Universität
Augsburg
University

Meta-Analysis in Finance:
Applications and Advances

DISSERTATION

zur Erlangung des akademischen Grades

Dr.-Ing.

eingereicht an der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

der Universität Augsburg

von

Jerome-Leonard Jacob Geyer-Klingeberg

Augsburg, März 2022

Erstgutachter: Prof. Dr. Andreas Rathgeber

Zweitgutachter: Prof. Ing. Evžen Kočenda, Ph.D.

Drittgutachter: Prof. Tom Stanley, Ph.D.

Termin der mündlichen Prüfung: 04.07.2022

Abstract

The volume of empirical research in finance exhibits a strong upward trajectory. On the one hand, the large number of research publications often produces contradictory results for the same phenomenon under examination. On the other hand, a credibility crisis calls into question the reliability of empirical research findings and the transparency of the academic research process that produces the results. Consequently, there is a need to objectively reflect and consolidate previous empirical research results and to correct biases that affect the validity of scientific evidence. Meta-analysis is a research approach that comes with this capability.

Meta-analysis is a secondary research method used to synthesize existing empirical results, to detect and explain consistencies and inconsistencies among research findings, and to identify and filter out distorting effects related to publication selection and model misspecification. Although meta-analysis is a standard tool for research synthesis and evidence-based decision-making in many related research disciplines such as economics and management science, it has rarely been applied in finance. The aim of this thesis is to structurally introduce meta-analysis in finance as a complement to primary research and traditional narrative reviews by providing an objective and statistical approach to the accumulation of scientific knowledge.

Chapter 2 provides a comprehensive overview of the opportunities that meta-analysis offers for finance research, recent applications of meta-analysis in finance, and the challenges and limitations associated with it. Chapter 3 presents an applied meta-analysis of 1,016 empirical effects obtained from 71 previous studies that estimate the impact of corporate financial hedging on firm value. In Chapter 4, a Monte Carlo simulation is conducted to compare the statistical properties of common weighting schemes in meta-regression analysis under the practical condition that multiple and dependent estimates are reported in the same study, which is the norm in finance. By introducing, applying, and further advancing meta-analysis in the context of finance research, this thesis aims to increase the awareness and acceptance of meta-analysis as well as to promote its future application in the field.

Bibliographic Record

Geyer-Klingeberg, Jerome: *Meta-Analysis in Finance: Applications and Advances*. Dissertation Thesis. University of Augsburg, Faculty of Mathematics, Natural Sciences, and Materials Engineering, Institute of Materials Resource Management. March 2022.

Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

AFA	American Finance Association
AJG	Academic Journal Guide by the Chartered Association of Business Schools
BIS	Bank of International Settlements
BLUE	Best Linear Unbiased Estimator
BMA	Bayesian Model Averaging
CI	Confidence Interval
CP	Commodity Prices
ERM	Enterprise Risk Management
ES	Effect Size
FAS	Financial Accounting Standards
FCFF	Free Cash Flow to the Firm
FEM	Fixed Effects Model
FV	Firm Value
FX	Foreign Exchange Rates
GDP	Gross Domestic Product
GLS	Generalized Least Squares
GMM	General Method of Moment
GW	Generalized Weights
IAS	International Accounting Standard
IFRS	International Financial Reporting Standards
IR	Interest Rates
ISO	International Organization for Standardization
IV	Instrumental Variable
MAER-Net	Meta-Analysis in Economics Research Network
MASEM	Meta-Analytic Structural Equation Modeling
MSE	Mean Squared Error
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
OTC	Over-The-Counter
PEESE	Precision-Effect Estimate with Standard Errors
PIP	Posterior Inclusion Probability
PMP	Posterior Model Probability
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REM	Random Effects Model
SE	Standard Error
SJR	Scimago Journal Ranking
TSSEM	Two-Stage Structural Equation Modeling
UIP	Unit Information Prior
WAAP	Weighted Average of Adequately Powered
WLS	Weighted Least Squares

List of Important Variables

α	True population effect size
$\hat{\alpha}$	Estimate of the true population effect size
β	Meta-regression coefficient
$\hat{\beta}$	Estimate of the meta-regression coefficient
e	Error term in the meta-regression
G	Gini coefficient
h	Index of the observations in a primary study, $h = 1, \dots, N$
HC	Continuous measure of corporate hedging activities
HD	Hedging dummy variable (= 1 for hedgers, = 0 for non-hedgers)
HP	Hedging premium calculated from the results reported in the primary studies
i	Index of the primary studies, $i = 1, \dots, L$
j	Index of the effect size estimates, $j = 1, \dots, K$
K	Total number of effect size estimates
l	Index of the moderator variables in the meta-regression, $l = 1, \dots, P$
L	Total number of primary studies
m	Number of effect size estimates reported in a primary study
M	Moderator variable for model misspecification
M'	Moderator variable for systematic heterogeneity
N	Total number of observations in a primary study
P	Total number of moderator variables in a meta-regression
r	Index of the number of meta-analyses examining the same effect, $r = 1, \dots, R$
R	Total number of meta-analyses conducted on the same effect
τ^2	Between-study variance
u	Error term in primary regression
V	Variance of the effect size estimate
w	Meta-regression weight
X	Independent variable in the primary study regression
y	Dependent variable in the meta-regression
z	Dependent variable in the primary study regression

List of Co-Authored Papers

During my doctoral studies, I have (co-)written thirteen papers that are published in referred academic journals, including eleven applied meta-analysis papers in finance and economics, as well as two primary studies in the field of commodity finance:[1]

1. Determinants of the WTI-Brent Price Spread Revisited
   The Journal of Futures Markets 41(5), pp. 736-757, 2021
   with Andreas Rathgeber

2. Rather Complements than Substitutes:
   Firm Value Effects of Capital Structure and Financial Hedging Decisions
   International Journal of Finance and Economics 26(4), pp. 4895-4917, 2021
   with Markus Hang, Andreas Rathgeber, and Stefan Stöckl

3. *Corporate Hedging and Firm Value: A Meta-Analysis
   The European Journal of Finance 27(6), pp. 461-485, 2021
   with Markus Hang and Andreas Rathgeber

4. Interaction Effects of Corporate Hedging Activities for a Multi-Risk Exposure:
   Evidence from a Quasi-Natural Experiment
   Review of Quantitative Finance and Accounting 56, pp. 789-818, 2021
   with Clemence Alasseur, Markus Hang, Andreas Rathgeber, and Lena Wichmann

5. The Impact of Speculation on Commodity Prices: A Meta-Granger Analysis
   Journal of Commodity Markets 22, 100148, 2021
   with Marie Hütter, Andreas Rathgeber, Florian Schmid, and Thomas Wimmer

6. *Meta-Analysis in Finance Research: Opportunities, Challenges, and
   Contemporary Applications
   International Review of Financial Analysis 71, 101524, 2020
   with Markus Hang and Andreas Rathgeber

7. Reporting Guidelines for Meta-Analysis in Economics
   Journal of Economic Surveys 34(3), pp. 469-475, 2020
   with Tomáš Havránek, Tom Stanley, Hristos Doucouliagos, Pedro Bom,
   Ichiro Iwasaki, Robert Reed, Katja Rost, and Robbie van Aert

8. What Drives Financial Hedging?
   A Meta-Regression Analysis of Corporate Hedging Determinants
   International Review of Financial Analysis 61, pp. 203-221, 2019
   with Markus Hang and Andreas Rathgeber

[1] The working papers of the published articles were presented in more than 20 sessions at various academic conferences. The full list of papers is available online on https://scholar.google.de/citations?user=eFdRttQAAAAJ.

9. It's Merely a Matter of Time: A Meta-Analysis of the Causality between Environmental Performance and Financial Performance
   Business Strategy and the Environment 28(2), pp. 257-273, 2019
   with Markus Hang and Andreas Rathgeber

10. Economic Development Matters: A Meta-Regression Analysis on the Relation between Environmental Management and Financial Performance
    Journal of Industrial Ecology 22(4), pp. 720-744, 2018
    with Markus Hang, Andreas Rathgeber, and Stefan Stöckl

11. Measurement Matters – A Meta-Study of the Determinants of Corporate Capital Structure
    Quarterly Review of Economics and Finance, 68, pp. 211-225, 2018
    with Markus Hang, Stefan Stöckl, and Andreas Rathgeber

12. Do Stock Markets React to Soccer Games? A Meta-Regression Analysis
    Applied Economics 50(19), pp. 2171-2189, 2018
    with Markus Hang, Stefan Stöckl, and Andreas Rathgeber

13. What do we Really Know about Corporate Hedging – A Meta-Analytical Study
    Business Research 11(1), pp. 1-31, 2018
    with Markus Hang, Stefan Stöckl, Andreas Rathgeber, and Matthias Walter

Two of the above papers, marked with an asterisk (*), comprise the foundation for this thesis. Chapter 3 is based on paper 6 and incorporates the introduction to meta-analysis, its opportunities, and challenges in the finance field. Chapter 4 is based on paper 3 and meta-analyzes the impact of corporate financial hedging on firm value. The respective sections of the published papers have been revised and extended for this thesis.

Chapter 1. Introduction

"[...] the foundation of science is the cumulation of knowledge from the results of many studies."

(Hunter and Schmidt, 2004: xxvii)

This chapter discusses the rapid growth of empirical research output in finance and the related heterogeneity of the empirical results (Section 1.1). It outlines the demand for meta-analysis in finance (Section 1.2) and presents the objectives and structure of the thesis (Section 1.3).

1.1. Volume, Velocity, Variety: The 3 V's of Empirical Research Output

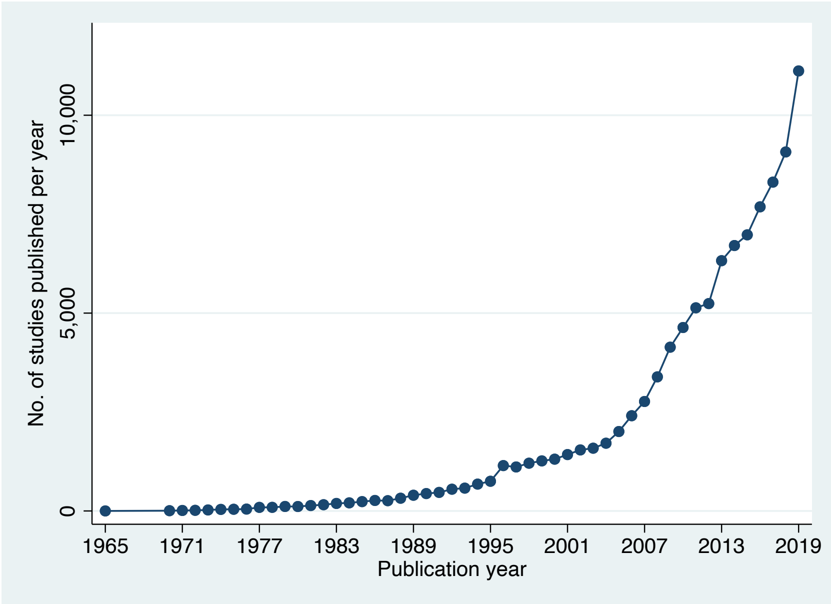
We are living in the age of big data. Along with the many bytes that we produce every day by sending emails, using social media, or taking photos, we are also surrounded by huge volumes of data in scientific research, both in terms of the data input we use for empirical analyses as well as the output we create as a result of scientific research. Here, I focus on the output side of empirical research. Adopting the characteristics of big data from computer science (McAfee and Brynjolfsson, 2012: 4-5) suggests three key features of big data related to empirical research output: the abundance of published research (volume), the time it takes to produce new empirical output (velocity), and the diversity of results (variety).

Figure 1 captures the volume and velocity of empirical research output. The graph illustrates the number of Scopus-listed publications that include alternative keywords for 'finance' and 'empirical analysis'. The number of articles increased from one publication in the year 1965 to 11,120 publications in the year 2019.[2] Accordingly, finance researchers have produced more empirical studies in the past seven years than in the entire preceding period. This trend is underlined by the record number of paper submissions reported by journal editors and conference organizers.[3] Volume and velocity are observable not only in the finance field as a whole, but also in its specific topic areas. For example, Harvey et al. (2016) compare 316 papers on cross-sectional return patterns, Hang et al. (2018) analyze 591 studies on the determinants of corporate capital structure, and Veld et al. (2020) aggregate 199 studies on the wealth effects of seasoned equity offerings. Extrapolating this growth, we could soon have research topics in finance with a thousand or more empirical studies examining the same phenomenon.

[2] It should be noted that Scopus also listed more journals over time.

[3] For example, according to an email sent by the conference organizers to all submitters on May 31, 2019, the American Finance Association received more than 2,000 paper submissions for its 2020 annual meeting.

Figure 1. Growth of Empirical Research Output in Finance



Notes: This graph shows the number of Scopus search results between 1965 and 2019 for the document type 'articles' in the subject area 'Economics, Econometrics and Finance' for the following search term: ('corporate finance' OR 'asset pricing' OR 'financial intermediation' OR 'banking') AND ('empiric*' OR 'data' OR 'sample').

Important reasons for the rapid increase in volume and velocity are the steadily improving availability and quality of economic data, the low cost of computing power to perform statistical analyses, and the fact that academic careers are strongly tied to research output and publications in prestigious journals. The growth in empirical research offers many opportunities to expand our scientific knowledge on a broad and solid foundation, but it also poses challenges.

One of these challenges is the variety of results that we can often observe for the same research question. In many of the major topic areas in finance, we see substantial differences in what researchers report about a specific phenomenon (among many others, Astakhov et al., 2019; Feld et al., 2013; Holderness, 2018). We often find studies on the same research question providing evidence for a significant positive, negative, or no effect for the variable of choice, while controlling for a set of other variables. In addition to the variation in reported effects, finance studies testing the same research phenomenon often vary widely in terms of their variable definitions (selection of dependent and independent variables, data transformations, etc.), data samples (countries, time period, industries, etc.), and the applied methods (statistical estimator, outlier treatment, endogeneity correction, etc.). This variety of reported results and applied methods reflects a critical discourse and thus progress in science. However, at the same time it is becoming increasingly difficult to find out what is really known in a particular research field and to what extent the observable heterogeneity of empirical results correlates with differences in study design, applied methods, and data samples.

Coupled with the challenge of heterogeneity is the dwindling credibility in empirical research. In many areas of science, we see growing skepticism towards single-approach findings and less confidence in the reliability of statistical test results produced under the pressure of reporting significant outcomes (among many others, Aguinis et al., 2017; Brodeur et al., 2016; Chordia et al., 2018; Christensen and Miguel, 2018; Harrison et al., 2017; Harvey et al., 2016; Ioannidis, 2005; Kim and Ji, 2015; Leamer, 1983; Morey and Yadav, 2018; Sterck, 2019; Wasserstein and Lazar, 2016). In his presidential address to the American Finance Association (AFA) in 2017, Campbell Harvey criticized the misinterpretation of *p*-values and the fact that empirical researchers are prone to publish articles with positive findings, which creates incentives for researchers to engage in data mining and *p*-hacking (Harvey, 2017: 1400-1401).[4] Hence, if a data set is large enough and consists of many variables, statistically significant relations among these variables can be detected by sheer coincidence or active searching.

