiko J. Rachinger

Chemnitz Economic Papers, No. 061, December 2023

Chemnitz University of Technology
Faculty of Economics and Business Administration
Thüringer Weg 7
09107 Chemnitz, Germany

Phone +49 (0)371 531 26000
Fax +49 (0371) 531 26019
https://www.tu-chemnitz.de/wirtschaft/index.php.en
wirtschaft@tu-chemnitz.de

Conventional Wisdom, Meta-Analysis, and Research Revision in Economics

Sebastian Gechert[1], Bianka Mey[2], Matej Opatrny[3], Tomas Havranek[4], T. D. Stanley[5], Pedro R.D. Bom[6], Hristos Doucouliagos[7], Philipp Heimberger[8], Zuzana Irsova[9] & Heiko J. Rachinger[10]

[1] *Chemnitz University of Technology; FMM Fellow.*
[2] *Chemnitz University of Technology*
[3] *Charles University Prague*
[4] *Institute of Economic Studies, Faculty of Social Sciences, Charles University Prague; CEPR, London; Meta-Research Innovation Center at Stanford*
[5] *Corresponding Author. Stanley@hendrix.edu. Department of Economics, Deakin University; Meta-Research Innovation Center at Stanford*
[6] *Deusto Business School, University of Deusto*
[7] *Department of Economics, Deakin University; IZA Bonn*
[8] *Vienna Institute for International Economic Studies (wiiw); FMM Fellow*
[9] *Anglo-American University, Prague*
[10] *Universitat de les Illes Balears, Mallorca*

December 22, 2023

Abstract. Over the past several decades, meta-analysis has emerged as a widely accepted tool to understand economics research. Meta-analyses often challenge the established conventional wisdom of their respective fields. We systematically review a wide range of influential meta-analyses in economics and compare them to 'conventional wisdom.' After correcting for observable biases, the empirical economic effects are typically much closer to zero and sometimes switch signs. Typically, the relative reduction in effect sizes is 45-60%.

Keywords. meta-analysis, systematic review, conventional wisdom

JEL classification. A14,B40,C10

1 Introduction

Dominant economic theories, seminal studies, and authoritative literature reviews often inform the conventional wisdom. Practical policy recommendation and implementation require knowledge of the specific values of important economic parameters; for example, the employment effects of minimum wage hikes, returns to education, the fiscal multiplier, the price elasticity of energy demand, or the intertemporal elasticity of substitution. Such conventional wisdom often defines the scope of public policy discussions and is used to calibrate economic models with these specific parameter values. However, conventional wisdom can also lead us astray.

Meta-analysis is the systematic and statistical analysis of all comparable empirical estimates of a specific parameter. It seeks to summarize, evaluate, and understand what we know about a given empirical economic question, phenomenon, policy parameter, or effect. Meta-analyses published in the *Journal of Economic Surveys* are compelled to follow guidelines that specify minimum standards for the coding, conducting, analyzing, and the reporting of quantitative surveys of economics research (Stanley, Doucouliagos, Giles, et al. 2013; Havránek, Stanley, et al. 2020). Meta-regression analysis (MRA) was developed specifically to explain and summarize the rich heterogeneity found among reported empirical economic estimates (Stanley and Jarrell 1989; Stanley 2001). By now, thousands of MRAs have been conducted on economic topics, with some hundred(s) of new studies produced each year (Havránek, Stanley, et al. 2020). MRA, with its ability to accommodate publication selection bias, was considered sufficiently important for understanding economics research to devote a special issue of the *Journal of Economic Surveys* (Roberts and Stanley 2005). Meta-analysis can reveal surprising truths about economics once publication selection and mis-specification biases have been identified and accommodated. Thus, a considerable number of meta-analyses in economics have questioned conventional wisdom in their respective fields.

This paper is a review of meta-analyses in the spirit of Ioannidis et al. (2017), Doucouliagos and Stanley (2013), Doucouliagos, Paldam, et al. (2018), and Gechert (2022). The purpose of this study is to compare the findings of influential meta-analyses to the 'conventional wisdom' about the same economic question or issue. What have we learned from meta-analyses of economics? How do their results differ from the conventional, textbook understanding of economics?

We identify 'influential' meta-analyses as those with at least 100 citations that were published in 2000 or later, and those that were recommended by a survey of members of the Meta-Analysis of Economics Research Network (MAER-Net) (https://www.maer-net.org/). Out of the full sample of 360 studies, 72 studies cover a general interest topic in economics and include original empirical estimates for a certain effect size. We narrow down further to those meta-analyses that provide both a simple mean of the original effect size and a corrected mean, controlling for publication bias or other biases. This gives us a final list of 24 studies covering the fields of growth and development, finance, public finance, education, international, labor, behavioral, gender, environmental, and regional/urban economics.

We compare the central findings of the meta-analyses to 'conventional wisdom' as classified by: (1) a widely recognized seminal paper or authoritative literature review; (2) the assessment of an artificial intelligence (AI), the GPT-4 Large Language Model (LLM); and (3) the simple unweighted average of reported effects included in the meta-analysis.

For 17 of these 24 studies, the corrected effect size is substantially closer to zero than commonly thought, or even switches sign. Statistically significant publication bias is prevalent in 17 of the 24 studies. Overall, we find that 16 of 24 studies show both a clear reduction in effect size and a statistically significant publication bias. Comparing the best estimate from the meta-analysis with the conventional wisdom from the reference study, the GPT-4 estimate, or the simple unweighted average, the relative reduction in the effect size is in the range of 45-60% in all three comparison cases. This is close to "Paldam's rule of thumb," according to which publication bias typically inflates the uncorrected mean of the effect size by a factor of two (Doucouliagos, Paldam, et al. 2018; Paldam 2022).

The paper is organized as follows: Section 2 reviews the related literature on the prevalence of publication selection bias. Section 3 describes how we selected the meta-studies in our final dataset and the information we collected. Section 4 provides a brief qualitative discussion of the contribution of selected meta-studies. Section 5 then shows the quantitative results from our survey. The final section concludes.

2 Publication selection bias: a renaissance

For many decades, publication selection bias has been widely recognized as a serious threat to the validity of empirical science (Sterling 1959; Rosenthal 1979; Lovell 1983; Hedges and Olkin 1985; DeLong and Lang 1992; Card and Krueger 1995; Ioannidis 2005; Stanley and Jarrell 2005; Stanley 2008; Stanley and Doucouliagos 2012; Stanley and Doucouliagos 2014, to cite but a few). Publication selection bias is the process of selecting which research findings to report based on their statistical significance or their consistency with conventional economic theory. Publication selection bias is the consequence of any type of preferential reporting of statistically significant findings, including the file drawer problem, publication bias, reporting bias, specification searching, questionable research practices, and p-hacking. As famously exposed by Leamer (1983), reported economic empirical findings are the consequence of the particular specification of innumerable combinations of independent variables, models, and methods (Sala-i-Martin 1997).

Evidence of exaggerated significance and effect size has been widely seen throughout the economics research literature. For example, a survey of 64,076 estimates from 159 areas of economics research found that reported results are typically exaggerated by a factor of two or more (Ioannidis et al. 2017). Two highly powered replication studies of multiple economics and behavioral experiments corroborate this doubling of effects size (Camerer, Dreber, Forsell, et al. 2016; Camerer, Dreber, Holzmeister, et al. 2018). Doucouliagos and Stanley (2013) show in a meta-meta-analysis of 87 empirical economics literatures that more competition and debate between rival theories (i.e., more pluralism) reduces publication bias. Methods to detect and correct publication selection bias were introduced and widely applied to economics in the May 2005 issue of the *Journal of Economics Surveys* (Roberts and Stanley 2005). Since then it has been standard to investigate publication selection bias when conducting meta-analyses in economics.

Recently, there has been a renaissance in documenting the effects of publication selection bias on reported economics research and in the development of new tools to identify and correct these biases. Franco et al. (2014) identify a severe under-representation of null findings, and simulations show how the meta-analysis of many smaller studies can reduce these biases (Hirschauer et al. 2022). Brodeur, Lé, et al. (2016) document the presence of p-hacking among 50,000 tests reported by top economics journals. Brodeur, Cook, et al. (2020) find that top journals are not exceptional in this. Moreover, they stress that alternative experimental designs differ in the magnitude of publication bias. Yet, top journals have the power to reduce this threat to the credibility of economics research. Askarov et al. (2023) uncover evidence from 345 economic meta-analyses that mandatory data-sharing policies at economics journals can be effective in reducing exaggerated effects and the severity of publication bias. Quite recently, Brodeur, Carrell, et al. (2023) investigate specific stages in the publication process and find that p-hacking is present prior to submission, somewhat mitigated by editors' desk decisions, but again enforced by reviewers who prefer statistically significant results. Frankel and Kasy (2022) develop a framework to discuss the trade-off between non-selective publication of findings and policy relevance under scarce journal capacity. An experiment confirms that a preference for statistically significant results is widely held among economics scholars

(Chopra et al. 2023). They promote pre-result reviews as a solution, which may be seen as part of a wider movement for more transparency and routine preregistration spearheaded by Christensen and Miguel (2018).

This study seeks to contribute to this rapidly growing literature by investigating how the findings from two dozen meta-analyses of specific economic areas of research compare to received conventional wisdom.

3 Data collection

To collect the required data and generate our final dataset, we followed several steps. First, to identify relevant studies, we searched Scopus, Google Scholar, and Web of Sciences (WoS). Second, we employed an expert list and surveyed MAER-net members regarding influential meta-analyses.

The database search proceeded as follows:

Scopus: We used the search string "meta AND analysis OR estimat" and selected several qualifiers: "Economics, Econometrics and Finance," "Articles in journals," "English language", and limited the keywords to "Meta-analysis" OR "Meta Analysis". Also, we employed two further eligibility criteria: published in 2000 or later and studies that have 100 cites or more. This yielded 164 studies.

WoS: We used the query "SU=Economics AND AK="meta-analysis" OR AK="meta" OR AK="meta analysis" OR AK="Meta- analysis" OR AK="Meta-Analysis" and applied the criteria on publication date and number of citations, which resulted in 57 studies from that source.[1]

Google Scholar: We used Harzing's Publish or Perish, employing the keywords "meta-analysis" and "economics," and set the sample period between 2000 and 2023, selecting 500 entries. After clearing for "Economic, econometric and finance published journal

[1] A summary of the WoS search query can be found at https://www.webofscience.com/wos/woscc/summary/5216412a-3195-4339-9976-3860847f306d-6f0209be/relevance/1

articles” and “English language,” 164 entries remained from that search. Meta-studies without an abstract and duplicates were dropped, which yielded a total of 333 candidate studies from the search queries.

In parallel, we conducted a simple voluntary expert survey in February 2023 to the members of the MAER-net community. We sent the survey to 150 members, asking the following questions:

- Do you think that there have been meta-studies that have overturned conventional economic wisdom? [YES/NO]
- Which meta-studies were most influential in terms of overturning conventional wisdom in economics? It would be especially helpful to us if you could also give some evidence/reasons for your answer.

Within the scheduled time of two weeks, we received 45 answers (a response rate of 30%). Of them, 29 (65.9%) answered *YES* to the first question, the remaining 16 answered *NO*. Regarding the second question, the experts suggested 27 additional candidate studies. Thus, in total our dataset comprises 360 meta-studies covering a broad range of research fields that have the potential for providing results possibly challenging conventional wisdom in their respective research field. Based on title and abstract screening, we coded these studies with the following qualifiers:

- = 1 if the study is closely related to economics; 0 otherwise.
- = 1 if the study topic is of general interest (subjectively chosen); 0.5 for unsure and 0 for not widely known.
- = 1 if the study empirically synthesizes primary studies; 0 otherwise.

Additionally, we broadly categorized them into “Agricultural/Ecological/Environmental,” “Behavioral,” “Health,” “Labor,” “Management,” “Meta_Analytical,” “Macro” and “Policy” to ensure that we captured a wide range of meta-studies.

For the next step, we continued with those studies that qualify in all three respects to ensure that they entail an important effect size estimate related to a conventional wisdom in economics. This procedure left us with 72 meta-analyses that underwent in-depth full-text screening to sample all information regarding relevant study characteristics and variables. This comprised the narrative as well as the meta-analytical point of view about the conventional wisdom.[2]

From the preliminary list of 72 studies, we excluded those that do not provide enough information for comparing an unweighted average of the effect sizes included and a measure of the underlying effect after correcting for publication bias or other biases. We further discarded all those meta-analyses that only report partial correlation coefficients (PCC), as they do not lend themselves to an economic interpretation of effect sizes. Our final sample consists of 24 meta-analyses, of which about 1/3 originated from the expert survey and 2/3 were found through database search. We review and analyse these studies in the following sections in terms of their impact on conventional wisdom in their respective research area.

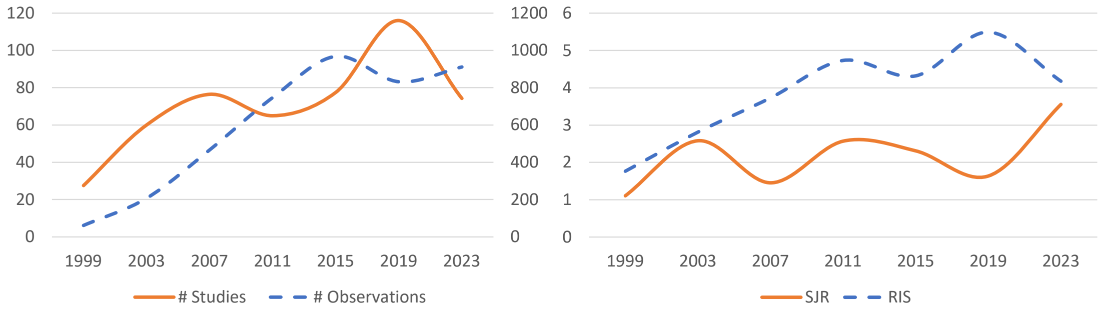
4 The growing relevance of meta-analysis in economics

In this section, we review the characteristics of the selected meta-studies and discuss in more detail some examples of influential meta-studies in specific economic fields. Figure 1 shows some overarching trends of the relevance of meta-analysis according to our wider sample of 72 studies. Figure 1a gives multi-year averages of the number of primary studies and observations included per meta-study according to their publication year. Figure 1b provides the multi-year averages of the impact factors according to the SCImago Journal Rank (SJR) and the Resurchify Impact Score (RIS).

Table 1 zooms in on the details of the 24 final studies, including their field and specific research question, the number of studies and estimates they cover, the number of cita-

[2]See the supplementary material in the online appendix for information on the full set of studies and the pre-selection of 72 studies.

Figure 1: Multi-year averages of characteristics of 72 pre-selected meta-analyses



(a) Number of primary studies and observations included

(b) Impact factors of journals where meta-analyses were published

Notes: The figure shows trends of the multi-year averages of the characteristics of the 72 pre-selected meta-studies. 1a: multi-year averages (based on publication year of the meta study) of the number of primary studies and observations included per meta-study. 1b: multi-year averages of the impact factors of journals where meta-analyses were published, according to the SCImago Journal Rank (SJR) and the Resurchify Impact Score (RIS). Source: https://www.resurchify.com/ranking

tions they have attracted, as well as recent impact factors and rankings of the respective journals in which they are published.