For example, Chordia et al. (2018) implement a data mining approach to generate more than two million trading strategies and test their performance. Applying traditional statistics reveals a large number of rejections of the null hypothesis of no profitability. In contrast, when they correct for *p*-hacking, trading strategies that are outperforming the benchmark are rare. In a similar vein, Kim and Ji (2015) review 161 papers that are published in the top four journals in financial economics. They find that the number of studies with statistically insignificant results is unreasonably low. Among the reviewed studies, 98% report results with statistical significance, which may only make sense if the null hypothesis is false for all studies. The authors consider this result to be a strong indication for publication selection bias in favor of statistical significance (Kim and Ji, 2015: 7). In a similar vein, Mitton (2022) recently investigated the consequences of methodological choices in a review of 495 empirical papers in corporate finance. The review of the literature reveals a wide variation in the applied empirical methodology, the set of control variables, and the methods used to deal with outliers. The author concludes that the wide range of methods provides the researcher with flexibility to match the theory under investigation and the characteristics of the data set with the appropriate methodology (Mitton, 2022: 528). However, the flexibility in methods could also be misused to intentionally produce preferred study results, especially statistically significant results. When given discretion over ten routine methodological decisions (including winsorizing, inclusion of industry dummy variables, or taking the log of the dependent variable), 70% of the randomly generated variables in the reported simulation are statistically significant

[4] I follow the classification by Mitton (2019: 1). Accordingly, data mining refers to the active search for significant results in large data sets across 'multiple hypotheses'. In contrast, *p*-hacking refers to the search for statistically significant results across 'multiple methodologies'.

determinants of corporate leverage at the 5% level. Accordingly, method choices can be exploited to obtain statistically significant results, even for unrelated variables.

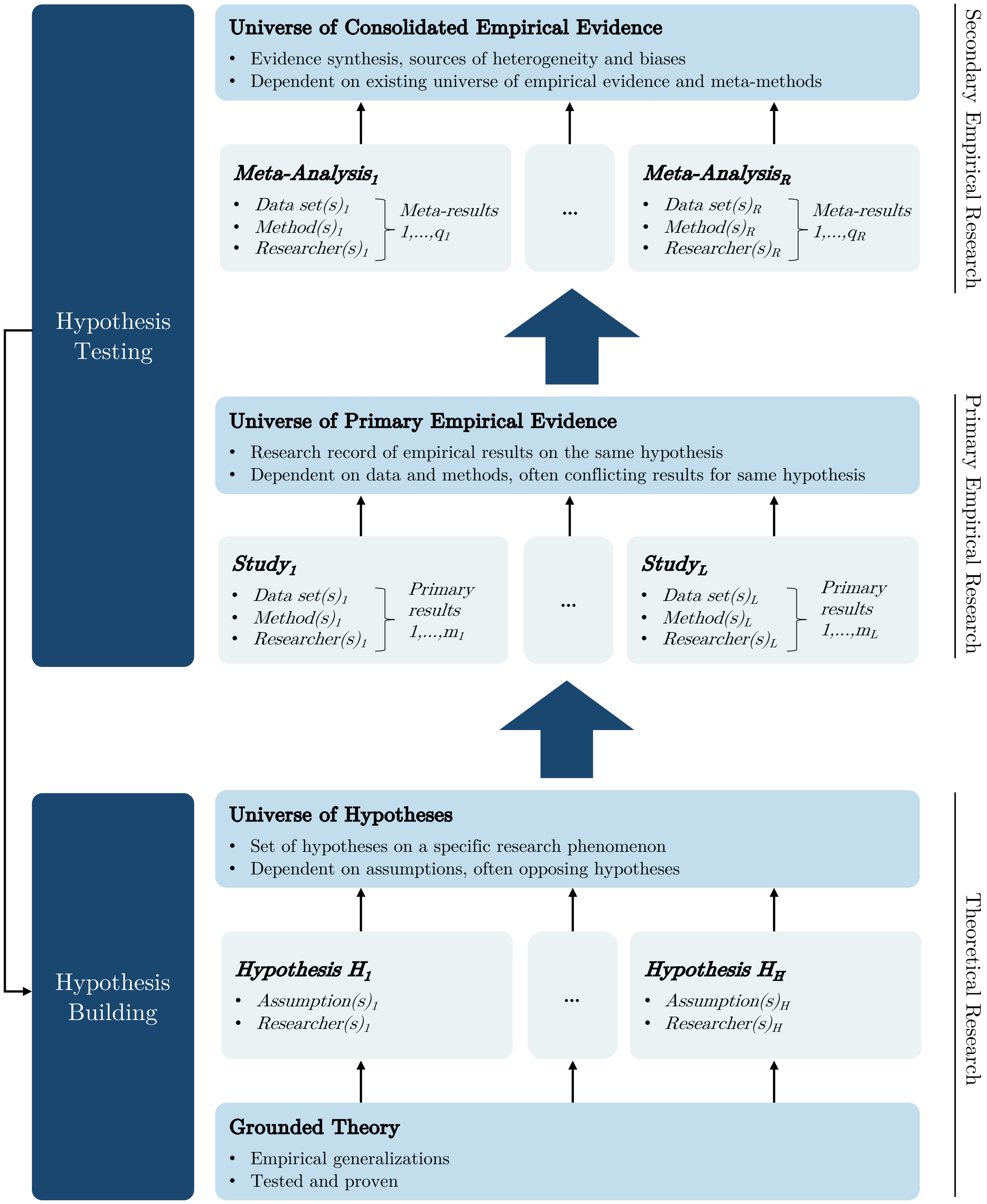
In summary, we observe an ever-growing amount of empirical literature, often with diverse and sample-specific findings. At the same time, scholars and the public might lose confidence in the credibility of empirical research output due to biases and questionable research methods. Therefore, it is increasingly challenging to find out what we really know about a particular phenomenon, what drives the variability of empirical findings, and whether we can rely on the scientific knowledge gained from reported empirical research. Searching for answers to these questions is not only relevant for academics, but also for decision-makers in politics and business. Indeed, the practical relevance of empirical research itself has been called into question.

1.2. Reflecting Empirical Research through Meta-Analysis

To deal with the challenges of an ever-increasing amount of research output, we need a research method that allows us to accumulate, compare, and correct conflicting and potentially biased research findings. A tool that offers these features is meta-analysis. Meta-analysis is a quantitative review technique that uses statistical methods to combine empirical results on the same research question taken from a set of primary studies and draw conclusions about whether and to what extent an effect has received empirical support (Glass, 1976: 3). In other words, meta-analysis is a method of 'big data analysis' that allows the piecemeal findings reported in several individual studies to be summarized and integrated into a bigger picture (Gurevitch et al., 2018: 175). In addition to aggregation, meta-analysis provides methods to detect and filter out preferential reporting of statistically significant results (Stanley, 2008: 108). Thus, meta-analysis has the potential to play an important role in contributing to the transparency and integrity of empirical research. Next to the analysis of publication selection bias, meta-analysis is also a tool to explore heterogeneity among research findings. Finance researchers make many choices in their research process, including the selection of the data set, the data preparation approach, and the estimation methods. Through meta-analysis, we can investigate whether these choices systematically affect the reported results and, if so, we can infer the average effect implied by the method that is assumed to be correct (Stanley and Doucouliagos, 2012: 81-89).

To embed meta-analysis in the universe of research methods, Figure 2 illustrates a simplified deductive research process, based on Mikolajewicz and Komarova (2019: 4) and Cameron and Price (2009: 76), with meta-analysis as the empirical approach for secondary research. Research, if it is hypothesis testing, begins with the hypotheses derived from existing, grounded theory or the results of previous hypothesis tests. Each hypothesis on a specific phenomenon is based on certain assumptions and the idiosyncratic characteristics of the researchers developing the hypothesis based on their predictions and observations. All available hypotheses form the universe of hypotheses, often with contradictory propositions for the same phenomenon.

Figure 2. Deductive Research Process



Notes: This figure shows a deductive scientific research process based on Mikolajewicz and Komarova (2019: 4) and Cameron and Price (2009: 76).

The researchers of empirical primary studies test the universe of (opposing) hypotheses using a specific set of data and methods. The results of the empirical studies are supporting or rejecting the initial hypothesis. All primary empirical results comprise the universe of primary empirical evidence. Often, the empirical results for a specific hypothesis are contradictory and conditional on the applied methods and data.

Meta-analysis, as a secondary empirical research approach, consolidates the scattered universe of empirical evidence into synthesized effects while simultaneously identifying biases in the primary research process and exploring sources of heterogeneity of previous primary studies' findings. Different meta-analyses can yield different results depending on the examined primary data and applied meta-methods. Empirical results from primary and secondary research can then support theory-building by confirming and rejecting earlier hypotheses, but also by suggesting new hypotheses.

Due to its distinct features, meta-analysis has become an established and frequently employed instrument for research synthesis and evidence-based decisions in many research fields (among many others, Combs et al., 2019; Eisend, 2017; Geyskens et al., 2009; Grewal et al., 2018; Gurevitch et al., 2018; Khilf and Chalmers, 2015). There are more than 50,000 meta-analytical studies in medicine (Ioannidis, 2016: 487), more than 2,000 meta-analyses have been conducted in economics (Havranek et al., 2020: 470), and more than 700 meta-analysis studies are published in management science (Buckley et al., 2013: 100). Meta-analyses in these areas are often widely acknowledged and frequently published in leading field journals such as the *Academy of Management Journal* (among others, Carney et al., 2011; Duran et al., 2015; Joshi et al., 2015), *American Economic Review* (among others, Card and Krueger, 1995), *Journal of Management* (among others, Lee et al., 2017; Marano et al., 2016; Shao et al., 2013), *Management Science* (among others, Noel et al., 1990; VanderWerf and Mahon, 1992), and *The Economic Journal* (among others, Card et al., 2010; Görg and Strobel, 2001). Despite its popularity and mainstream acceptance in many academic disciplines, meta-analysis is still rarely used in finance research. However, some recent examples successfully demonstrate its potential value for the field (among others, Feld et al., 2013; Holderness, 2018; Kysucky and Norden, 2016; van Ewijk et al., 2012).

1.3. Objectives and Structure of the Thesis

Given the critical role of meta-analysis in advancing scientific knowledge and the slightly evolving adoption of meta-analysis methods in finance, the central objective of this thesis is to illustrate the application of meta-analysis in finance, to highlight its advantages and challenges, and to investigate the performance of meta-methods under the practical conditions typically encountered by meta-researchers in finance. Through this contribution, I aim to have a positive impact on the awareness and acceptance of meta-analysis as a research method in finance and promote its future application in the field.

The thesis is divided into five chapters, which are shown in Figure 3.

Figure 3. Structure of the Thesis

Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5
Motivation for meta-analysis in finance	Introduction of meta-methods and review of meta-analysis research in finance	Application of meta-analysis on the corporate hedging and firm value nexus	Comparison of meta-regression weights by Monte Carlo simulation	Summary and conclusions

Notes: This figure illustrates the structure of the thesis.

Chapter 2 introduces the methods of traditional meta-analysis, meta-regression analysis, and meta-analytic structural equation modeling. This is followed by a structured literature review of 76 existing meta-analysis studies in finance. From the literature review, I derive the key opportunities and challenges of meta-analysis in finance. Although similar introductions of meta-analysis for specific research areas are available in other domains, such as management, marketing, and economics (among others, Anderson and Kichkha, 2017; Eisend, 2017; Geyskens et al., 2009), to the best of my knowledge there has been no attempt to conceptually introduce meta-analysis and its methods to the finance research field.

Chapter 3 illustrates the application of the meta-methods introduced in the previous chapter. To this end, meta-analysis is used to study the impact of corporate financial hedging on firm value. Corporate hedging behavior, along with its impact on firm valuations, has attracted widespread attention in the corporate finance literature. Despite a large body of empirical evidence investigating firm-level data to determine whether corporate hedging is a value-enhancing strategy, the literature remains largely unsettled on the overall effect of hedging on firm value. Previous empirical studies show mixed effects ranging from firm value discounts to large value premiums of hedging firms, while little is known about the drivers of the large heterogeneity in the existing research findings. In this chapter, meta-regression analysis is applied to aggregate and compare 1,016 hand-collected findings reported in 71 primary studies. Testing of publication selection bias and heterogeneity analysis are used to analyze distorting effects from the selective reporting of preferred results and to explore the data and method choices that determine the variability in reported firm value effects of corporate hedging.

Chapter 4 examines practical challenges related to the weighting schemes applied in meta-regression when meta-researchers collect not only one but many effect size estimates from the same study. Multiple estimates reported per study are rather common in economics and finance research, where studies often report many estimates for the same effect under examination using alternative models, robustness checks, or sensitivity analyses. A Monte Carlo simulation is designed in the way that it creates a synthetic panel data set of primary studies with varying numbers of reported effect size estimates per study. In this setup, I evaluate the statistical properties of inverse variance weighting in meta-regression as compared to two alternative weighting schemes that consider the number of reported effect size estimates in the weights. The simulation allows determining how the performance of the inverse variance weighting and the two alternative weights changes when the unbalancedness of the panel data increases, as some primary studies report a large number of effect size estimates, when the sample overlap between effect size estimates from the same study becomes larger, and when studies actively select statistically significant outcomes.

Lastly, Chapter 5 presents a summary of the key findings for each chapter and concludes the thesis with an outlook.

Chapter 2. Meta-Analysis for Transparency in Finance Research[5]

"What we need is some objective and critical methodology to integrate conflicting research findings and to reveal the nuggets of 'truth' that have settled to the bottom."

(Stanley and Doucouliagos, 2012: 2)

This chapter introduces the concepts and methods of meta-analysis (Sections 2.1 and 2.2). It reviews the current state of the art for meta-research in finance (Section 2.3), derives opportunities offered by meta-analysis (Section 2.4), and discusses related challenges (Section 2.5). The chapter ends with a summary (Section 2.6). The meta-methods introduced in this chapter are applied in Chapter 3 and are investigated in more detail in Chapter 4.

2.1. Introduction to Meta-Analysis

2.1.1. Terms and Definitions

The term 'meta-analysis' needs to be defined in the larger context of secondary research methods. Traditionally, there are two common approaches to reviewing the primary literature in a research field: narrative reviews and systematic reviews (Pae, 2015: 418). Narrative reviews are often conducted by domain experts as a qualitative summary of the literature on a (broader) research question, along with interpretations and criticisms (Garg et al., 2008: 253). However, the interpretations in narrative reviews usually depend on the author's judgment and subjective choices about which studies were given more emphasis than others (Akobeng, 2005: 845). In contrast to narrative reviews, systematic reviews are considered to be more rigorous and less prone to subjective judgements (Booth et al., 2016: 2). Systematic reviews follow a predetermined and reproducible procedure to search, screen, select, appraise, and summarize all existing studies that have addressed the same specific and narrowly defined research question (Greenhalgh et al., 2018: 1). Systematic reviews can be qualitative or quantitative, depending on the type of data collected from the studies. If studies report comparable statistical results, a systematic review can be complemented by meta-analysis (Booth et al., 2016: 186). Accordingly, quantitative systematic reviews are usually meta-analyses. However, not every meta-analysis is a systematic review unless it follows its strict and systematic procedure.