The following broad picture emerges: meta-analyses have covered many important topics in economics, and they provide well-cited benchmark estimates for their respective fields. While they were less popular in economics about 20 years ago, many of them have been published in top-tier journals in recent years. With the strong turn towards empirical evaluation in economics (Angrist and Pischke 2010; Angrist, Azoulay, et al. 2017; Paldam 2021) and the exploding number of available primary empirical studies, the demand for quantitative summaries has increased. At the same time, the meta-datasets have grown substantially over the years, pointing to the tremendous workload required to produce high-quality meta-analyses.

Apart from this broad assessment, what can we learn from single meta-studies in specific fields? Our set of studies covers important topics like finance, economic growth and development, public finance, education and inequality, international, labor, behavioral,

Table 1: Relevance and characteristics of 24 selected meta-studies

Meta-Study	Field	Research Question	Studies	Estimates	Cited	RIS	SJR
Abreu et al. (2005)	growth / development	convergence rate	48	619	158	4.8	1.8
Ashenfelter et al. (1999)	education / inequality	rate of return to schooling	27	96	825	2.4	1.5
Bandiera et al. (2021)	gender	gender-wise pay-incentive effects	15	17	15	5.4	
Bom and Ligthart (2014)	public finance / fiscal policy	private output elasticity of public capital	68	578	186	4.8	1.8
Disdier and Head (2008)	international	elasticity of trade volume to proximity	103	1467	1594	5.0	8.3
Doucouliagos and Stanley (2009)	labor	employment elasticity to minimum wage	64	1424	621	2.9	1.6
Doucouliagos, Stanley, and Giles (2012)	health	value of a statistical life	37	39	50	3.5	2.1
Feld and Heckemeyer (2011)	international	semi-elasticity of FDI to tax changes	45	704	125	4.8	1.8
Fidrmuc and Korhonen (2006)	business cycles	business cycle correlation of CEECs	35	463	151	2.4	1.2
Gechert (2015)	public finance / fiscal policy	fiscal multipliers	104	1069	295	1.2	0.6
Gechert et al. (2022)	macro	capital-labor substitution elasticities	121	3186	36	1.6	2.5
Havránek and Irsova (2011)	international	semi-elasticity of domestic firms' productivity to foreign presence	57	3626	240	3.7	3.6
Havránek (2015)	behavioral	elasticities of intertemporal substitution in consumption	169	2735	327	4.1	5.5
Havránek, Irsova, et al. (2022)	education / inequality	elasticity of substitution between skilled and unskilled labor	77	682	11	5.5	8.4
Imai et al. (2021)	behavioral	present-bias parameter in hyperbolic discounting	28	220	63	3.5	5.1
Kaiser et al. (2022)	finance	treatment effect financial education on knowledge and behavior	68	677	208	7.8	10.4
Koetse et al. (2008)	macro	capital-energy substitution elasticities	34	317	300	8.8	2.5
Labandeira et al. (2017)	environment	price elasticities of energy demand	428	1976	434	7.4	2.1
Longhi et al. (2005)	labor	wage elasticity of native workers to immigration	18	348	146	4.8	1.8
Melo et al. (2009)	regional / urban	urban agglomeration elasticities	34	729	260	2.5	1.1
Nijkamp and Poot (2005)	labor	wage elasticity to unemployment	17	208	114	4.8	1.8
Reynaud and Lanzanova (2017)	environment	value of lake ecosystem services	133	699	112	6.0	1.8
Rose and Stanley (2005)	international	effect of currency union on trade	34	754	168	4.8	1.8
Vooren et al. (2019)	labor	ALMP effect on labor market outcomes	57	645	118	4.8	1.8

Notes: The table presents characteristics of the final list of 24 meta-studies. Studies and Estimates give the number of primary studies and single estimates collected in the meta analysis. Cited: number of times the meta study was cited according to Google Scholar as of March 2023. SJR = SCImago Journal Rank, RIS = Resurchify Impact Score as of the year 2021. Source: https://www.resurchify.com/ranking gender, environmental and regional/urban economics. In the following, we will briefly summarize one example per field.

We start with *labor economics*, one of the most intensely researched fields in meta-analysis. At least four of the 24 studies in our selection are related to labor markets. Here, we focus on the effects of minimum wages. The elasticity of employment to changes in minimum wages is a classic example where meta-analysis has contributed to overturning conventional wisdom. It is also the topic that first raised awareness about publication bias in economics. Minimum wages were conventionally thought to reduce employment of low-wage workers according to the neoclassical theory of the labour market and the findings of influential empirical studies (e.g. Brown 1999). Card and Krueger (1994) shook this consensus with a quasi-experimental study that found either no adverse employment effects from raising the minimum wages in one of two adjoining US states, or even a positive employment effect. Card and Krueger (1995) also conducted a modest meta-analysis of 19 effects and attributed the exaggerated minimum wage effect to publication bias. Their work laid the basis for a dramatic overhaul of labor market theories and policy prescriptions. Another well-cited meta-analysis in this field, much more comprehensive and rigorous, is carried out by Doucouliagos and Stanley (2009). Both studies agree that there is severe publication bias in the literature, strongly overstating the negative employment effects of minimum wage hikes. Doucouliagos and Stanley (2009) also show that, after correction, no adverse employment effect remains. Since these meta-analyses, policy makers have become less skeptical about raising minimum wages.

Related to labor economics is the question of *gender* differences in the workplace. Bandiera et al. (2021) provide an excellent example of a very recent meta-analysis, published in the new and policy-oriented outlet *American Economic Review: Insights*. Bandiera et al. (2021) ask whether women indeed respond less to performance pay than men, a common assumption informed by studies on gender-specific risk aversion and self-confidence. Bandiera et al. (2021) hypothesize this effect is largely driven by self-selection of women (men) into less (more) competitive environments. Thus, they narrow their choice of primary studies to lab and field experiments that rule out or control for self-selection and provide at least two distinct treatments, one of which provides clearly higher-powered incentives. Their dataset contains 17 such high-quality experiments. In line with the methods proposed by Stanley, Doucouliagos, and Ioannidis (2017) and Furukawa (2019), selecting only high-quality studies can be another promising approach to correct for publication bias. Indeed, Bandiera et al. (2021) do not apply established econometric techniques of detecting publication bias and an underlying effect. Nevertheless, in their study set, they find on average that women perform better than men (even though the effect is not statistically significant) in performance-pay settings when self-selection is ruled out. This clearly contradicts the conventional wisdom of earlier studies, nuances the results, and establishes new standards to identify gender-specific treatment effects.

Studying responses to performance pay is also related to *behavioral economics*, another field where meta-analysts have made important contributions that have challenged stylized facts. A recent and excellent example is Imai et al. (2021), who consider 28 published articles with 220 estimates about present bias. Present bias means that peoples' implicit discount factors are larger for near-term comparisons than for decisions farther in the future (Laibson 1997). A typical effect of this psychological phenomenon would be procrastination in financial decisions. Present bias is allegedly often detected in experiments and has become a standard assumption in behavioral economics. Imai et al. (2021) show that the simple average from their primary studies indeed points to present bias, with an average of the parameter $\beta = 0.96$, statistically-significantly different from $\beta = 1$, the null hypothesis of no present bias. Note that the setting differs from typical hypothesis testing against a zero effect. When the null hypothesis is $\beta = 1$ and the expected alternative is $\beta < 1$, publication selection that discards statistically insignificant results and those with $\beta > 1$ would likely lead to a downward-biased average estimate of β. Indeed, this is what Imai et al. (2021) detect in their dataset. Considering various tests, they report a moderate publication bias and find the corrected mean, averaged over various tests in their full sample, to be actually 0.99, which is not statistically-significantly different from 1. Thus, present bias may be an overrated phenomenon, less prevalent than previously thought.