[5] Parts of this section are published in the paper *Meta-Analysis in Finance Research: Opportunities, Challenges, and Contemporary Applications*, International Review of Financial Analysis 71(101524), 2020, co-authored by Markus Hang and Andreas Rathgeber.

According to Glass (1976: 3), a meta-analysis is the *"statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings.*" Meta-analysis originates from medical research, where studies are often based on small samples because large clinical trials are both costly and time-consuming (Leandro, 2005: 3-4). When a clinical study has a small sample size, which entails low statistical power, it might fail to find statistically significant relations due to large standard errors of the estimates (DerSimonian and Laird, 1986: 177). The application of meta-analysis to synthesize results from a set of small-sample studies increases the precision of the estimated effects and hence has the power to uncover important findings that could not have been found in a single study based on a few subjects (Leandro, 2005: 6).

Different research areas adopted meta-analysis in different ways and adjusted its methods and concepts to their discipline. For example, the management literature is largely focused on the aggregation of correlational data reported in previous empirical studies and the testing of structural models at the meta level (Geyskens et al., 2009: 399). In contrast, the economics literature focuses on explaining the sources of heterogeneity in prior regression results along with identifying and correcting for frequent biases (Stanley and Doucouliagos, 2012: 3-4). Meta-analysis in finance research is a new discipline, which is inspired by meta-analytic applications in management and economics. To ensure a common understanding of the terms and definitions used throughout this thesis, Table 1 provides an overview of the meta-analytical terminology.

Table 1. Glossary of Meta-Analysis Terms

Term	Description
Effect Size	A quantitative measure of the magnitude of a phenomenon that is used for the purpose of addressing a specific research question (Kelley and Preacher, 2012: 140).
Heterogeneity	Variability in study outcomes between and within studies (Nagendrababu et al., 2020: 238).
Misspecification Bias	Bias arising from the inclusion of incorrect variables, the omission of important regressors, measurement errors, or incorrect functional forms (Asteriou and Hall, 2011: 180).
Moderator Variable	A third variable that conditions the direction and/or strength of the relation between two variables (Baron and Kenny, 1986: 1174).
Primary Study	An original empirical study reporting effect size estimates for the phenomenon under examination (Stanley and Doucouliagos, 2012: 13).
Publication Selection Bias	Bias arising when the accessible literature and its results are systematically unrepresentative of the population of all the research results that have been produced in a research area (Rothstein et al., 2005: 1).

Notes: This table defines the key terms related to meta-analysis that are used throughout this thesis.

Effect Size. Effect sizes are the unit of analysis to be aggregated and compared in meta-analysis (Nagendrababu et al., 2020: 235). Since the true population effect size, α, is usually unknown, primary studies commonly report estimates for the population effect size, $\hat{\alpha}$. The measures used as effect sizes differ across research areas. For business and economics research, Stanley and Doucouliagos (2012: 22) distinguish between statistical and economic effect size measures. 'Statistical' effect sizes are unitless measures of the relationship between two variables, such as (partial) correlation coefficients. 'Economic' effect sizes capture the effect of economic interest, such as elasticities or other measures that quantify the percentage change in a dependent variable. The standard error of the effect sizes, $SE(\hat{\alpha})$, indicates the uncertainty that is included in the effect size measurement.

Heterogeneity. To be aggregated in a meta-analysis, the effect size estimates should be observed from a group of similar studies (Borenstein et al., 2009: 18). However, the effects reported in empirical research often vary. This variation, also known as sampling error, is naturally inherent when researchers take samples instead of analyzing the true population (Stanley and Doucouliagos, 2012: 80-81). If the variation of the effect size estimates observed in a specific sample is greater than expected due to sampling error, the sample is heterogenous (Higgins and Thompson, 2002: 1539). There are different sources of heterogeneity. Meta-analysts commonly distinguish two groups of heterogeneity: 'structural' heterogeneity and 'methodological and data-related' heterogeneity (Havranek and Irsova, 2011: 238; Rusnak et al., 2013: 52). Structural heterogeneity results from real differences among the effect size estimates, for example, when two studies find systematically different results for companies from different industries, countries, or geographical regions. Methodological and data-related heterogeneity arises when effect size estimates vary due to differences in the methodology, model specifications, and data preparation procedures applied in the primary studies.

Misspecification Bias. Research in economics and finance is mostly observational rather than experimental. While in experimental sciences, sampling into treatment and control groups often overcomes the need for control variables, researchers analyzing observational data cannot directly control for experimental variation (Koetse et al., 2005: 2). Therefore, econometrics attempts to achieve statistical control by including other variables besides the key factors under study that are assumed to influence the dependent variable. However, there might be more variables that cannot be considered, for example, because the data does not exist. Thus, bias caused by omitted variables is a common risk to empirical models. Model misspecification is a broader term that refers not only to the problem of omitted variables, but also to incorrect functional forms of econometric models or measurement errors (Asteriou and Hall, 2011: 181-191). Misspecification bias occurs when misspecifications affect the results of a study in such a way that they represent a distorted sample of the true population effect.

Moderator Variable. A moderator variable is a third variable that impacts the direction and/or the strength of the relationship between two variables (Baron and Kenny, 1986: 1174). In other words, moderators can be used to find out whether a particular relation between two variable is constant or sensitive to contextual factors. Moderator variables can be categorical or continuous. In meta-analysis, moderators are usually characteristics of primary studies that are correlated with the effect size estimates in the primary studies (Hall and Rosenthal, 1991: 438). For example, if the impact of oil price changes on stock returns is significantly smaller in published journal articles than in unpublished working papers, then the publication status of a study is a moderator variable for the impact of oil on stocks.

Primary Study. An empirical primary study is the original analysis of data (Glass, 1976: 3). Primary studies might report one or several effect size estimates. A meta-analysis is the re-analysis of the effect size estimates observed from a series of primary studies. The term primary study is used independent of the publication outlet and therefore covers published journal articles, book chapters, working papers, conference proceedings, as well as unpublished empirical results. The latter group of non-journal articles is sometimes referred to as 'grey literature' (Borenstein et al., 2009: 279).

Publication Selection Bias. Publication selection bias arises when researchers or editors and reviewers discard undesired results from publication, also called 'publication selection' or the 'file drawer problem' (Begg and Berlin, 1988: 419; Rothstein et al., 2005: 1; Stanley, 2005: 309-310). Undesirable outcomes might be statistically insignificant effects, outcomes without support of the ex-ante hypothesis, outcomes that are inconsistent with theoretical predictions, or results that do not agree with what was found in the previous literature. If such 'negative' results end up in the file drawer, the sample of available empirical results does not correctly reflect the population effect. If uncontrolled, the active selection of preferred statistical results can bias the accumulated effects in a meta-analysis (Doucouliagos and Stanley, 2013: 320-321).

2.1.2. Steps for Conducting Meta-Analysis

Meta-analysis comprises several steps, which are depicted in Figure 4. The standard meta-analysis approach in economics and finance can be divided in two main parts: data search and data preparation (Steps 1–7), and the meta-evaluation (Steps 8–12).

Figure 4. Meta-Analysis Flowchart

Data Search and Data Preparation

1. Definition of the research question(s) and hypotheses
2. Definition of criteria for the inclusion and exclusion of primary studies
3. Literature search in electronic databases, working paper repositories, and reference lists
4. Filtering the sample of relevant primary studies against inclusion and exclusion criteria
5. Definition of statistical effects and study characteristics to be collected
6. Manual coding of primary studies by extracting the relevant information
7. Data preparation including the calculation of effect sizes and definition of moderators

Meta-Evaluation

8. Calculation of descriptive statistics and aggregation of findings
9. Analysis and correction of publication selection bias
10. Analysis of heterogeneity and misspecification bias
11. Investigation of robustness tests and sensitivity analyses
12. Reporting and interpretation of the findings

Notes: This figure illustrates the steps of literature search, data preparation, and statistical analysis to be conducted in a meta-analysis in economics and finance.

Depending on the specific research area and the applied meta-methods, these steps may be conducted in a different order, steps may be omitted, or additional steps may be added. Chapter 3 of the thesis presents a comprehensive application of meta-analysis in corporate finance, including all steps outlined in Figure 4.

Data Search and Data Preparation. Before starting a meta-analysis, the researcher defines the research question(s) and hypotheses to be addressed. Some scholars even promote the pre-registration of meta-studies[6] to reduce the risk of data mining and tendentious reporting of results (Christensen and Miguel, 2018: 953; Quintana, 2015: 2). In the next step, the data search starts with the definition of the selection criteria, which could be inclusion or exclusion criteria. This is typically a list of objective conditions allowing the meta-researcher to decide whether to include or exclude a primary study from the meta-sample (Nagendrababu et al., 2020: 235). The selection criteria ensure that the primary studies included in the sample are

[6] The pre-registration information should include the initial meta-analysis protocol (Quintana, 2015: 2), which, according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, consists of details such as the study rationale, study eligibility criteria, search strategy, moderator variables, risk of bias, and the statistical approach for the meta-analysis (Shamseer et al., 2005: 24).

similar in the sense that their differences can be coded by the meta-analyst (Stanley and Doucouliagos, 2012: 13).

The definition of the selection criteria is followed by the collection of the population or a representative sample of the literature that examines the phenomenon under investigation. The literature search usually encompasses a thorough review of electronic databases for research articles (like EBSCOhost, Google Scholar, Proquest's ABI/INFORM, or the Social Science Research Network working paper database), a backward and forward search through the papers' references lists, and contacting study authors to gain access to recent or unpublished results. To obtain the information required for the application of meta-analysis methods, the meta-analyst extracts the relevant data by manually 'coding' the primary studies identified in the literature search. This step is typically the most time-consuming part of a meta-analysis and often demands several months of reading papers and extracting information.

For the literature search and the coding, it is essential to maintain high quality standards to avoid systematic bias or errors in the meta-data set. Recommended processes and best practices for the data search and coding are well documented in many handbooks and guidelines (Havranek et al., 2020; Shamseer et al., 2005; Stanley et al., 2013; Stanley and Doucouliagos, 2012: 12-37) and therefore are not discussed further in this thesis.

Meta-Evaluation. When the data is extracted from the primary studies and processed, meta-methods can be applied to combine the previous statistical results into an overall effect and to compare the summarized effects for different subsamples. In this step, it is important to test and correct for biasing effects caused, for example, by selective reporting of statistically significant results or the model misspecification in primary studies. Meta-analysis provides various graphical and analytical tools for the detection of biases (Stanley and Doucouliagos, 2010).

Another step in the meta-evaluation stage is the investigation of the influencing factors that are responsible for the structural as well as the methodological and data-related heterogeneity among the empirical results between and within studies. The calculation of summary effects, the investigation of publication selection, and the exploration of the sources of heterogeneity are typically the main goals of meta-analyses in economics and finance. Robustness and sensitivity analyses ensure the stability of the meta-results against different methodological choices and changes in the data set.

Several guidelines and best practices support the interpretation of the meta-results (among others, Doucouliagos, 2011; Doucouliagos and Stanley, 2013). A more detailed explanation of the different meta-methods follows in the next sections.

2.2. Meta-Analysis Models

Anello and Fleiss (1995) provide a general classification of meta-analysis models depending on the primary goal of the analysis. They distinguish between 'analytic' and 'exploratory' meta-models. The goal of analytic meta-analysis is to combine the statistical results from a series of primary studies into a pooled effect, either separately for each phenomenon of interest (univariate model) or simultaneously if more than one phenomenon is of interest (multivariate model) (Anello and Fleiss, 1995: 111-113). In contrast, the goal of an exploratory meta-analysis, also referred to as moderator or heterogeneity analysis, is to explore why effect sizes vary from study to study or within the same study (Anello and Fleiss, 1995: 113). There are three contemporary approaches to conducting analytic and exploratory meta-analysis in economics and finance: (i) traditional meta-analysis, (ii) meta-regression analysis, and (iii) meta-analytic structural equation modeling. The three approaches are summarized in Table 2 and are further explained in the following subsections. The approaches are not mutually exclusive, but meta-analysts often apply several methods in the same study.

Table 2. Contemporary Approaches to Meta-Analysis in Economics and finance

	Traditional Meta-Analysis	Meta-Regression Analysis	Meta-Analytic Structural Equation Modeling
Approach	Uses weighted averages to combine effect size estimates obtained from a set of individual studies.	Uses regression analysis to explore a set of moderator variables (independent variables) as drivers of the heterogeneity observed in the effect size estimates (dependent variable) across a set of individual studies.	Uses structural models and path analysis to explain the simultaneous relation among a group of variables obtained from a set of individual studies.
Main Purpose	Integrate effect size estimates from multiple studies into a pooled effect to draw overall conclusions.	Identify moderator variables as sources of effect size variation and correct biases from publication selection and model misspecification.	Test simultaneous theoretical relationships/models including mediating factors.
Common Effect Size	Pearson correlations, mean differences	Partial correlations, elasticities	Pearson correlations
Main Opportunity	Provide an estimate for the population effect across the literature.	Reveal why effect size estimates vary and detect biases affecting credibility of empirical research results.	Allow for the testing of complex structural models based on the results of many studies.
Main Challenge	Accounting for (complex) heterogeneity and publication selection bias.	Accounting for multi-dimensional dependence structure in the meta-sample.	Accounting for (complex) heterogeneity and publication selection bias.
Guidelines	Borenstein et al. (2010) Hunter and Schmidt (2004)	Havranek et al. (2020) Stanley et al. (2013)	Bergh et al. (2016) Jak (2015)

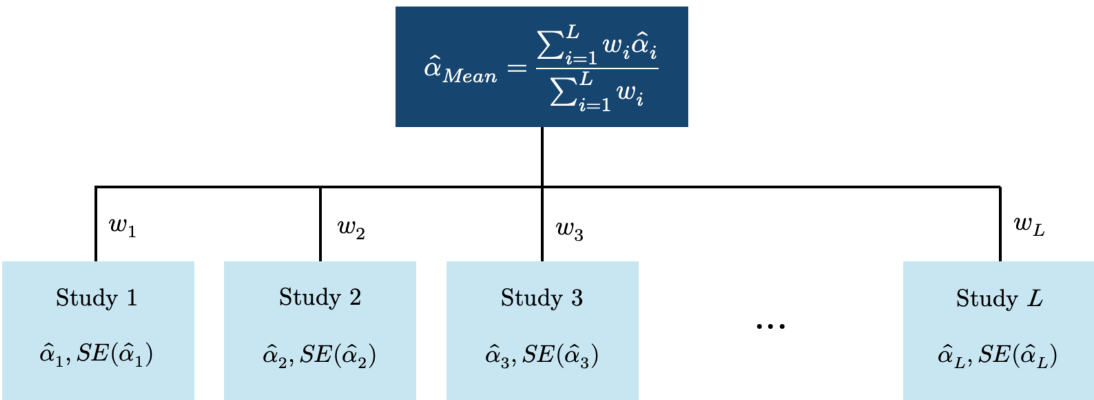
Notes: This table provides an overview of the main meta-methods applied in economics and finance research.

2.2.1. Traditional Meta-Analysis

Traditional meta-analysis methods focus on the aggregation of statistical effect sizes. To estimate the true population effect, α, meta-analysts calculate weighted averages of the effect size estimates, $\hat{\alpha}$, observed from a sample of primary studies (Figure 5). There are several methodological approaches for traditional meta-analysis. One of the most popular methods is the Hedges-Olkin meta-analysis (Hedges and Olkin, 1985), which is discussed here.

Given is a set of $i = 1, \dots, L$ empirical primary studies investigating the same phenomenon or hypothesis, for example, the abnormal return effects after announcements of equity issuances by publicly traded firms. Each study reports one[7] effect size estimate, $\hat{\alpha}_i$, along with the corresponding standard error of the effect size estimate, $SE(\hat{\alpha}_i)$, or statistics to re-calculate the standard error (e.g., t-statistics or the number of observations used for the estimation). Traditional meta-analysis aims at deriving the best estimate for the unknown population effect size, α, by calculating the weighted mean effect size estimate, $\hat{\alpha}_{Mean}$.

Figure 5. Illustration of the Traditional Meta-Analysis Method



Notes: This figure illustrates the aggregation of results from a set of individual studies by calculating weighted mean effects.

A key parameter in traditional meta-analysis is the weight w_i assigned to each effect size estimate. There are two models for traditional meta-analysis that use different weights to compute the mean effect size estimate: the fixed effects meta-analysis model (FEM) and the random effects model (REM). The FEM assumes that each study reports an estimate for the same underlying population effect, i.e., the true effect is the same in all studies (Borenstein et al., 2009: 61). Accordingly, the variation of the effect size estimates across studies is attributed to the random sampling error (Hedges and Vevea, 1998: 486). Thus, the FEM assumes that there is no additional variation beyond sampling error, i.e., there is no heterogeneity.

[7] Model extensions of the traditional approach exist that account for multiple estimates reported in the same study (Card, 2012: 279-305).

Hedges and Olkin (1985: 110) show that for the FEM, the optimal study weights, w_i, are given by the inverse of the effect size estimate's squared standard error, $SE(\hat{\alpha}_i)^2$:

$$w_i^{\text{FEM}} = \frac{1}{SE(\hat{\alpha}_i)^2}. \tag{1}$$

This weighting scheme indicates that more precise effect size estimates, which are those with lower standard errors, receive larger weights and hence have a greater impact on the weighted mean effect.

In the case of real differences between studies, the effect size estimates will be different even if all studies would have an infinitely large sample size (Riley et al., 2011: 965). For example, regional differences in governance structures could lead to fundamentally different reactions of shareholders to equity issuance announcements in different countries. In the case of heterogeneous effect sizes, the FEM would be an incorrect meta-analysis model.

Genuine heterogeneity in the sample of effect size estimates causes an additional source of variation beyond sampling error, which must be taken into account in the weighting scheme (Borenstein et al., 2009: 73). In contrast to the FEM, the REM assumes that effect size estimates are drawn from study-specific populations. In the REM, the meta-analytic weights take heterogeneity into account by splitting the variation of effect size estimates into within-study variation, $SE(\hat{\alpha}_i)^2$ (capturing the sampling error), and an estimate of the between-study variation, $\hat{\tau}_{\mathrm{i}}^2$ (capturing the variance of the effect size estimates between study-specific populations). In the REM, study weights are assigned with the goal of minimizing both components of the variance:

$$w_i^{\text{REM}} = \frac{1}{SE(\hat{\alpha}_i)^2 + \hat{\tau}_{\mathrm{i}}^2}, \tag{2}$$

where the between-study variation, τ_i^2, is usually unknown, but can be estimated by a method of moments (DerSimonian and Laird, 1986) or (restricted) maximum likelihood estimator (Viechtenbauer, 2005) to receive the estimate $\hat{\tau}_{\mathrm{i}}^2$.

In addition to inverse variance weighting, traditional meta-analysis also applies alternative weights. For example, Stanley and Doucouliagos (2015) show that in the case of publication selection bias an unrestricted weighted least squares (WLS) estimator is superior to the conventional REM and better than FEM if there is heterogeneity. Moreover, standard errors are sometimes not reported in the primary studies and re-calculations from other reported statistics might be unreliable. In this case, meta-analysts may also use the primary study's sample size, journal impact factors, or the number of the study's citations as quality indicators to assign larger weights to 'better' studies (Stanley and Doucouliagos, 2012: 46).

After selecting and calculating appropriate weights, the estimated value for the population effect size is the weighted mean of the effect size estimates, $\hat{\alpha}_i$:

$$\hat{\alpha}_{Mean} = \frac{\sum_{i=1}^{L} w_i \hat{\alpha}_i}{\sum_{i=1}^{L} w_i}, \tag{3}$$

where $\hat{\alpha}_{Mean}$ and corresponding confidence intervals (CIs) provide an estimate for the overall effect implied by the literature.

Next to the traditional meta-analysis approach by Hedges and Olkin (1985), there are alternative approaches to summarize effect size estimates across studies. For example, the artifact-correcting meta-analysis by Hunter and Schmidt (1990) has been widely used in management sciences. This approach corrects effect size estimates for a variety of statistical artifacts[8] and biases (e.g., due to measurement errors or computational errors). However, in contrast to self-reported psychometric data, studies in economics and finance are typically based on independently verifiable economic data. Accordingly, the nature of the data does not necessarily require artifact correction (Heugens et al., 2009: 493). As opposed to the inverse variance of the effect size variance, as applied in weighting by Hedges and Olkin (1985), Hunter and Schmidt (1990) use the number of observations in each primary study sample as weights. Additional differences between the traditional meta-analysis approaches are discussed in more detail by Johnson et al. (1995) and Hunter and Schmidt (2004).

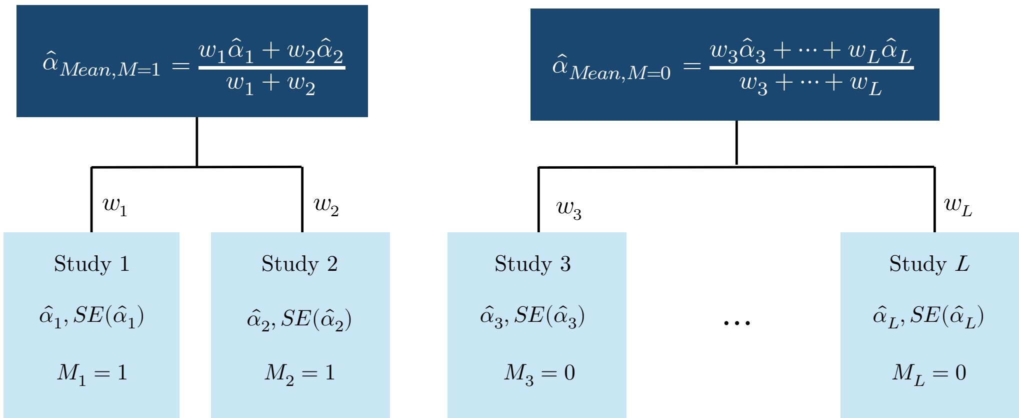
In addition to the computation of an overall effect, traditional meta-analysis is also used to analyze the combined effect sizes of different 'subgroups' of studies. In subgroup analysis, the sample is divided into two or more separate groups. The split into the subgroups is based on a specific criterion or moderator variable M. Studies that share a common characteristic are assigned to the same subgroup (e.g., all studies that examine U.S. firm data compared to all studies that examine data from other countries). The approach for subgroup meta-analysis is illustrated in Figure 6.

After splitting the sample of primary studies into subgroups, the mean effect size of each subgroup can be calculated as previously described in Eqs. (1), (2) and (3). The variance of the difference between the subgroup effects is used to compute CIs and significance tests to analyze whether differences between subgroups, and hence heterogeneity, are significant. If heterogeneity is present in the sample, the variation within the subgroups should decrease when there is significant variation between subgroups (Rosenbusch et al., 2013: 346).

[8] Hunter and Schmidt (2004: 33) define artifacts as *"study imperfections [...] to remind ourselves that errors in study results produced by study imperfections are artifactual or man-made errors and not properties of nature."* Thus, artifacts drive the reported effect size estimates away from what the actual effect would be in the case that the study would have been conducted perfectly, i.e., without errors.

Accordingly, the analysis of the significance of the differences between subgroups is a test of whether the criterion used to split the subgroups is a moderator for heterogeneity among the effect size estimates (Shadish and Sweeney, 1991: 884).

Figure 6. Illustration of the Traditional Meta-Analysis Method with Subgroups



Notes: This figure illustrates subgroup meta-analysis by calculating weighted mean effects for subgroups of studies sharing a common study characteristic M.

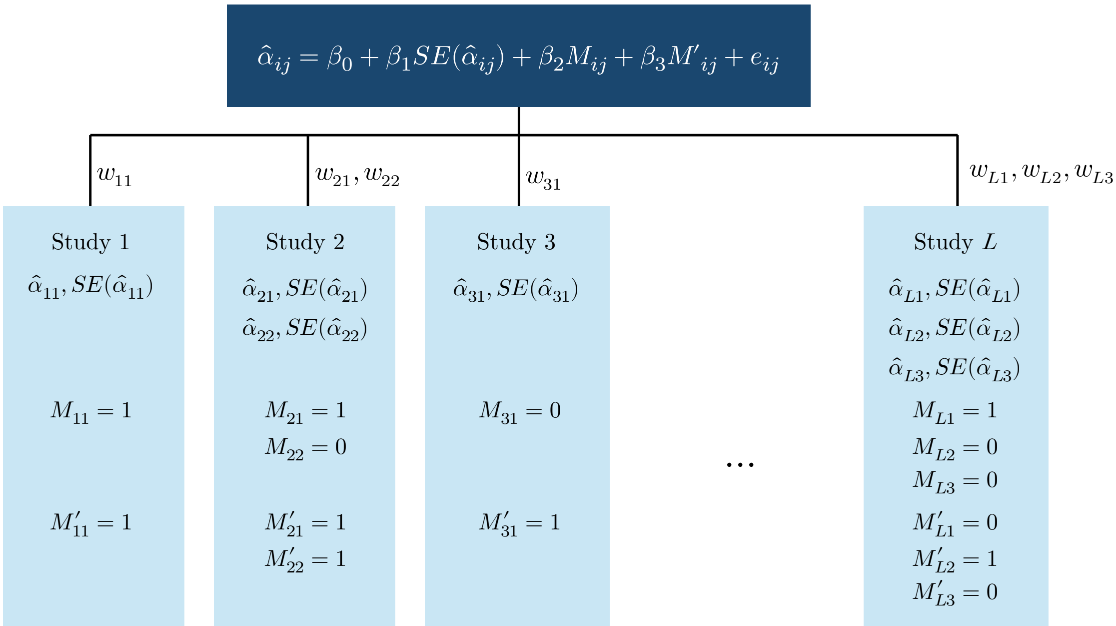
A decisive disadvantage of subgroup analysis is that it cannot properly estimate the simultaneous effect of different heterogeneity drivers while also correcting for publication selection. In economics and finance research, however, we commonly find selective reporting of preferable results along with empirical results that are driven not by a single study feature, but rather by a number of different observable and unobservable sources of heterogeneity. The next section proceeds with meta-regression analysis, which explicitly models the joint impact of many drivers of heterogeneity and sources of biases in a regression-based framework.

2.2.2. Meta-Regression Analysis

In contrast to medicine and natural sciences that use clinical trials or natural experiments, research studies in economics and finance are commonly based on observational data using pre-existing databases. Another distinct feature is the large variation in research designs and data that we commonly find in these research fields, as opposed to the often applied 'gold standard' of randomized controlled trials in medical sciences (Hariton and Locascio, 2018: 1716). Meta-regression analysis is a set of meta-analysis methods that have been developed to consider these specific characteristics of observational research in economics and finance.

Meta-regression is defined as *"the regression analysis of regression analyses"* (Stanley and Jarrell, 1989: 161). In an empirical research environment where studies routinely report regression parameters, meta-regression provides a method for integrating and explaining diverse findings obtained from a set of prior regression models. Meta-regression is also a tool that explicitly accounts for heterogeneity as well as frequent biases, especially selective reporting of statistically significant results and model misspecification. The general concept of meta-regression with two dummy moderator variables, M and M', is illustrated in Figure 7.

Figure 7. Illustration of Meta-Regression Analysis



Notes: This figure illustrates the meta-regression method with two moderator variables. Meta-regression uses the effect size estimates from a series of primary studies as the dependent variable while the standard error of the effect size estimates and study-level moderator variables M and M' serve as independent variables.

Given is a set of $i = 1, \dots, L$ empirical primary studies investigating the same phenomenon or research hypothesis. Each study reports one or several estimates $j = 1, \dots, m_i$ obtained from a regression model:

$$\boldsymbol{z} = \boldsymbol{X\alpha} + \boldsymbol{u}, \tag{4}$$

where $\boldsymbol{z}$ is the dependent variable in the primary study (e.g., the corporate debt financing), $\boldsymbol{X}$ is a matrix of explanatory variables (e.g., the determinants of corporate debt financing), and $\boldsymbol{u}$ represents the random error term. Eq. (4) shows the simplest form of a primary regression, which can be extended by fixed effects, instrumental variables (IV), or other advanced regression methods. α_1 is an element of the vector of regression coefficients, $\boldsymbol{\alpha}$. This coefficient measures the magnitude of a certain effect under examination (e.g., the impact of tax rates on corporate debt financing). A collection of L primary studies produces a sample of estimates, $\hat{\alpha}_{1,i}$, which are denoted y_{ij}. The index j accounts for the fact that empirical studies in economics and finance routinely report more than one regression parameter, e.g., for different robustness analyses and subsample tests.

Since regression coefficients are not always comparable across studies, for example due to different scaling of variables or different functional forms, y_{ij} usually represents transformations of regression parameters, such as elasticities, semi-elasticities, or partial correlation coefficients Stanley and Doucouliagos (2012: 22-29).[9] In addition to the regression parameters, the standard errors of the regression estimates, $SE(y_{ij})$, must be given in the primary studies or, alternatively, the meta-researcher must be able to recalculate them from other reported statistics. Moreover, the primary studies reporting the regression estimates are often different in terms of their study design and sample characteristics. Such study differences are coded by a set of moderator variables denoted by the vectors $\boldsymbol{M}$ and $\boldsymbol{M'}$.