As pointed out, present bias has strong implications for *finance*, a field for which there is another recent representative meta-analysis, the one by Kaiser et al. (2022). This meta-study looks at the impact of financial education on financial knowledge and downstream behaviors. Kaiser et al. (2022) take into account 68 primary studies, which provide almost 700 observations on the impact of financial education. For the sake of homogeneity, we focus here on their findings about the impact on financial knowledge, as typically measured by standardized improvements in test scores. For this subset, Kaiser et al. (2022) establish that the publication probability of statistically insignificant results is low, pointing to inflationary publication bias. The simple average of treatment effects from financial education is 0.19 in their sample of primary studies, a bit lower than the well-known study by Bruhn et al. (2016) reports. Correcting for publication bias further reduces the underlying effect to 0.13. Thus, one might conclude from the evidence in Kaiser et al. (2022) that the effects of financial education may have been overstated in the past.

Ever since environmental issues and climate change have become mega topics, they have attracted a tremendous amount of empirical work. Meta-analysis can provide welcome summaries about important relations in *ecological and environmental economics* to experts and policy makers alike. A classic question in this field is the price elasticity of energy demand, which is central to the steering effects of energy taxes and emission certificates. Labandeira et al. (2017) accumulates the evidence of more than 400 studies that provide nearly 2,000 estimates of energy price changes in transport, heating, and electricity use. We focus here on the long-term elasticity, which provides a more comprehensive measure of steering effects. An early and influential survey by Dahl and Sterner (1991), finds an average long-term elasticity of -0.8 for gasoline demand. In contrast, the systematic meta-dataset in Labandeira et al. (2017) calculates a simple average of -0.52. Unfortunately, Labandeira et al. (2017) do not study the impact of publication bias or employ any other correction methods to arrive at a best practice estimate.

We now turn to *international economics*, another widely covered field in meta-analysis. Four out of the 24 studies in our final selection contribute to this field, including the most-cited meta-analysis in our selection (Disdier and Head 2008). This study addresses the question of how geographical distance affects bilateral trade flows. Their aim is to identify a "typical distance effect" and factors of heterogeneity. They do so by collecting almost 1,500 estimates from more than 100 primary studies. The distance effect is usually estimated as θ, "the negative of the elasticity of bilateral trade with respect to distance" (Disdier and Head 2008, p.39), in a gravity equation. The simple average estimate from their study, about 0.9, is on the lower end of the range of estimates according to conventional literature surveys. However, it still confirms the typical "puzzle" in the literature that distance is much more influential on trade flows than would be expected from mere transport costs alone. Disdier and Head (2008) also address publication bias, using a simple OLS regression of θ estimates on their standard errors. They only find a weakly positive correlation, where publication bias is actually statistically insignificant. The bias-corrected estimate is therefore close to the simple average, around 0.8. That is, the puzzling distance effects on trade are slightly weakened, though still existent, according to this meta-analysis. Disdier and Head (2008) surmise that the impact of publication bias might be weaker in this literature since distance is often not the main variable of interest but a mere control variable in gravity models. It is also not a direct policy concern as, for example, the effects of minimum wages on employment.

The effect of distance on trade is somewhat related to agglomeration effects, a central topic in *regional and urban economics*. The study by Melo et al. (2009) asks whether the literature on agglomeration finds a genuine effect on productivity and which factors may explain differences in outcomes. They exploit more than 700 elasticity estimates from 34 primary studies and find that the effects vary a lot with country-specifics and industrial coverage, among other factors. The simple average of all elasticity estimates is around 0.06, similar to the central finding of the seminal study by Ciccone and Hall (1996). However, Melo et al. (2009) detect asymmetric publication bias in favor of more positive effects. The bias-corrected estimate falls to around 0.04, not overthrowing but clearly reducing the productivity benefits of dense agglomerations from a suspected strong effect to a more moderate effect.

Agglomeration effects may be supported or hampered by public infrastructure and other public capital. There are also several meta-analyses in the realm of *public finances and fiscal policies*. The most cited one in our selection, by Bom and Ligthart (2014), focuses on the productivity of public capital. It collects 68 studies with almost 600 estimates of the output elasticity of public capital. The simple mean of the elasticity according to the meta-analysis is about 0.19, only about half of the large estimates found in the seminal article by Aschauer (1989). Moreover, Bom and Ligthart (2014) detect positive publication bias in the literature, and arrive at a best publication bias corrected estimate of 0.11. Nevertheless, this is a sizeable and statistically significant average effect of public capital on output, which leads the authors to conclude that public capital is in short supply in OECD countries and could be extended to the benefit of societal welfare.

Public capital and institutions may be one of the driving forces of *economic growth and development* of countries. The study by Abreu et al. (2005) provides an excellent example of an early meta-analysis in economics that covers a highly relevant topic, the "legendary" measure of $\beta = 2\%$ rate of conditional convergence between income levels of poor and rich countries. This value of 2%, which was established for several conditions and samples by, among others, Sala-i-Martin (1996), has been a major stylized fact in the growth literature for many years. It posed a puzzling case to the baseline neoclassical growth model of Solow, Swan and Ramsey, which would predict a much faster catching-up of poor countries through capital accumulation. Abreu et al. (2005) collect about 600 estimates from the empirical literature on β-convergence. They show that the dispersion of estimates is indeed wide, questioning the confidence in a single measure of 2% as a "natural constant." Moreover, Abreu et al. (2005) find a systematic relation to unobserved heterogeneity in technology levels that, if taken into account, raises the value of β. At the same time, they detect statistically significant publication bias that inflates the average of reported estimates. They conclude with a corrected average convergence rate of 2.9%, which is, however, subject to strong heterogeneity. This finding points to the necessity to take into account country-specific or region-specific circumstances and institutions that cannot be captured by a universal growth model.

A more fundamental parameter that is related to growth and development, as well as other topics in *macroeconomics* in general, is the elasticity of substitution between capital and labor (σ) in production functions, which has been studied by Gechert et al. (2022). The size of the elasticity has important implications for growth and business cycle models, the effectiveness of monetary and tax policies, or the functional distribution of incomes. In many macroeconomic models, the elasticity is conveniently assumed to be equal to $\sigma = 1$ (the Cobb-Douglas case). More flexible CES approaches tend to assume a value of 0.5. Gechert et al. (2022) collect more than 3,000 estimates from 121 studies. They show that indeed a simple mean of all estimates is close to the Cobb-Douglas case with $\hat{\sigma} = 0.9$. However, publication bias is prevalent in this literature, where negative values are implausible and an attractor for large positive estimates exists. Correcting for this bias and following some best practices from the literature leads to a consensus estimate of $\sigma = 0.3$, strongly rejecting the conventional Cobb-Douglas assumption. Under these conditions, labor and capital are gross complements. Thus, wage rises in relation to the costs of capital may not lead to a strong replacement of labor by capital. Consequently, alternative explanations have to be found for the secular decline in the labor share. If the elasticity of substitution is far below one, the fall in the labor share cannot easily be explained by capital deepening in a neoclassical growth model, as in Piketty and Zucman (2014). Directed technical change or an increase in market concentration are alternative explanations that do not hinge on high values of σ.

If capital-labor substitution is a fundamental parameter in macroeconomics, so is the elasticity of substitution between skilled and unskilled labor, a key concept not only in macroeconomics but also in *education and inequality economics*. The recently published meta-analysis by Havránek, Irsova, et al. (2022) considers this relation, which is often assumed to be 1.5 in model parameterization. This would imply that skilled and unskilled workers are gross substitutes, though not too strongly. For reasons of identification, most primary studies actually estimate the negative inverse of the elasticity, which, under the conventional assumptions, would amount to -2/3. As an important new feature, Havránek, Irsova, et al. (2022) take into account both publication bias, which would typically lead to inflated estimates of the inverse elasticity (i.e., a downward biased elasticity), and attenuation bias, which would draw the inverse elasticity towards zero (i.e., an upward biased elasticity). Their central finding is that publication bias trumps attenuation bias, and that an unbiased average estimate of the negative inverse should rather be around -1/4, i.e., a strong substitution elasticity of close to 4. This implies that skill-biased technical change has a strong effect on the relative demand for skilled labor and the skill premium, stronger than was previously held.