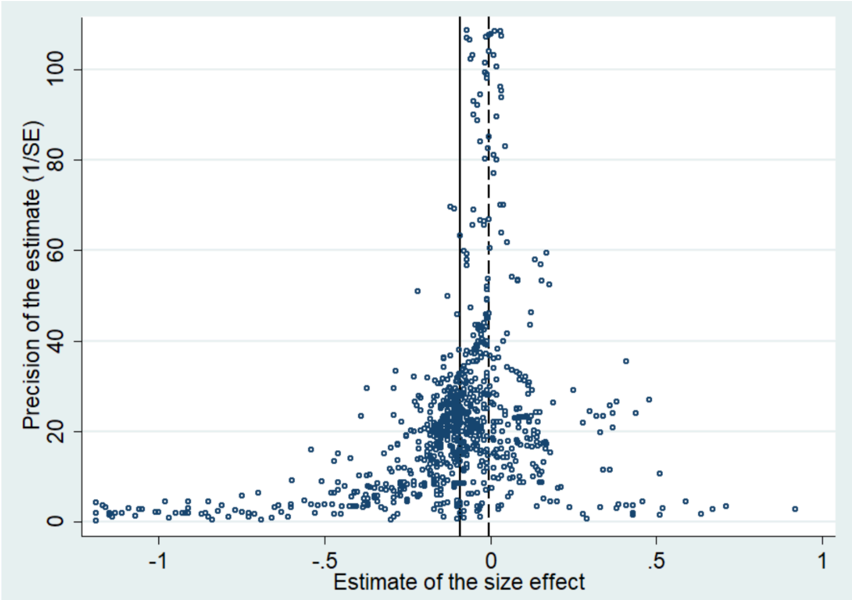
In the meta-regression model, the effect size estimates, y_{ij}, are regressed on a set of explanatory variables, usually $SE(y_{ij})$, M_{ij} and M'_{ij}, which quantify heterogeneity and common sources of bias. This goes much further than averaging effect sizes in the traditional meta-analysis approach, as meta-regression simultaneously models the impact of many variables in a multiple regression framework. A meta-regression model can be defined as follows:

$$y_{ij} = \beta_0 + \beta_1 SE(y_{ij}) + \sum_{l_1=2}^{P_1+1} \beta_{l_1} M_{l_1,ij} + \sum_{l_2=\mathrm{P}_1+2}^{P_1+P_2+1} \beta_{l_2} M'_{l_2,ij} + e_{ij}, \qquad (5)$$
$$\text{with } e_{ij} \sim N(0, SE(y_{ij})^2).$$

Eq. (5) accounts for publication selection by including $SE(y_{ij})$ as an explanatory variable in the meta-regression. In the absence of publication selection, the observed effect sizes, y_{ij}, and their standard errors, $SE(y_{ij})$, should be independent quantities. However, if primary study authors actively change their model specification or data sets until they find estimates that are large enough to offset high standard errors, correlation between the effect size estimates and their standard errors occurs (Stanley et al., 2008: 280). The relation between the primary studies' regression estimates and their (inverse) standard errors is often depicted in a scatter diagram called the 'funnel plot'. Without publication selection, this graph should resemble a symmetrical distribution around the most precise estimates (those with the lowest standard errors), forming an inverted funnel. In the case of publication selection, the graph will be truncated.[10] Figure 8 presents an exemplary funnel plot reported in the meta-analysis by Astakhov et al. (2019: 1473), who examine 102 primary studies reporting estimates for the effects of firm size on stock returns. The funnel plot appears asymmetric with many estimates concentrated in the left tail, suggesting publication selection towards negative firm size effects.

[9] See Stanley and Doucouliagos (2012: 22-29) for an overview of commonly applied effect sizes in economics.
[10] Asymmetry in the funnel graph might also be driven by methodological or structural heterogeneity.

Figure 8. Funnel Plot of 102 Studies Reporting Firm Size Effects on Stock Returns



Notes: This graph shows the scatter plot of effect size estimates for the relation between firm size and stock returns as reported in the meta-analysis by Astakhov et al. (2019: 1473).

According to Egger et al. (1997: 632), testing H_0: $\beta_1 = 0$ in Eq. (5) is a test for the truncation in the funnel plot. The corresponding regression estimate, $\hat{\beta}_1$, measures the direction and magnitude of publication selection. The estimated value for the intercept, $\hat{\beta}_0$, is the mean effect size across all studies when $SE(y_{ij})$ is close to zero ($SE(y_{ij}) \rightarrow 0, E(y_{ij}) \rightarrow \beta_0$), conditional on the assumption that all other moderator variables are zero (Havranek et al., 2015b: 402). Thus, rejecting the null hypothesis, H_0: $\beta_0 = 0$, is a test for the existence of a genuine effect beyond publication selection conditional on the values for $M_{l_1,ij}$ and $M'_{l_2,ij}$ (Stanley et al., 2008: 281). Meta-analysts in economics and finance usually estimate Eq. (5) without $M_{l_1,ij}$ and $M'_{l_2,ij}$ in a first stage to identify publication selection and to estimate the corrected mean effect beyond bias (Stanley, 2008: 108). Then, in a second stage of the analysis, the other meta-explanatory variables are added to the first stage model.

Meta-regression also provides a means by which to quantify the sensitivity of primary results to variations in model specification (Stanley and Jarrell, 1989: 165). Therefore, meta-regression studies typically include a set of $l_1 = 2, \ldots, P_1 + 1$ dummy variables, denoted $M_{l_1,ij}$ in Eq. (5), which indicate whether a specific control variable is present or absent in the original regression model or whether a specific functional form was used (e.g., level-level, log-log, or log-level). The estimated meta-regression coefficients, $\hat{\beta}_{l_1}$, measure the sensitivity of the examined effect to changes in the control variables or functional form of the primary regression model.

Moreover, the moderator variables $M'_{l_2,ij}$ represent a set of meta-independent variables that quantify relevant study characteristics and explain variation across the collected results related to differences in data and method choices. The estimated meta-coefficients, $\hat{\beta}_{l_2}$, reflect the average effect of a given study characteristic on the effect sizes. Accordingly, the explanatory variables in the meta-regression are interpreted as moderators of the effect size estimates, y_{ij}.

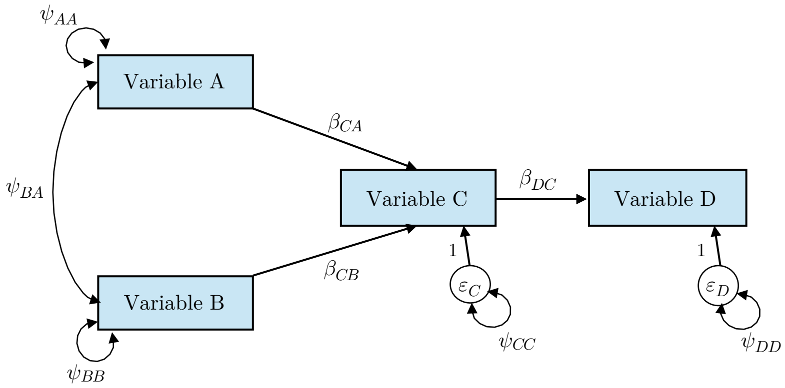
The error variance in Eq. (5) is usually not constant due to differences in the variances of the effect size estimates. To accommodate heteroscedasticity in the error term, the meta-regression model is commonly estimated with a WLS regression using the inverse of the squared standard errors, $1/SE(y_{ij})^2$, as weights in the regression (Stanley and Doucouliagos, 2012: 111). Other common weights in meta-regression are the sample size of the study or the inverse of the number of effect sizes extracted from each study. The latter is used to consider the unbalancedness of the meta-data set arising from the extraction of multiple reported effect sizes from the same study (Zigraiova and Havranek, 2016: 953). The statistical properties of the different weighting schemes in meta-regression analysis are examined in more details in Chapter 4.

2.2.3. Meta-Analytic Structural Equation Modeling

Beyond traditional meta-analysis and meta-regression, the adoption of structural equation modeling for meta-analysis allows for simultaneous testing of the relationships between multiple variables (Cheung, 2015; Cheung and Chan, 2005; Viswesvaran and Ones, 1995). This group of methods is referred to as meta-analytic structural equation modeling (MASEM). While the traditional meta-analysis approach examines a single bivariate relationship, MASEM considers the interactions with other variables, allows for the evaluation of mediating effects among these variables, and tests entire structural models, even those models that have not been tested in the primary studies before (Landis, 2013: 252). In this way, MASEM enables the meta-researcher to compare the explanatory power of competing theory frameworks and to contribute to the identification of boundaries, structures, and shortcomings of theoretical models (Bergh et al., 2016: 478-479).

MASEM begins with the definition of a theoretical model for the relations among a set of variables. To judge whether the model is supported by empirical evidence, the theoretical model is fitted on the pooled data from the primary studies. In this way, MASEM benefits from the data reported in all studies, even if the studies in the sample do not investigate all relations between the variables featured in the structural model to be tested (Viswesvaran and Ones, 1995: 866). An exemplary model describing the hypothesized relations among four example variables is presented in Figure 9 (Jak, 2015: 5-7).

Figure 9. Exemplary Path Model for Meta-Analytic Structural Equation Modeling



Notes: This figure illustrates an exemplary path model with four variables of interest (A, B, C, D). The effect of A and B on D is mediated by C.

Given is a set of $i = 1, \ldots, L$ empirical primary studies investigating one or more bivariate relations among a set of variables. For the sake of simplicity, there are four variables of interest (A, B, C, D) in this example (Jak, 2015: 5-7). A and B are exogenous variables, while C and D are endogenous variables. The effect of A and B on D is mediated by C. Moreover, there is a non-zero covariance among the two exogenous variables. Direct effects between variables are denoted by β, variances and covariances are ψ, and ε represents the residual factors. The variances of the residual factors are the unexplained variances of the two endogenous variables. Accordingly, part of the variance in C is explained by A and B. The remaining variance is the residual variance. To ensure the identification of the model, the regression coefficient of the variable on the residual factor is not estimated but commonly fixed at one (Jak, 2015: 5).

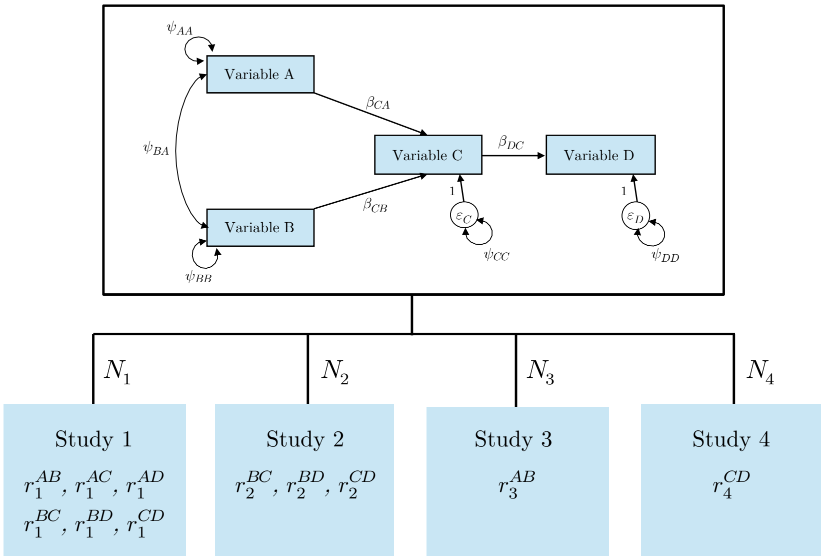
After defining the structural model relations, the meta-data is collected from the primary studies. Each study in the sample reports Pearson correlation coefficients for one or more relations among the set of variables. Moreover, the sample size needs to be obtained from each primary study. In the example with the four variables, only the first study reports the full correlation matrix among all variables. The other three studies contain correlation information only for one or two of the relations among the set of variables. The correlation coefficients collected from the primary studies can be arranged in study-specific correlation matrices:

$$\boldsymbol{R}_1 = \begin{bmatrix} 1 & & & \\ r_{AB} & 1 & & \\ r_{AC} & r_{BC} & 1 & \\ r_{AD} & r_{BD} & r_{CD} & 1 \end{bmatrix}, \boldsymbol{R}_2 = \begin{bmatrix} 1 & & \\ r_{BC} & 1 & \\ r_{BD} & r_{CD} & 1 \end{bmatrix}, \qquad (6)$$

$$\boldsymbol{R}_3 = \begin{bmatrix} 1 & \\ r_{AB} & 1 \end{bmatrix}, \boldsymbol{R}_4 = \begin{bmatrix} 1 & \\ r_{CD} & 1 \end{bmatrix}$$

When the structural model is defined and the correlation coefficients are collected from the set of primary studies, MASEM is typically conducted in two stages as illustrated in Figure 10 (Cheung, 2015: 215). In stage one, the Pearson correlation coefficients are combined into a pooled meta-analytic correlation matrix. In stage two, the structural model is fitted onto this pooled correlation matrix.

Figure 10. Illustration of Meta-Analytic Structural Equation Modeling



Notes: This figure illustrates an exemplary meta-analytic structural equation modeling approach. MASEM pools a set of correlation coefficients among several variables into a meta-analytic correlation matrix and then fits a structural model onto the pooled data.

Stage One. In contrast to classical structural equation modeling methods, MASEM typically fits the model to the meta-analyzed correlation matrix rather than the covariance matrix (Cheung and Chan, 2009: 29). Several methods exist to compute the pooled correlation matrix from the correlations obtained from the set of primary studies (Jak, 2015: 21-28). The simplest approach is equivalent to the traditional meta-analysis methods described in Eqs. (1), (2), and (3). It takes the correlation coefficients from the primary studies and pools them into one overall effect by calculating weighted mean correlations, either using the FEM or REM weights. The weighted means are calculated separately for each bivariate relation. Advanced multivariate pooling methods, such as the generalized least squares (GLS) approach (Becker, 1992: 347-348) or the two-stage structural equation modeling (TSSEM) method (Cheung and Chan, 2005: 34-37), consider the dependence structure among the effect size estimates. Beyond non-independent effects, a second issue might arise if the correlation coefficients are not homogenous across studies, which means that heterogeneity exists. Similar to traditional meta-analysis, a REM approach can be applied to pool the correlation matrix and to consider heterogeneity (Cheung, 2015).

Stage Two. To fit the structural model to the meta-analyzed correlation matrix, the model parameters are arranged as matrices. The matrix **B** captures the path coefficients, matrix **P** represents variances and covariances, and matrix **I** is an identity matrix (Jak, 2015: 10).

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_{CA} & \beta_{CB} & 0 & 0 \\ 0 & 0 & \beta_{DC} & 0 \end{bmatrix}, \mathbf{P} = \begin{bmatrix} \psi_{AA} & & & \\ \psi_{BA} & \psi_{BB} & & \\ 0 & 0 & \psi_{CC} & \\ 0 & 0 & 0 & \psi_{DD} \end{bmatrix},$$
$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7}$$

The general hypothesis to be tested by fitting the structural model to the meta-data is that the population covariance matrix $\mathbf{\Sigma}$ is equal to the model implied covariance matrix $\mathbf{\Sigma_{Model}}$. The three matrices **B**, **P**, and **I** serve as input to compute the model implied covariance matrix (Cheung, 2015: 17):

$$\mathbf{\Sigma_{Model}} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{P}(\mathbf{I} - \mathbf{A})^{-1\mathrm{T}}. \tag{8}$$

Since the population covariance matrix $\mathbf{\Sigma}$ is unavailable to the meta-researcher, MASEM makes use of the sample covariance matrix **S**. To estimate the model parameters in **B** and **P**, the difference between the model implied covariance matrix $\mathbf{\Sigma_{Model}}$ and the observed covariance matrix **S** is minimized. This leads to the following discrepancy function that can be estimated by maximum likelihood (Cheung, 2015: 26):

$$F_{ML} = log|\mathbf{\Sigma_{Model}}| - log|\mathbf{S}| + \mathrm{trace}(\mathbf{S}\mathbf{\Sigma}^{-1}_{\mathbf{Model}}) - p, \tag{9}$$

where p represents the number of variables in the model. If the model perfectly fits to the data, F_{ML} would be zero and $\mathbf{\Sigma_{Model}} = \mathbf{S}$.

Besides the general MASEM approach, there are alternative procedures to the two stages described above. A common method is the TSSEM (Cheung and Chan, 2005), where multigroup structural modeling is applied to pool the correlation coefficients in stage one and WLS estimation is used to fit the model to the pooled data in stage two. For a detailed description of the different MASEM approaches, see Cheung (2015) and Jak (2015).

Similar as for standard structural equation modeling, MASEM applies several measures to evaluate the fit of the structural model. For example, the $\chi 2$ goodness-of-fit statistic quantifies the discrepancy between the sample and the model-implied covariance (Hu and Bentler, 1999: 2). Furthermore, various incremental, absolute, and comparable fit indices are available to supplement the $\chi 2$ test, such as the Comparative Fit Index (Bentler, 1990) or the root mean square error of approximation (Steiger and Lind, 1980).

So far, the MASEM methods have been frequently applied in management sience (Bergh et al., 2016). In finance research, MASEM is rarely used (Hang et al., 2021b; van Essen et al., 2015b).

The previously presented meta-analysis methods are independent of the specific research discipline and can be applied in all areas of economics, finance, and business research. The remainder of this thesis will focus specifically on the applications and advances of meta-analysis in finance research.

2.3. Review of Previous Meta-Analyses in Finance

This section presents a literature review of the existing meta-analyses in finance and derives opportunities and challenges accompanying the application of meta-methods.

2.3.1. Literature Search and Study Selection

To identify existing meta-analyses in finance, I conducted a comprehensive literature search. In the first step, I used keywords to search in the following electronic databases: ABI/INFORM Complete, Business Source Premier, EconLit, and Google Scholar. The search term combines two groups of keywords: those related to the meta-analytic nature of the study ('meta-analysis', 'meta-regression', 'quantitative review'), and those with a focus on the main topic fields in finance research ('finance', 'asset pricing', 'financial intermediation', 'financial markets', 'investment', 'capital markets'). In a second step, I screened all articles published in the leading 35 journals in finance with a ranking of 4*, 4, or 3 in the 2018 Academic Journal Guide (AJG) issued the Chartered Association of Business Schools (CABS, 2018). In the third step, the reference lists of all relevant articles identified in the first two steps were investigated to identify further research publications.

During the literature search process, I excluded several studies from the sample for the following reasons:

(1) The unit of analysis is not the results reported in empirical primary studies, but rather the performance measures of funds (Coggin et al., 1993; Coggin and Hunter, 1993), forecasts of analysts (Coggin and Hunter, 1983), or the effects of cloud cover on stock returns across several stock exchanges (Keef and Roush, 2007).

(2) The authors call their approach meta-analysis, although it is not a quantitative integration of the literature using one of three methodological frameworks introduced in

the previous section. For example, Harvey et al. (2016) and Harvey (2017) use the term meta-analysis to compare factors for the cross-section of expected returns across different studies. However, they do not present a research synthesis using the common statistical methods for meta-analysis.

(3) The study presents a real meta-analysis, but the examined research question cannot be assigned to finance research. However, it should be noted that there is no clear cutoff for which studies should be included or excluded. This holds especially for topic areas that lie in-between finance and other research fields, especially economics and management science. For example, the examination of the relation between ownership structure and firm performance covers aspects from both management and finance research. I decided to add all meta-analyses examining such cross-disciplinary topics if their list of included primary studies covers a substantial number of research papers published in traditional finance journals. Nevertheless, this approach is subjective, and thus the selected list of meta-analyses in finance depends on the author's judgement.

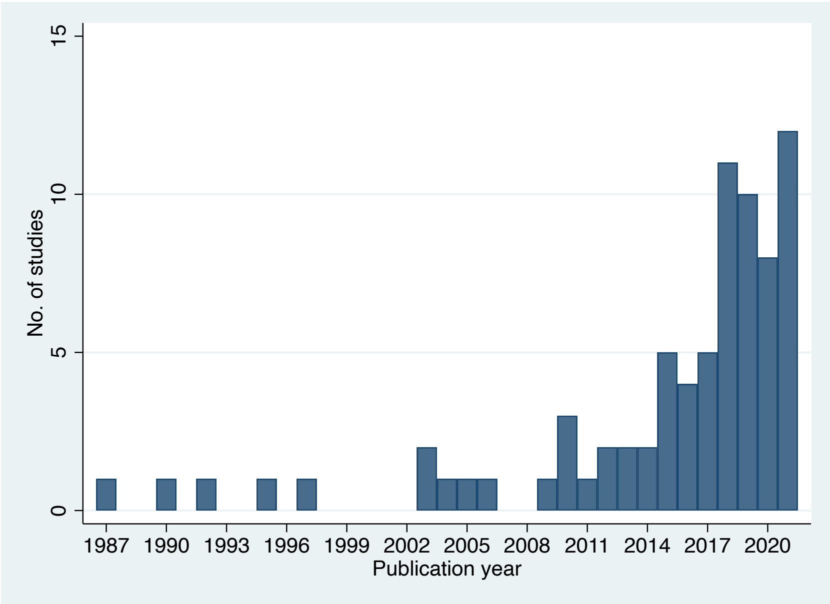
After applying the selection criteria on the set of collected search results, the final sample consists of 76 original papers using meta-analysis to aggregate and compare empirical results in finance. The last study was added to the sample in January 2022.

2.3.2. Descriptive Statistics

Figure 11 shows the number of meta-analyses in finance broken down by publication year. The earliest published study is by Coggin and Hunter (1987), who aggregate the results of two primary studies reporting estimates for the impact of risk factors in asset pricing. Despite this early application of meta-analysis, there were only a few studies published from then until 2014. Interestingly, there has been an increasing trend in recent years, with 41 meta-analyses published on finance-related topics since 2018. The rather small number of previous meta-analyses compared to other related research areas, especially economics and management, and the fact that half of the existing studies have been published in the last four years illustrate that meta-analysis in finance is still a relatively young discipline.

I see three main reasons for the growing interest in meta-analysis research in finance. First, the amount of empirical research output has increased significantly over the past decade. This provides a large and diverse body of empirical results that can be meta-analyzed. Second, the rising demand for transparency in empirical research calls for methods such as meta-analysis to detect and correct distortions like publication selection or p-hacking. Third, meta-analysis has been successfully applied in various related areas like economics, management, accounting, and marketing. This might generate a ‘spill-over’ effect into finance research.

Figure 11. Growth of Meta-Analysis Studies in Finance



Notes: This graph shows the number of studies applying meta-analysis in finance per publication year, as of 09/01/2022.

To break down the topics being addressed in the prior meta-analyses, I group them into three topic fields: (1) asset pricing, (2) corporate finance, and (3) financial intermediation. The full list of all identified meta-analyses is reported in Table A.1 in the appendix. Within the topic fields, I further classify studies into topic areas.[11] As a result of this classification, we can see that the majority of meta-analyses examine research questions in corporate finance (49 studies), followed by asset pricing (20 studies), and financial intermediation (7 studies). Within corporate finance, four topic areas are dominant: corporate governance (15 studies), raising capital (8 studies), capital structure (8 studies), and risk management (8 studies). A reason for the predominant use of meta-analysis in corporate finance might be the various fundamental theories and hypotheses, often with ambiguous empirical evidence from a large body of literature. This may create demand for aggregation and comparison using meta-analysis. Also, the examined effects are often similar in terms of the applied statistical models producing comparable quantities, which lays the foundation for a meaningful synthesis through meta-analysis.

Regarding the publication outlets, I find that 66 meta-analyses are published in referred journals, while 10 studies are unpublished working papers, conference papers, or book chapters.

[11] These topic areas are the same as for the list area assignments of the AFA annual meeting.

Among the published papers, several articles appeared in leading field journals like the *Journal of Financial Economics* (Holderness, 2018), *Management Science* (Kysucky and Norden, 2016), *Journal of Banking and Finance* (Feld et al., 2013), or *Journal of Empirical Finance* (van Ewijk et al., 2012). Given these prestigious journals accepting meta-analyses for publication and the recent increase in the overall number of meta-analyses might indicate a breakthrough and portend a broader acceptance of meta-analysis in finance research in the future.

A breakdown of the publication outlets shows that 32% are original finance journals, 30% are management journals, 22% are economics journals, 1% are accounting journals. The remaining 15% are working papers and book chapters. Accordingly, a large number of finance-related meta-analyses are published in management journals. The dominance of management journals may be due the fact that there are various meta-analyses on corporate governance topics included in the sample, which cover elements of both finance and management. Another factor could be that editors and reviewers of management journals are more willing to accept meta-analysis papers because the method is widely used and acknowledged in their field (Geyskens et al., 2009).

Table 3 provides summary statistics for several dimensions characterizing the study design and methodology applied in the 76 meta-analyses. Based on a systematic review of these studies, I derive opportunities and challenges of meta-analysis in finance, which are discussed in the subsequent Sections 2.4 and 2.5.

2.4. Opportunities of Meta-Analysis in Finance

Knowledge Accumulation. By combining evidence from many seemingly conflicting primary studies, meta-analysis draws a statistical summary of the research field under examination. Meta-analysis creates a *"store of accumulated knowledge"* (Grewal et al., 2018: 9), which is important for the development of an overall understanding of the answers to important research questions. The demand for knowledge accumulation even increases with a growing volume of empirical research papers, which are difficult to aggregate and compare in an objective way without statistical methods like meta-analysis. The estimation of an overall population level effect size, which is the main objective of earlier meta-analyses in finance (69 out of 76 meta-analyses), builds on previous empirical findings produced by many different researchers using a range of different methodologies and data sources over various time periods and countries. Due to this variety of research setups used as input, meta-analysis offers a level of independence and rigor over traditional primary studies.

Table 3. Quantitative Summary of 76 Meta-Analyses in Finance

Study Attribute	Statistic / No. of Studies	Study Attribute	No. of Studies
Publication Year		***Analysis of Publication Selection Bias***	
Mean	2015	No analysis of publication selection bias	36
Median	2018	Analytically and graphically	23
Min.	1987	Analytically	11
Max.	2021	Graphically	6
Number of Studies in the Meta-Sample*		***Analysis of Misspecification Bias***	
Mean	60	No analysis of misspecification bias	56
Median	37	Moderator variables in the meta-regression	20
Min.	2		
Max.	613		
Number of Observations in the Meta-Sample*		***Analysis of Heterogeneity*****	
Mean	461	Moderator variables in the meta-regression	57
Median	184	Subgroup analysis	26
Min.	3	No analysis of heterogeneity	5
Max.	6,312		
Primary Goal of the Meta-Analysis**		***Meta-Method*****	
Estimation of summary effects	69	Meta-regression analysis	56
Analysis of methodological/data heterogeneity	64	Hedges and Olkin meta-analysis	20
Analysis of structural heterogeneity	40	Hunter and Schmidt meta-analysis	13
Analysis of publication selection bias	32	Meta-analytic structural equation model	4
		Lipsey and Wilson meta-analysis	2
		ANOVA/ANCOVA	1
		Multivariate meta-analysis	1
Effect Size**			
Pearson correlations	26		
Partial correlations	24		
Abnormal returns	11		
Regression coefficients	6		
Elasticities	5		
Other effect sizes	17		

Notes: This table shows the summary of the study characteristics of 76 meta-analyses in finance, as of 09/01/2022. * Some studies report several distinct meta-analyses in the same study (e.g., for the determinants of the corporate debt-equity ratio). The number of studies and the number of observations consider that some studies report different samples for multiple meta-analyses reported in the same study. ** Multiple counts per study possible due to the application of different methods and robustness analyses.

Precision. An important advantage of meta-analysis comes from the way it handles sampling error. Sampling error is the random deviation from the overall population, which arises when researchers draw samples, for example, by selecting a set of firms. The larger the sample in a study, the smaller the sampling error and the more precise are the effect size estimates (Combs, 2010: 10; Hackshaw, 2008: 1142). As a good meta-analysis combines the power of many studies, positive and negative sampling errors average out and the overall sampling error becomes smaller than in any of the included individual studies, especially those with small samples (Hunter and Schmidt, 2004: 84).

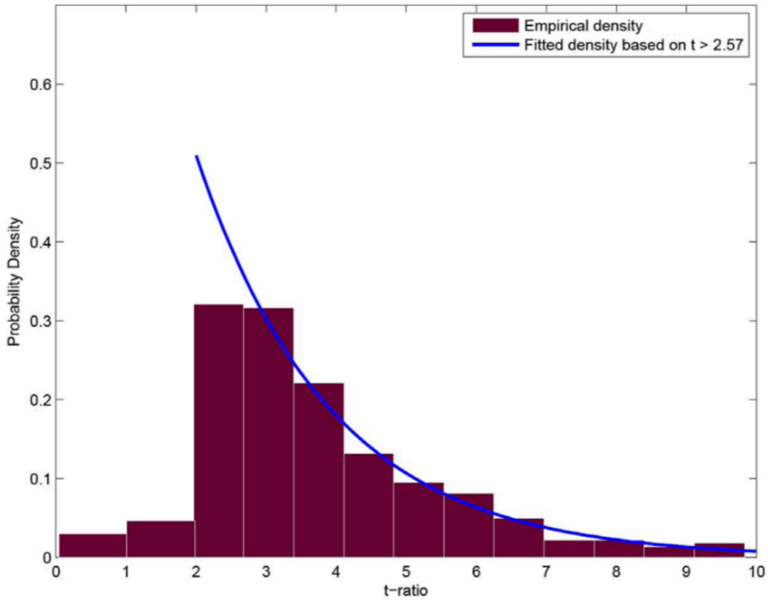
According to Table 3, the median meta-analysis in finance includes 37 primary studies and 184 effect size estimates. The sample maximum is 6,321 effect size estimates and 613 primary studies. This underpins the power of meta-analysis to bring together and compare a large body of literature on a particular phenomenon while minimizing sampling error. However, the gain in precision in finance might be lower than in other research fields. In contrast to empirical research in economics, which is often based on data with lower frequency (e.g., quarterly data on economic growth), many data sets in empirical finance, especially in capital markets research, are often available with higher frequency (e.g., daily or intraday stock prices). This increases sample sizes within primary studies and thus also the precision of the effect size estimates. Nevertheless, also in finance, the aggregation within a meta-analysis has the power to further decrease sampling error by pooling studies with larger and smaller sample sizes.