5 Quantifying relative research revision by meta-analysis

Many of the aforementioned examples, even though they consider very different research questions, seem to share a common pattern: the parameter of interest, after a thorough and comprehensive collection of empirical evidence, and after accounting for publication selection bias as well as influential control variables, is often smaller in absolute terms than the common wisdom as derived from an influential primary study, a classic literature review, mere conventions, or when considering simply the unweighted average from the meta-sample. Such a pattern has already been documented in the meta-meta-analyses of Ioannidis et al. (2017) and Doucouliagos, Paldam, et al. (2018), who show that effect sizes systematically appear inflated in several literatures if publication bias is prevalent.

We assess this pattern more systematically for our selection of 24 meta-studies. Table 2 compares the corrected mean from the meta-analysis with (i) a narrative reference study, (ii) the answer from an artificial intelligence (AI), and (iii) the unweighted simple mean from the meta-analysis.

(i) For each of the meta-studies, we searched for a conventional wisdom point estimate from a narrative reference study. This seminal paper can be a recent conventional literature survey or a highly cited and well-published primary study that set the tone for follow-up primary studies in the respective literature. Importantly, the narrative study needs to provide a preferred estimate of the parameter of interest. The reference study is cited in column (2), and its qualitative as well as quantitative assessment are given in columns (3) and (4) of Table Table 2.

(ii) Alternatively, we also asked an AI, specifically the large language model (LLM) GPT-4[3], for a best possible point estimate of the parameters of interest in our 24 meta-studies. The generic question to the AI for each of the 24 fields reads as follows:

> *Please provide an estimate of the effect of [research question of the meta-analysis] based on all relevant literature up to year [publication date of the meta-analysis]. That is, the estimate should reflect the state of knowledge prior to the publication of the meta-analysis [title of the meta-analysis] on*

[3] GPT-4 has the advantage that it is well-established and has access to an up-to-date database. While it is not open-access, it proved more powerful in providing a quantitative assessment than open-access alternatives like ChatGPT or the Bing LLM.

Table 2: Conventional wisdom and results from the 24 selected meta-analyses

Meta-Study	Seminal Study Conventional Wisdom (CW) (1) Reference	(2) Qualitative	(3) Quant	GPT4 (4) AI CW	Meta-Finding (5) Simple Mean	(6) Correct. Mean	(7) Pub Bias
Abreu et al. (2005)	Sala-i-Martin (1996)	+: poor countries catch up	2.00	2.00	4.30	0.30	yes
Ashenfelter et al. (1999)	Psacharopoulos (1994)	+: school years increase earnings	0.09	0.09	0.07	0.07	yes
Bandiera et al. (2021)	Gneezy et al. (2003)	−: women respond less to performance pay	-0.28	[n/a]	0.08	0.07	[n/a]
Bom and Ligthart (2014)	Aschauer (1989)	+: public capital enhances productivity	0.39	0.18	0.19	0.11	yes
Disdier and Head (2008)	Anderson and Newell (2003)	+: bilateral trade increases with proximity	1.30	0.95	0.91	0.80	no
Doucouliagos and Stanley (2009)	Brown (1999)	−: higher minimum wage reduces employment	-0.08	-0.10	-0.19	0.04	yes
Doucouliagos, Stanley, and Giles (2012)	OECD (2012)	+: large benefits from improving health/safety	3.90	6.00	9.50	1.66	yes
Feld and Heckemeyer (2011)	Bénassy-Quéré et al. (2005)	−: higher tax rates reduce FDI	4.79	2.50	3.35	1.74	yes
Fidrmuc and Korhonen (2006)	Artis and Zhang (1997)	+: synchronous business cycles of CEECs and Euro Area	0.60	0.60	0.15	0.16	no
Gechert (2015)	Ramey (2019)	+: tax cuts strongly increase GDP	2.50	0.65	0.54	0.61	no
Gechert et al. (2022)	Knoblach and Stöckl (2020)	+: close to unity (Cobb-Douglas)	0.75	0.95	0.90	0.30	yes
Havránek and Irsova (2011)	Javorcik (2004)	+: spillovers from foreign affiliates to local firms	0.38	0.75	0.88	0.18	yes
Havránek (2015)	Hall (1988)	+: higher r shifts consumption to future	0.50	0.35	0.50	0.07	yes
Havránek, Irsova, et al. (2022)	Cantore et al. (2017)	− (inverse): $\lvert\varepsilon\rvert<1$ (skilled and unskilled labor gross substitutes)	-0.67	-0.57	-0.56	-0.27	yes
Imai et al. (2021)	Augenblick et al. (2015)	1-β>0: people are present-biased	0.07	0.20	0.04	0.01	yes
Kaiser et al. (2022)	Bruhn et al. (2016)	+: benefits of greater financial knowledge	0.23	0.20	0.19	0.13	yes
Koetse et al. (2008)	Berndt and Wood (1979)	+/−: C-E complements or substitutes	0.43	0.50	0.47	0.46	[n/a]
Labandeira et al. (2017)	Dahl and Sterner (1991)	−, $\lvert\varepsilon\rvert<1$: gasoline normal inelastic good, substantial long-run ε	-0.80	-0.70	-0.53	-0.53	[n/a]
Longhi et al. (2005)	Card (2001)	−: higher labor supply reduces wages	-0.15	-0.15	-0.12	-0.04	no
Melo et al. (2009)	Ciccone and Hall (1996)	+: agglomeration enhances productivity	0.06	0.04	0.06	0.04	yes
Nijkamp and Poot (2005)	Blanchflower and Oswald (2003)	−: wage curve downward-sloping	-0.10	-0.10	-0.12	-0.08	yes
Reynaud and Lanzanova (2017)	Egan et al. (2009)	+: ecosystem services increase valuation of lakes	153	[n/a]	315	153	yes
Rose and Stanley (2005)	Rose (2000)	+: currency unions increase trade	1.20	1.15	0.86	0.39	yes
Vooren et al. (2019)	Heckman et al. (1999)	+: ALMP improve labor market outcomes (long run)	0.03	0.10	0.02	0.004	yes

Notes: The table compares the findings of the 24 selected meta-analyses with those from a reference study in the respective field and the conventional wisdom estimate from GPT-4.

the same topic. The estimate should take into account all available scientific studies, not just one prominent study. At the same time, the estimate should rigorously summarize the conventional wisdom in the literature in year [publication date of the meta-analysis]. Answer like an economist and expert in this field. Provide the best possible point estimate of the effect together with the corresponding 95% confidence intervals.

The answer regarding the point estimate given by the AI is documented in column (5). Note that the AI sometimes only provides a range of estimates, of which we take the simple average. In two cases, the AI did not respond with a quantitative assessment. Nevertheless, in most cases, the AI gave an informative and deliberative answer, including a point estimate and a confidence interval. The full answers are provided in the supplementary material. On average, the AI's point estimate is quite close to the results from our own selection of seminal conventional studies (which was done beforehand on a different laptop).

(iii) Our third comparison (in column 6) is the simple unweighted mean of estimates included in the meta-analysis, which is usually given in the descriptive statistics of the meta-study. Such an unweighted average of a broad set of primary studies does not account for any corrections for publication bias or best practices. One might expect this measure to differ substantially from the estimate of the narrative reference study. While this is partly the case for individual research questions, on average, the figures do not differ too much. This might point to the performative power of seminal studies in setting an established reference value for the parameter of interest.

The three reference values can be compared to the corrected mean from the meta-study as documented in column (7). Usually, this corrected mean refers to an estimate from the meta-study after correcting for publication bias and/or defining a best practice estimate. Meta-analyses have applied various approaches to such corrections in the past, and only recently, has the field converged to established guidelines and standard test procedures (Stanley, Doucouliagos, Giles, et al. 2013; Havránek, Stanley, et al. 2020; Irsova et al. 2023). Thus, there is no single coherent way for extracting the corrected mean from the respective meta-analysis. Primarily, we referred to a preferred estimate from the meta-study and we document our choice in the supplementary material if more than one such candidate estimate is available in the meta-study.

It turns out that quite often the corrected mean from the meta-analysis is substantially closer to zero (or to the null hypothesis) than all of the three comparison measures. Very often, this lower value is driven by some sort of publication bias: 17 of the 24 studies detect a statistically significant publication bias (column 8).

In order to compare and quantify this pattern across studies, we set up Relative Research Revision ($R3$) indices for our three comparison metrics. The $R3$ is calculated as follows:

$$R3_i^j = \frac{MCM_i - CW_i^j}{CW_i^j} \tag{1}$$

where MCM_i is the meta corrected mean from field i, and CW_i^j is the conventional wisdom according to comparison metric j (from the narrative study, the AI, or the simple mean of the meta-study). The $R3$ index has the following useful properties: it gives the percentage change of the absolute value of the conventional wisdom effect size due to the meta-analysis. The percentage change is positive in cases when MCM_i and CW_i^j have the same sign and MCM_i exceeds the CW_i^j in absolute value (an upward revision). It is negative and between 0% and -100%, when MCM_i is closer to zero than CW_i^j (a downward revision). It exceeds -100% in cases of a sign reversal of the conventional wisdom.[4]

[4]Note that the $R3$ index can be transformed into the research inflation (RI) index of Ioannidis et al. (2017), which is defined as $RI = \frac{CW}{MCM} - 1$ and thus corresponds to $RI = \frac{-R3}{1+R3}$. The $R3$ index is more useful in our case as it signals downward revisions towards zero and reversals with the same negative sign and monotonously increasing magnitude, while upward revisions receive a positive sign. For the RI index, upward revisions and reversals would have the same sign, which would render the average of the index ambiguous.

Table 3: Relative Research Revision (R3) indices

Meta-Study	R3 Seminal	R3 AI	R3 Meta
Abreu et al. (2005)	-85%	-85%	-93%
Ashenfelter et al. (1999)	-24%	-24%	-7%
Bandiera et al. (2021)	-124%	[n/a]	-18%
Bom and Ligthart (2014)	-73%	-39%	-44%
Disdier and Head (2008)	-38%	-16%	-12%
Doucouliagos and Stanley (2009)	-155%	-141%	-122%
Doucouliagos, Stanley, and Giles (2012)	-57%	-72%	-83%
Feld and Heckemeyer (2011)	-64%	-31%	-48%
Fidrmuc and Korhonen (2006)	-73%	-73%	6%
Gechert (2015)	-75%	-6%	13%
Gechert et al. (2022)	-60%	-68%	-67%
Havránek and Irsova (2011)	-53%	-76%	-80%
Havránek (2015)	-85%	-79%	-85%
Havránek, Irsova, et al. (2022)	-60%	-53%	-51%
Imai et al. (2021)	-83%	-94%	-72%
Kaiser et al. (2022)	-43%	-36%	-32%
Koetse et al. (2008)	7%	-8%	-2%
Labandeira et al. (2017)	-34%	-25%	0%
Longhi et al. (2005)	-72%	-72%	-64%
Melo et al. (2009)	-35%	11%	-33%
Nijkamp and Poot (2005)	-23%	-23%	-35%
Reynaud and Lanzanova (2017)	0%	[n/a]	-51%
Rose and Stanley (2005)	-68%	-67%	-55%
Vooren et al. (2019)	-87%	-96%	-80%
Median	**-62%**	**-60%**	**-50%**
Mean	**-61%**	**-53%**	**-46%**

Notes: The table presents the calculations of the three relative research revision (R3) indices for the 24 final meta-analyses according to the information in Table 2.

The results for the $R3$ indices are given in Table 3. For 11 of the 24 studies, the downward revision is -50% or more extreme, consistently among all three $R3$ indices. In 17 cases, at least one of the $R3$ measures indicates such a strong downward revision. In two instances, the corrected effect size even switches sign. While the three $R3$ indices differ for each single case, considering their means and medians shows that they are astonishingly similar, falling within a close range from about -45 to -60%. Note that this pattern does not differ much between studies that entered through the expert survey and those from the database search. That is, the average corrected mean from a meta-study in our sample reduces the conventional-wisdom effect size by about half. This is a confirmation of Paldam's rule of thumb that the various incentives for publication selection inflate the average estimate typically by a factor of 2 (Ioannidis et al. 2017; Paldam 2022). It also resonates with Camerer, Dreber, Holzmeister, et al. (2018) who show that highly-powered replication studies of experiments in social sciences report, on average, only half of the effect size of the original study.

6 Conclusion

In a survey of meta-analyses in the spirit of Ioannidis et al. (2017), Doucouliagos and Stanley (2013), Doucouliagos, Paldam, et al. (2018), and Gechert (2022), we have found that many meta-analyses overturned conventional wisdom in their specific fields by exploiting comprehensive datasets of empirical estimates and by detecting publication bias. On average, estimates shrink by about half in absolute terms when comparing the unweighted average and the mean beyond publication bias, confirming "Paldam's rule" (Ioannidis et al. 2017; Paldam 2022). This finding also resonates with Camerer, Dreber, Holzmeister, et al. (2018), who show that highly-powered replication studies of experiments in social sciences report, on average, only half the effect size of the original study.

Our analysis lends support to the potential of meta-analysis to bring forward improvements regarding a more robust calibration of model-parameters, as well as the economic effect sizes of policy interventions. This could lead both to more accurate policy recommendations and to rethinking theoretical channels of impact.

Future research might decompose the contributions of different motives of publication bias, stemming from theory conformism, or a search/preference for statistically significant results. Moreover, the use of AI in meta-analysis surely has more potential to be exploited in the future. As more data points become available, it might also be interesting to evaluate whether the results of meta-analyses influence the range of empirical estimates that are published afterwards, and as such have a transforming impact on the prevalence of publication bias. Finally, meta-analysts themselves might establish pre-analysis standards in order to make the process of meta-data collection and analysis more transparent and replicable.

Acknowledgments

Gechert, Heimberger and Mey acknowledge financial support by the European Macro Policy Network (EMPN); Irsova and Havranek acknowledge support from the Czech Science Foundation (#24-11583S) and from NPO "Systemic Risk Institute" LX22NPO5101, funded by the European Union—Next Generation EU (Czech Ministry of Education, Youth and Sports, NPO: EXCELES).

References

Abreu, Maria, Henri L. F. de Groot, and Raymond J. G. M. Florax (2005). "A Meta-Analysis of b-Convergence: the Legendary 2%". In: *Journal of Economic Surveys* 19.3, pp. 389–420.

Anderson, Soren and Richard Newell (2003). "Simplified marginal effects in discrete choice models". In: *Economics Letters* 81.3, pp. 321–326.

Angrist, Joshua D., Pierre Azoulay, Glenn Ellison, Ryan Hill, and Susan Feng Lu (2017). "Economic Research Evolves: Fields and Styles". In: *American Economic Review* 107.5, pp. 293–297.