Identification and Correction of Publication Selection Bias. Publication selection bias has been widely regarded as a serious threat to the validity of empirical research results (Andrews and Kasy, 2019; Begg and Berlin, 1988; Ferguson and Brannick, 2012; Harrison et al., 2017; Hunter and Schmidt, 2004; Roberts and Stanley, 2005; Rothstein et al., 2005; Stanley, 2005; Stanley and Doucouliagos, 2014; Thornton and Lee, 2000). As Doucouliagos and Stanley (2013) find in a review of 87 meta-analyses that analyze 3,599 primary studies in total, most fields of economics research (including financial economics) are affected by publication selection.

In finance, Harvey et al. (2016: 11) detected strong selective reporting of significant research results in factor studies explaining the cross-section of expected returns. Figure 12 shows the distribution of the *t*-values reported in the 313 factor studies (Harvey, 2017: 1401). The number of studies reporting *t*-values between 2.0 and 2.57 is almost identical to the number of studies reporting *t*-values in the interval between 2.57 and 3.14, *"which only makes sense if there is publication bias"* (Harvey, 2017: 1401). Accordingly, it appears that authors of factor studies systematically neglect insignificant results and preclude them from being published.

When the 'consumers' of finance research are unaware of the alternative results that end up in the drawer due to publication selection and thus remain unpublished, their conclusions about the significance and robustness of the truncated sample of published results may be severely biased (Mitton, 2022: 2). If publication selection is present in a research field, averaging across primary studies in a meta-analysis will create biased mean effects (Stanley, 2005: 311). However, meta-analysis, especially meta-regression analysis, provide powerful means for identifying and even filtering out publication selection bias. Thereby, meta-analysis contributes to the transparency and increased credibility of empirical research results. This is a clear advantage over primary research as several biases cannot be controlled at the level of individual studies; this particularly holds for publication selection bias that *"is caused by the process of conducting empirical research itself"* (Stanley and Doucouliagos, 2012: 4).

Figure 12. Histogram of *t*-statistics from 313 Factor Studies in Asset Pricing



Notes: Distribution of *t*-statistics from 313 studies examining factors in asset pricing reported by Harvey (2017: 1401).

There are 40 out of 76 meta-analyses in finance that explicitly control for the impact of publication selection and associated biases via graphical analysis or statistical testing. Most of these studies discover a medium-to-strong impact of publication selection bias. For example, Astakhov et al. (2019) conduct a meta-analysis of the firm size effect on stock returns. On average across 102 prior studies, smaller firms outperform larger stocks by 5.08% in annualized terms. However, after correcting for publication selection, the mean size premium is only 1.72% (Astakhov et al., 2019: 1464).[12] Accordingly, due to publication selection, the literature as a whole exaggerates the firm size effect by a factor of three.

Identification and Correction of Model Misspecification. Primary studies in finance use a wide range of model specifications (Mitton, 2022: 542). Changes in the primary study research design are often regarded as innovation over prior work (Koetse et al., 2005: 2). Especially in observational research where even the most sophisticated econometric models *"cannot eliminate all the potentially confounding influences"* (Stanley and Doucouliagos, 2012: 3), the selection of the control variables that are added to the model and the variables that are omitted (consciously or unconsciously) can have a strong impact on the results. Often there are a

[12] This refers to the implied difference between the 10^{th} and the 90^{th} percentile of NYSE stocks.

number of variables that are used consistently throughout the literature (e.g., controls for firm size), while other control variables are more specific to the research question to be addressed in a primary study. Data constraints and the desire to be 'different' often lead to varying sets of control variables. Omitted-variable bias, which occurs when important variables are left out of the primary study model, is a common threat in observational research.

When the empirical literature is a mix of correctly and incorrectly specified models, this could also cause biased results in at the meta level (Stanley and Jarrell, 1989: 162). However, meta-analysis, and especially meta-regression, are designed to detect and control for misspecification bias. In the meta-regression model in Eq. (5), the vector $\boldsymbol{M}$ accounts for potential biases arising from the exclusion of relevant control variables from the original regression model in Eq. (4). Meta-regression reveals the sensitivity of empirical effects against omitted variables and changes in other model specifications. Even a synthetic overall effect assuming the 'ideal' model specification can be estimated with meta-regression.[13]

As Table 3 shows, 20 out of the 76 prior meta-analyses in finance explicitly test for misspecification bias by coding the presence or absence of important control variables in the primary regression using (dummy) moderator variables. For example, Feld et al. (2013: 2854) find that six control variables in the primary study regression model significantly determine the reported marginal effect of tax on the corporate debt level. As finance studies often include a fixed standard set of several control variables, meta-analysts may also include a dummy variable indicating whether this standard set of control variables is included or not.

Exploration of Heterogeneity. The results in Table 3 suggest that the analysis of methodological and data-related heterogeneity is a major goal of the previous meta-analyses in finance. Heterogeneity is inherently present in empirical finance research, which is mostly a non-experimental field of science, where study design and data selection are different across different studies and authors. With the rapid increase of empirical research and the progress in data availability and statistical methods, heterogeneity even grows over time. With meta-analysis, researchers can leverage the study level differences to find an explanation for the variation in effect size estimates and to understand the relationship between those effects and the study design characteristics (Borenstein et al., 2009: 107-109). Such an analysis of heterogeneity can resolve conflicts in the literature and determine important moderating factors.

[13] See Section 3.7.5 for an applied example of how meta-regression results can be used to predict best practice estimates.

In Eq. (5), the vector $\boldsymbol{M'}$ represents the moderator variables for methodological and data-related heterogeneity. The estimated meta-regression coefficients capture the impact of the moderator variables on the examined effect size. As many moderators are dummy or categorical variables indicating the presence of a certain study characteristic (e.g., whether a model controls for endogeneity or not), the estimated coefficients commonly reflect the average impact on the reported effect size estimates if the study design deviates from the base group in that aspect, all other factors being equal (Abdullah et al., 2015). Thus, significant meta-regression coefficients can be interpreted as an indication that a specific study characteristic changes the focal relationship, which means that it increases or decreases the effect size estimates exhibited by models with this attribute (Stanley and Doucouliagos, 2012: 85). For example, van Ewijk et al. (2012: 825) find in their meta-regression that the equity premium gathered from 24 primary studies is on average 1.31% larger when ex-post as compared to ex-ante methods are applied and 3.54% smaller before the year 1910 as compared to more recent time periods.

The evaluation of the 76 meta-analyses in finance reveals a group of frequently applied moderator variables that are included in the meta-regression models in addition to controls for publication selection and misspecification (as described in the two paragraphs above):

- Measurement of the focus variables: for example, variations in the definition and measurement of the dependent and independent variables.

- Data characteristics: for example, the number of observations, average sample year, or data frequency.

- Method choices: for example, control for endogeneity, type of estimator, inclusion of fixed effects, or robust estimation of standard errors.

- Publication characteristics: for example, the number of citations, journal impact factor, or the publication status of the study.

As data in finance is often derived from similar databases, such as Bloomberg or Thomson Reuters, I also recommend including the data source of the primary study as a moderator variable. Moreover, real economic differences across countries, industries, or company sizes can be explained via meta-regression. In finance, we often see that the majority of studies refer to U.S. data. Hence, I recommend adding a binary variable for studies focusing on the U.S. as compared to other countries or world regions.

The final selection of the moderators, of course, depends on the specific research question and the data observable from the primary studies.

Theory Advancement. Meta-analysis can be helpful in testing different or alternative competing theories on an aggregated level. The MASEM approach presented in Section 2.2.3 even allows for the comparison of complex theoretical models and the testing of their empirical support. Moreover, the meta-analyst can add previously unexamined data from external sources to test new theory-driven hypotheses about moderating effects. For example, meta-analysis can be employed for testing cross-country boundary conditions that drive the direction and size of a certain effect, which would otherwise demand an extensive collection of data over many different countries and time periods. For example, Kysucky and Norden (2016) collect a sample of 101 primary studies on the benefits of relationship lending. The authors collect additional data on bank competition from the World Bank. Based on the country and time period examined in the primary studies, they assign the bank competition information to the effect size estimates and find that the level of bank competition is an essential moderator for the realization of benefits from relationship lending (Kysucky and Norden, 2016: 99).

Predictions and Benchmarks. The results from meta-analysis, especially from meta-regression, can be used to make predictions and to derive practical implications. The meta-analyst can use the meta-regression coefficients from Eq. (5) to estimate the underlying true effect conditional on the values of the examined moderator variables. By inserting best practice values for the moderators, meta-analysis can create a synthetic study. For example, Feld et al. (2013: 2854) predict the marginal tax impact on debt for a synthetic benchmark study by setting all significant dummy meta-regressors to a value of one (i.e., the study is in line with the benchmark) and all insignificant as well as the continuous moderators are set to the sample mean (i.e., referring to the average study). This benchmark study predicts a marginal tax effect on the debt ratio of 0.27.

Future enhancements in primary studies can be evaluated against the benchmark given by meta-analysis. Moreover, Bayesian approaches, which often rely on theoretical distributions of priors, can refer to meta-analysis to provide the empirical distribution of all reported empirical effects in the literature and to generate alternative priors from this empirical distribution of effects (Moral-Benito, 2012: 805-807). Meta-analysis results are also helpful for reviewers, editors, and readers of scientific papers to have a reference point derived from previous literature when evaluating the findings of a new study. Moreover, meta-analysis results can be used to discover new research questions and demand for further research.

Evidence-Based Decision Making. To make a practical impact, empirical research must be approachable for practitioners, politicians, and other decision-makers such as corporate managers. The growing body of literature, distorting biases, and the variation among empirical results constitute challenges for the practical impact of research. Meta-analysis summarizes the state of the art in a specific domain, explains differences in its results, detects and corrects biases. This makes research findings more accessible for finance practitioners, as it allows them to integrate aggregated scientific outcomes into their decision-making and to rely on the evidence provided by the literature as a whole instead of a selected individual results.

2.5. Challenges of Meta-Analysis in Finance and How to Address Them

Choice of Effect Size. The effect size is the central unit for defining the results of different studies on the same scale and thus making them comparable. Therefore, the choice of the effect size is crucial for an impactful meta-analysis (Stanley and Doucouliagos, 2012: 22-29).

Among the previous meta-analyses in finance (Table 3), several studies analyze economic effects, such as abnormal returns or elasticities (among others, Holderness, 2018; van Ewijk et al., 2012; Veld et al., 2020). For example, Rahim et al. (2014: 390) find that the mean cumulative abnormal return across 35 event studies is -1.14% for the announcement of convertible debt offerings and -0.02% for warrant-bond offerings. Moreover, Feld et al. (2013: 2851) show that the average marginal tax effect on the corporate debt ratio is 0.27. Accordingly, the leverage-to-equity ratio increases by 2.7 percentage points if the marginal tax rate increases by 10 percentage points.

In contrast to economic effect sizes, 50 of the prior meta-analyses examine partial or zero-order correlations, which are unitless measures for the association between two variables and are therefore suitable for aggregation over studies. Nevertheless, the aggregated findings do not reveal the economic magnitude of the effect under examination, but rather the direction and size of the statistical relation. Statistical measures, such as partial correlations or Pearson correlations, do not allow for a meaningful interpretation of the economic magnitude of the accumulated effects. Thus, economic measures are the effect sizes to be preferred over statistical effect sizes as they enable a more powerful interpretation of the findings and increase the contribution of meta-analysis to practical decision-making. Sometimes economic measures can even be re-calculated by the meta-analyst based on the information provided in the primary studies. Statistical effect size measures can serve as a second-best option, especially if economic effects are not available, for example, due to missing data.

Sample Composition. As with any empirical analysis, the quality of the input data determines the reliability of the outcomes. 'Garbage-in-garbage-out' is an issue that is often discussed in the course of meta-analysis (Borenstein et al., 2009: 381). This challenge spans across multiple layers that determine the best practices to address the issues related to data quality.

First, if results are obviously incorrect (e.g., a *t*-value does not correspond to the statistical significance level indicated by the asterisks), such observations should be corrected if possible or should be omitted from the meta-analysis if they cannot be corrected. Second, meta-samples often include observations from primary studies published in top journals, obervations published in lower-ranked journals, and even studies that have not been published in a journal. Such a mix of studies is rather common, especially in well-developed research fields where the initial seminal papers are often published in leading field journals and follow-up work examining the same research question using different data, time horizons, or methodologies are spread across other journals, conference proceedings, or working papers.

Following Stanley and Doucouliagos (2012: 19), meta-analysts *should "err on the side of inclusion"* of all available studies rather than being selective, as selection requires an obvious criterion as to what a 'good' study is. Also, from a statistical point of view, the meta-researcher should aim at including all available effect size estimates to increase the sample size and thereby the efficiency of the meta-analysis estimates (Stanley and Doucouliagos, 2012: 72). If there are measurable differences, for example, among more prominently published papers and studies in lower-ranked journals, meta-analysis can detect and control for their impact. Essential for the inclusion of all available primary studies is that meta-analysis can explicitly control for both paper quality characteristics (e.g., by including the number of citations or the journal ranking as a moderator variable) and methodological quality (e.g., by including a dummy indicating whether a study controls for endogeneity). In all cases, the meta-analyst should inform the reader how quality differences are accounted for and controlled.

Econometric Issues. Meta-analysts in finance typically code more than one effect size estimate per study. According to Table 3, the median is 184 observations from 37 studies. This results in about 5 effect size estimates per primary study. However, sampling multiple estimates per study leads to 'within-study correlation' among the effect size estimates. Moreover, in contrast to experimental research, studies in empirical economics are usually non-independent and commonly rely on similar and overlapping data sets. Consequently, samples from different studies and authors are often related, e.g., because they examine data of the same or similar companies (e.g., S&P 500 firms). This introduces 'between-study correlation' but could also cause other types of dependencies like between-author or between-database correlation.

The sources and consequences of dependencies in a meta-analysis are similar to those in a primary study. For example, a global panel data set in a primary study may include interdependencies arising from the clustering of firm observations taken from the same country, identical time periods, or multiple observations of the same firm across several years (Petersen, 2008: 435-437). To handle potential dependencies, meta-regression analysis can apply the same remedies as a primary study. In particular, panel regression models, robust standard errors, and multi-level models are commonly applied methods to address data dependence problems in a meta-analysis (Hedges et al., 2010; Rusnak et al., 2013: 49; Stanley and Doucouliagos, 2012: 112-117). Meta-analysis in finance should explicitly account for the different levels of dependencies, at least by using clustered standard errors at the level of the individual studies (Havranek et al., 2020: 472).