Angrist, Joshua D. and Jörn-Steffen Pischke (2010). "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics". In: *Journal of Economic Perspectives* 24.2, pp. 3–30.

Artis, Michael J. and Wenda Zhang (1997). "International Business Cycles and the ERM: Is There a European Business Cycle?" In: *International Journal of Finance & Economics* 2.1, pp. 1–16.

Aschauer, David Alan (1989). "Is public expenditure productive?" In: *Journal of Monetary Economics* 23.2, pp. 177–200.

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek (1999). "A review of estimates of the schoolingrearnings relationship, with tests for publication bias". In: *Labour Economics* 6.

Askarov, Zohid, Anthony Doucouliagos, Hristos Doucouliagos, and T. D. Stanley (2023). "The Significance of Data-Sharing Policy". In: *Journal of the European Economic Association* 21.3, pp. 1191–1226.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger (2015). "Working over Time: Dynamic Inconsistency in Real Effort Tasks *". In: *The Quarterly Journal of Economics* 130.3, pp. 1067–1115.

Bandiera, Oriana, Greg Fischer, Andrea Prat, and Erina Ytsma (2021). "Do Women Respond Less to Performance Pay? Building Evidence from Multiple Experiments". In: *American Economic Review: Insights* 3.4, pp. 435–454.

Bénassy-Quéré, Agnès, Lionel Fontagné, and Amina Lahrèche-Révil (2005). "How Does FDI React to Corporate Taxation?" In: *International Tax and Public Finance* 12.5, pp. 583–603.

Berndt, Ernst R. and David O. Wood (1979). “Engineering and Econometric Interpretations of Energy-Capital Complementarity”. In: *American Economic Review* 69.3, pp. 342–354.

Blanchflower, David G. and Andrew J. Oswald (2003). *The Wage curve*. Cambridge, MA: MIT Press.

Bom, Pedro R. D. and Jenny E. Ligthart (2014). “What Have We Learned From Three Decades Of Research On The Productivity Of Public Capital?” In: *Journal of Economic Surveys* 28.5, pp. 889–916.

Brodeur, Abel, Scott Carrell, David Figlio, and Lester Lusher (2023). “Unpacking p-Hacking and Publication Bias”. In: *American Economic Review* 113.11, pp. 2974–3002.

Brodeur, Abel, Nikolai Cook, and Anthony Heyes (2020). “Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics”. In: *American Economic Review* 110.11, pp. 3634–3660.

Brodeur, Abel, Mathias Lé, Marc Sangnier, and Yanos Zylberberg (2016). “Star Wars: The Empirics Strike Back”. In: *American Economic Journal: Applied Economics* 8.1, pp. 1–32.

Brown, Charles (1999). “Minimum wages, employment, and the distribution of income”. In: *Handbook of labor economics*. Ed. by Orley Ashenfelter and David E. Card. Vol. Chapter 32. Handbooks in Economics. Amsterdam: Elsevier, pp. 2101–2163.

Bruhn, Miriam, Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia (2016). “The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil”. In: *American Economic Journal: Applied Economics* 8.4, pp. 256–295.

Camerer, Colin F., Anna Dreber, Eskil Forsell, et al. (2016). “Evaluating replicability of laboratory experiments in economics”. In: *Science* 351.6280, pp. 1433–1436.

Camerer, Colin F., Anna Dreber, Felix Holzmeister, et al. (2018). "Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015". In: *Nature human behaviour* 2.9, pp. 637–644.

Cantore, Cristiano, Filippo Ferroni, and Miguel A. León-Ledesma (2017). "The dynamics of hours worked and technology". In: *Journal of Economic Dynamics and Control* 82, pp. 67–82.

Card, David E. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration". In: *Journal of Labor Economics* 19.1, pp. 22–64.

Card, David E. and Alan B. Krueger (1994). "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania". In: *American Economic Review* 84.4, pp. 772–793.

– (1995). "Time-Series Minimum-Wage Studies: A Meta-analysis". In: *American Economic Review* 85.2, pp. 238–243.

Chopra, Felix, Ingar Haaland, Christopher Roth, and Andreas Stegmann (2023). "The Null Result Penalty". In: *Economic Journal* early online.

Christensen, Garret and Edward Miguel (2018). "Transparency, Reproducibility, and the Credibility of Economics Research". In: *Journal of Economic Literature* 56.3, pp. 920–980.

Ciccone, Antonio and Robert E. Hall (1996). "Productivity and the Density of Economic Activity". In: *American Economic Review* 86.1, pp. 54–70.

Dahl, Carol and Thomas Sterner (1991). "Analysing gasoline demand elasticities: a survey". In: *Energy Economics* 13.3, pp. 203–210.

DeLong, J. Bradford and Kevin Lang (1992). "Are All Economic Hypotheses False?" In: *Journal of Political Economy* 100.6, pp. 1257–1272.

Disdier, Anne-Célia and Keith Head (2008). "The Puzzling Persistence of the Distance Effect on Bilateral Trade". In: *Review of Economics and Statistics* 90.1.

Doucouliagos, Hristos, Martin Paldam, and T. D. Stanley (2018). “Skating on thin evidence: Implications for public policy”. In: *European Journal of Political Economy* 54, pp. 16–25.

Doucouliagos, Hristos and T. D. Stanley (2009). “Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis”. In: *British Journal of Industrial Relations* 47.2, pp. 406–428.

– (2013). “Are All Economic Facts Greatly Exaggerated? Theory Competition And Selectivity”. In: *Journal of Economic Surveys* 27.2, pp. 316–339.

Doucouliagos, Hristos, T. D. Stanley, and Margaret Giles (2012). “Are estimates of the value of a statistical life exaggerated?” In: *Journal of Health Economics* 31.1, pp. 197–206.

Egan, Kevin J., Joseph A. Herriges, Catherine L. Kling, and John A. Downing (2009). “Valuing Water Quality as a Function of Water Quality Measures”. In: *American Journal of Agricultural Economics* 91.1, pp. 106–123.

Feld, Lars P. and Jost H. Heckemeyer (2011). “FDI and Taxation: A Meta-Study”. In: *Journal of Economic Surveys* 25.2, pp. 233–272.

Fidrmuc, Jarko and Iikka Korhonen (2006). “Meta-analysis of the business cycle correlation between the euro area and the CEECs”. In: *Journal of Comparative Economics* 34.3, pp. 518–537.

Franco, Annie, Neil Malhotra, and Gabor Simonovits (2014). “Publication bias in the social sciences: unlocking the file drawer”. In: *Science* 345.6203, pp. 1502–1505.

Frankel, Alexander and Maximilian Kasy (2022). “Which Findings Should Be Published?” In: *American Economic Journal: Microeconomics* 14.1, pp. 1–38.

Furukawa, Chishio (2019). *Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method*. Tech. rep.

Gechert, Sebastian (2015). “What fiscal policy is most effective? A meta-regression analysis”. In: *Oxford Economic Papers* 67.3, pp. 553–580.

Gechert, Sebastian (2022). "Reconsidering macroeconomic policy prescriptions with meta-analysis". In: *Industrial and Corporate Change* 31.2, pp. 576–590.

Gechert, Sebastian, Tomáš Havránek, Zuzana Irsova, and Dominika Kolcunova (2022). "Measuring capital-labor substitution: The importance of method choices and publication bias". In: *Review of Economic Dynamics* 45.July, pp. 55–82.

Gneezy, U., M. Niederle, and A. Rustichini (2003). "Performance in Competitive Environments: Gender Differences". In: *The Quarterly Journal of Economics* 118.3, pp. 1049–1074.

Hall, Robert E. (1988). "Intertemporal Substitution in Consumption". In: *Journal of Political Economy* 96.2, pp. 339–357.

Havránek, Tomáš (2015). "Measuring Intertemporal Substitution: The importance of method choices and selective reporting". In: *Journal of the European Economic Association* 13.6, pp. 1180–1204.

Havránek, Tomáš and Zuzana Irsova (2011). "Estimating vertical spillovers from FDI: Why results vary and what the true effect is". In: *Journal of International Economics* 85.2, pp. 234–244.

Havránek, Tomáš, Zuzana Irsova, Lubica Laslopova, and Olesia Zeynalova (2022). "Publication and Attenuation Biases in Measuring Skill Substitution". In: *Review of Economics and Statistics*, pp. 1–37.

Havránek, Tomáš, T. D. Stanley, et al. (2020). "Reporting guidelines for meta-analysis in economics". In: *Journal of Economic Surveys* 34.3, pp. 469–475.

Heckman, James J., Robert J. Lalonde, and Jeffrey A. Smith (1999). "The Economics and Econometrics of Active Labor Market Programs". In: *Handbook of labor economics.* Ed. by Orley Ashenfelter and David E. Card. Vol. Chapter 31. Handbooks in Economics. Amsterdam: Elsevier, pp. 1865–2097.

Hedges, Larry V. and Ingram Olkin (1985). *Statistical methods for meta-analysis.* Orlando: Academic Press.

Hirschauer, Norbert, Sven Grüner, and Oliver Mußhoff (2022). *Fundamentals of Statistical Inference*. Cham: Springer International Publishing.

Imai, Taisuke, Tom A. Rutter, and Colin F. Camerer (2021). “Meta-Analysis of Present-Bias Estimation using Convex Time Budgets”. In: *Economic Journal* 131.636, pp. 1788–1814.

Ioannidis, John P. A. (2005). “Why most published research findings are false”. In: *PLoS medicine* 2.8, e124.

Ioannidis, John P. A., T. D. Stanley, and Hristos Doucouliagos (2017). “The Power of Bias in Economics Research”. In: *Economic Journal* 127.605, pp. 236–265.

Irsova, Zuzana, Hristos Doucouliagos, Tomáš Havránek, and T. D. Stanley (2023). “Meta–analysis of social science research: A practitioner’s guide”. In: *Journal of Economic Surveys*.

Javorcik, Beata Smarzynska (2004). “Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages”. In: *American Economic Review* 94.3, pp. 605–627.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban (2022). “Financial Education Affects Financial Knowledge and Downstream Behaviors”. In: *Journal of Financial Economics* 145.2, pp. 255–272.

Knoblach, Michael and Fabian Stöckl (2020). “What Determines the Elasticity of Substitution between Capital and Labor? A Literature Review”. In: *Journal of Economic Surveys* 34.4, pp. 847–875.

Koetse, Mark J., Henri L.F. de Groot, and Raymond J.G.M. Florax (2008). “Capital-energy substitution and shifts in factor demand: A meta-analysis”. In: *Energy Economics* 30.5, pp. 2236–2251.

Labandeira, Xavier, José M. Labeaga, and Xiral López-Otero (2017). “A meta-analysis on the price elasticity of energy demand”. In: *Energy Policy* 102, pp. 549–568.

Laibson, David I. (1997). “Golden Eggs and Hyperbolic Discounting”. In: *Quarterly Journal of Economics* 112.2, pp. 443–477.

Leamer, Edward E. (1983). “Let’s Take the Con Out of Econometrics”. In: *American Economic Review* 73.1, pp. 31–43.

Longhi, Simonetta, Peter Nijkamp, and Jacques Poot (2005). “A Meta-Analytic Assessment of the Effect of Immigration on Wages”. In: *Journal of Economic Surveys* 19.3, pp. 451–477.

Lovell, Michael C. (1983). “Data Mining”. In: *Review of Economics and Statistics* 65.1, pp. 1–12.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland (2009). “A meta-analysis of estimates of urban agglomeration economies”. In: *Regional Science and Urban Economics* 39.3, pp. 332–342.

Nijkamp, Peter and Jacques Poot (2005). “The Last Word on the Wage Curve?” In: *Journal of Economic Surveys* 19.3, pp. 421–450.

OECD (2012). *Mortality Risk Valuation in Environment, Health and Transport Policies.* OECD.

Paldam, Martin (2021). “Methods Used in Economic Research: An Empirical Study of Trends and Levels”. In: *Economics* 15.1, pp. 28–42.

– (2022). *Research methods in economics.* Tech. rep.

Piketty, Thomas and Gabriel Zucman (2014). “Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010”. In: *Quarterly Journal of Economics* 129.3, pp. 1255–1310.

Psacharopoulos, George (1994). “Returns to investment in education: A global update”. In: *World Development* 22.9, pp. 1325–1343.

Ramey, Valerie A. (2019). “Ten Years After the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research?” In: *Journal of Economic Perspectives* 33.2, pp. 89–114.

Reynaud, Arnaud and Denis Lanzanova (2017). "A Global Meta-Analysis of the Value of Ecosystem Services Provided by Lakes". In: *Ecological economics : the journal of the International Society for Ecological Economics* 137, pp. 184–194.

Roberts, Colin J. and T. D. Stanley, eds. (2005). *Meta-regression analysis: Issues of publication bias in economics*. Malden, Mass.: Wiley-Blackwell.

Rose, Andrew K. (2000). "One money, one market: the effect of common currencies on trade". In: *Economic Policy* 15.30, pp. 8–45.

Rose, Andrew K. and T. D. Stanley (2005). "A Meta-Analysis of the Effect of Common Currencies on International Trade*". In: *Journal of Economic Surveys* 19.3, pp. 347–365.

Rosenthal, Robert (1979). "The file drawer problem and tolerance for null results". In: *Psychological Bulletin* 86.3, pp. 638–641.

Sala-i-Martin, Xavier X. (1996). "The Classical Approach to Convergence Analysis". In: *The Economic Journal* 106.437, p. 1019.

– (1997). "I Just Ran Two Million Regressions". In: *American Economic Review* 87.2, pp. 178–183.

Stanley, T. D. (2001). "Wheat from Chaff: Meta-Analysis as Quantitative Literature Review". In: *Journal of Economic Perspectives* 15.3, pp. 131–150.

– (2008). "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection". In: *Oxford Bulletin of Economics and Statistics* 70.1, pp. 103–127.

Stanley, T. D. and Hristos Doucouliagos (2012). *Meta Regression Analysis in Economics and Business*. New York: Routledge.

– (2014). "Meta-regression approximations to reduce publication selection bias". In: *Research Synthesis Methods* 5.1, pp. 60–78.

Stanley, T. D., Hristos Doucouliagos, Margaret Giles, et al. (2013). “Meta-Analysis of Economics Research Reporting Guidelines”. In: *Journal of Economic Surveys* 27.2, pp. 390–394.

Stanley, T. D., Hristos Doucouliagos, and John P. A. Ioannidis (2017). “Finding the Power to Reduce Publication Bias”. In: *Statistics in Medicine* 36.10, pp. 1580–1598.

Stanley, T. D. and Stephen B. Jarrell (1989). “Meta-regression Analysis: A Quantitative Method of Literature Surveys”. In: *Journal of Economic Surveys* 3.2, pp. 161–170.

– (2005). “Meta-regression Analysis: A Quantitative Method of Literature Surveys”. In: *Journal of Economic Surveys* 19.3, pp. 299–308.

Sterling, Theodore D. (1959). “Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance - Or Vice Versa”. In: *Journal of the American Statistical Association* 54.285, p. 30.

Vooren, Melvin, Carla Haelermans, Wim Groot, and Henriëtte van den Maassen Brink (2019). “THE EFFECTIVENESS OF ACTIVE LABOR MARKET POLICIES: A META-ANALYSIS”. In: *Journal of Economic Surveys* 33.1, pp. 125–149.