Quality. The number of meta-analyses across all disciplines has rapidly increased in recent years. At the same time, criticisms have evolved that challenge the quality and correct application of meta-analysis methods. For example, Ioannidis (2016) criticizes the mass production of redundant and misleading meta-analyses in medicine and Nelson and Kennedy (2009) discuss the use and abuse of meta-analysis in environmental and natural resource economics. The bottom line of these critiques is that it takes a rigorous and high-quality meta-analysis to realize the opportunities outlined in the previous section.

Many how-to guides and best practices on conducting and reporting meta-analyses are available in different research domains that help authors to achieve high standards and rigor. However, as in any field of research, reviewers and consumers of meta-studies also need to contribute to a high standard of quality by challenging and improving meta-analysis applications. Related guidelines for meta-regression in economics have been published by the Meta-Analysis in Economics Research Network (MAER-Net), which is an international network of scholars committed to improving economic science through meta-analysis (Havranek et al., 2020; Stanley et al., 2013). Meta-analysts in finance should refer to these guidelines during all steps of their meta-work. Making data and codes available to the public is another essential aspect of transparency and replicability of meta-research.[14]

Data Retrieval. Another challenge that often occurs in meta-analysis is the limited availability of the required data reported in the primary studies (e.g., to calculate elasticities as effect sizes). Sometimes missing information can be gathered from web appendices, earlier working paper versions, by email requests to the primary study authors, or by re-calculating missing values from other reported data. It is obvious that this is an important challenge, one

[14] See, for example, the data repository of the Deakin Lab for Meta-Analysis of Research: http://www.deakin.edu.au/business/research/delmar/databases.

which can be overcome when finance journals routinely require authors of empirical work to report all information required for replication and meta-analysis in their paper or in a technical appendix. This is a routine that is already common in other research fields such as management science.

Contribution. The statistical aggregation of empirical studies, including the exploration of heterogeneity and the analysis of publication selection or misspecification bias, is a unique advantage of meta-analysis that comes with an inherent value for a research field. However, to be impactful, a meta-analysis must clearly elaborate its contribution, also against large-sample and cross-country primary studies. It is important to illustrate how the meta-results encode dissidence of previous findings and have theoretical implications.[15] The analysis of specific moderator variables that are subject to ongoing discussions, for example, related to specific methodological choices or variable measurement, might support the progress in a field. Moreover, if there is a lack of cross-country studies due to data collection problems for large international samples, meta-analysis can help to create insights into country-level moderators by integrating the findings of many single-country studies (among others, Holderness, 2018; Kysucky and Norden, 2016). Also, the inclusion of international variables collected from additional databases can reveal new insights in a research field (e.g., macroeconomic data matched to the studies based on the examined time period and country). In a similar vein, meta-analysis can aggregate data from both very early studies and more recent studies, which enables the meta-researchers to derive long-run estimates across the entire time period analyzed in the previous literature. Meta-analysts in finance should clearly communicate their contribution, and this contribution should be more than just a statistical integration of previous findings from the literature.

2.6. Summary

Over the past two decades, we have observed a massive increase in the volume of empirical research in finance, with results being spread across journals, unpublished working papers, and book chapters. Despite or even because of the large amount of empirical output, the findings from finance research have not always been conclusive and often depend on certain sample characteristics and methodological choices by the primary studies' authors. Additionally, we observe growing concerns about the replicability and reliability of reported empirical results. The inconclusiveness and skepticism of single-approach findings call for more powerful and

[15] In the management literature, for example, the leading field journal, the *Academy of Management Journal*, only accepts meta-analyses for publication when they come with theoretical implications: https://aom.org/research/publishing-with-aom/author-resources/submitting-to-amj.

generalized hypothesis testing, which may be achieved by review techniques such as meta-analysis.

The chapter introduced meta-analysis to finance researchers by discussing its strengths and challenges. In general, meta-analysis collects and synthesizes existing research results to determine patterns of heterogeneity, to reach generalizations, and to identify as well as correct common biases in empirical research. Through its transparent and objective way of aggregating a large set of primary studies from different authors and data sources, it can play a valuable role in strengthening the evidence of financial phenomena and increasing credibility of empirical research findings. The review of 76 prior meta-analyses in finance illustrates the ability of meta-analysis to bring coherence to a pool of inconsistent findings and to explore why studies produce different results. With meta-analysis, we can also learn about the impact that methodological and data-related choices in finance have on a certain empirical result. Biases in empirical research, especially the preferential publication of significant and strong results (publication selection) and the incorrect specification of econometric models (misspecification), might produce a literature conveying a false impression regarding the underlying effect in question. Meta-analysis provides methods for finance researchers to identify, assess, and correct specific biases. At the same time, econometric issues in the meta-model, questions about data composition, as well as a clear message of the contribution through the meta-analysis must be addressed to make meta-research impactful.

Taken together, meta-analysis, if applied carefully, can be a powerful tool that facilitates progress in finance research. There are several promising topic areas in corporate finance, asset pricing, and banking, where a systematic aggregation of empirical evidence could be fruitful. Especially in those research fields with controversy in theory and a rich set of empirical studies with mixed evidence using similar but differently specified methods, the application of meta-analysis has the potential to provide valuable insights. Also, for event studies, meta-analysis can be impactful as it allows for the calculation of overall abnormal return effects across several country-level studies and the testing of new moderating factors. The next chapter presents an application of meta-analysis in finance, including all steps outlined in this chapter.

Chapter 3. Corporate Financial Hedging and Firm Value: An Applied Meta-Analysis[16]

"Empirical results in existing studies are ambiguous with respect to the valuation impact of corporate hedging."

(Choi et al., 2013: 240)

This chapter presents an application of meta-analysis in corporate finance. The meta-analysis methods presented in Chapter 2 are applied to the empirical literature examining the firm value effects of corporate hedging. The chapter starts with a summary of the theoretical background of the hedging and firm value relation (Section 3.2) and a review of the related empirical literature (Section 3.3), followed by a description of the data collection (Section 3.4). Afterwards, the specification of the meta-regression model (Section 3.5) and the related moderator variables (3.6) are defined. This is followed by the presentation of the results (Section 3.7) and the related discussion of the findings (Section 3.8). The chapter concludes with a summary (Section 3.9).

3.1. Motivation

Today, many companies operate in globalized and volatile markets. The corporate environment is often characterized by fluctuations in economic and financial conditions, such as unexpected changes in foreign exchange rates (FX), interest rates (IR), and commodity prices (CP). These sources of financial risk can have a large impact on the value of a firm and may even cause financial distress or bankruptcy (Berk and DeMarzo, 2014: 985). Corporate hedging as part of the overall risk management approach is an important strategy to control a firm's financial risk exposure and to ensure ongoing business success (Brealy et al., 2011: 648). Among the many approaches to hedging a firm's risks, financial derivatives are a widely used hedging instrument that allows firms to reduce their cash flow exposure to market price volatility arising from changes in FX, IR, and CP (Giambona et al., 2018: 805).

Whether corporate hedging is valuable for companies constitutes an intensely discussed research question in the corporate finance literature over the past 30 years (among many others,

[16] Parts of this section are published in the paper *Corporate Hedging and Firm Value: A Meta-Analysis*, The European Journal of Finance 27(6), pp. 461-485, 2021, co-authored by Markus Hang and Andreas Rathgeber.

Allayannis and Weston, 2001; Bartram et al., 2011; Bessler et al., 2019; Carter et al., 2006; Jin and Jorion, 2006; MacKay and Moeller, 2007; Pérez-González and Yun, 2013). While traditional finance theory claims that under the specific conditions of a perfect capital market, corporate hedging is irrelevant to firm value (Modigliani and Miller, 1958), more recent theories identify situations in which hedging can be a value-enhancing strategy (Bessembinder, 1991; DeMarzo and Duffie, 1991; Froot et al., 1993; Smith and Stulz, 1985). However, when it comes to the empirical investigation of these theories, the findings are rather mixed and do not allow for any final conclusions:

> *"The evidence of studies directly analyzing the value effect of corporate hedging is thus fairly mixed and inconclusive as well, suggesting the need for further empirical, and possibly theoretical, analysis on this issue."* (Aretz and Bartram, 2010: 362)

> *"Contrary to the theoretical literature, the empirical evidence on the implications of hedging for firm value is mixed."* (Disatnik et al., 2014: 740)

> *"As a result, the motives and value of corporate hedging are still in doubt, and positive and normative theory is underdeveloped."* (MacKay and Moeller, 2007: 1380)

Despite a large body of empirical studies investigating firm-level data to determine whether corporate hedging adds value to firms, the literature remains largely unsettled, particularly with regard to two dimensions. First, the empirical estimates for the impact of hedging on corporate values range from large positive premiums to zero and even negative effects (Table 4). Second, the design of the empirical studies that produced these varying outcomes is rather diverse in terms of the applied econometric methods, measurement of the hedging and the firm value variable, the time period of the sample, examined countries, and other aspects of data and methodology.

Table 4. Hedging Premiums Reported in Selected Empirical Studies

Authors	Journal	Hedging Premium*
Allayannis and Weston (2001)	Review of Financial Studies	4.1%
Allayannis et al. (2012)	Journal of International Economics	5.6%
Carter et al. (2006)	Financial Management	6.2%
Fauver and Naranjo (2010)	Journal of Corporate Finance	-14.3%
Jin and Jorion (2006)	Journal of Finance	-2.1%
Kim et al. (2006)	Journal of Corporate Finance	0.8%
MacKay and Moeller (2007)	Journal of Finance	-3.2%
Pérez-González and Yun (2013)	Journal of Finance	9.5%

Notes: This table shows the hedging premium, which is the percentage mark-up in firm value of hedging firms as compared to non-hedging firms. * Median of all estimates for the hedging premium reported per study.

The discordance of the empirical evidence and the variability in study designs make it difficult to assess what overall hedging premium the literature implies and how the inconsistency in the reported results is driven by heterogeneity in data and methods. To address these questions, I apply the meta-analysis methods introduced in Chapter 2 of the thesis to a sample of 71 primary studies providing 1,016 effect size estimates for the impact of corporate hedging on the value of non-financial companies. This meta-analysis contributes to the literature in several ways:

First, I aggregate the wide research record on the hedging and firm value relation and present accumulated hedging premiums for different types of risk exposures, while building on recent discussions by Harvey et al. (2016) concerning data mining and publication selection in empirical finance studies and taking advantage of the capability of meta-analysis to detect and correct publication selection bias. The aggregated hedging premiums can be interpreted as a 'consensus' effect combining all available empirical information reported in the prior literature. Such a consensus and its influencing factors might also be of great interest for corporate decision-makers when defining hedging strategies.

Second, I determine how the authors' choices regarding methods and data influence the estimates of the hedging premiums reported in their studies. The findings from the heterogeneity analysis disentangle the large variability of existing empirical results by showing which study characteristics determine the hedging premium. In this regard, I also extend the limited evidence on the country-level determinants of the value effects of corporate hedging by testing a new set of previously unexamined macroeconomic variables as contingency factors for the hedging premium.

3.2. Theoretical Background

This section introduces the basic concepts of financial risk and risk management. Moreover, the risk management process is described together with a closer look at corporate hedging as one of the key instruments for risk treatment. The section ends with the general motivation of companies to engage in corporate hedging.

3.2.1. Taxonomy of Financial Risk

The Oxford English Online Dictionary defines risk as *"(Exposure to) the possibility of loss, injury, or other adverse or unwelcome circumstance"* (Oxford English Dictionary, 2022). Risk touches on almost all aspects of corporate activities. Accordingly, companies are subject to a wide range of risks including changes in consumer demand, varying costs for raw materials, bankruptcy of contract partners, resignation of employees, disruptions in the production processes, breakdown of equipment, the development of new technologies by competitors, and many more (Berk and DeMarzo, 2014: 985).

The finance literature classifies the various corporate risks into four main categories (Christoffersen, 2003: 4-5; McNeil et al., 2015: 5):

- ***Credit Risk*** is the risk that the counterparty of an outstanding investment, such as loans and bonds, cannot or can only partly fulfill its obligations on the due date.

- ***Liquidity Risk*** is the risk of conducting transactions in markets with low liquidity leading to a lack of marketability of an investment that cannot be bought or sold (quickly enough) to prevent or minimize financial loss.

- ***Market Risk*** is the risk of changes in the value of a financial position or portfolio due to unexpected changes in the market value of the underlying components on which the position or portfolio depends. These price changes may relate to equity, FX, IR, or CP.

- ***Operational Risk*** is the risk of financial loss resulting from inadequate or failed internal processes, people, and systems or from external events.

Depending on the industry and business purpose, every company holds an individual risk portfolio, implying a different impact of the four abovementioned risk categories (Vernimmen et al., 2009: 387-389). For example, a financial institution might have a larger exposure to credit and liquidity risk from buying and selling financial products, whereas a manufacturing company might have a greater exposure to market and operational risks occurring from buying raw materials and selling manufactured products in international markets. Usually, companies have an obvious interest in actively managing their risk exposure in order to prevent unforeseen negative impact on the business. Therefore, risk management is an essential element of the corporate strategy and planning (Brealy et al., 2011: 645-668).

3.2.2. Risk Management

Risk management includes *"coordinated activities to direct and control an organization with regard to risk"* (International Organization for Standardization, 2018). ISO 31000, which is the international standard for risk management, describes risk management as a process that consists of several key elements including communicating, assessing, treating, reporting, and monitoring risk. The ISO risk management process is illustrated in Figure 13 (International Organization for Standardization, 2018).

Figure 13. The Risk Management Process according to ISO 31000:2018

Notes: This figure illustrates the risk management process defined in the ISO 31000:2018 – Risk Management Guidelines (International Organization for Standardization, 2018).

The individual elements of the risk management process contain aspects of various research disciplines including controlling, accounting, governance, strategic management, and finance (International Organization for Standardization, 2018):

- ***Communication & Consultation*** is about assisting relevant stakeholders in understanding risk, promoting awareness of risk, and bringing different areas of expertise together for each step in the risk management process.

- ***Scope, Context, Criteria*** refer to customizing the risk management process to the corporate circumstances including objectives, tools, and resources.

- ***Risk Assessment*** is the core element of the risk management process to identify, analyze, and evaluate risk.

- ***Risk Treatment*** includes the selection and implementation of strategies to address risk.

- ***Monitoring & Review*** aims at assuring and improving the quality and effectiveness of the risk management process and its implementation.

- ***Recording & Reporting*** comprises documentation, communication of risk management outcomes, and the provision of information for decision-making.

There is a wide range of research on Enterprise Risk Management (ERM) and success factors for its implementation (among many others, Beasley et al., 2005; Hopkin, 2018; Nocco and Stulz, 2006). The finance literature has a long tradition of risk assessment and risk treatment. Value-at-Risk, volatility modeling, dynamic conditional correlations, or copulas are just a few of many univariate and multivariate models in finance to assess financial risk (Christoffersen, 2003: 65-216). Moreover, various strategies and instruments exist to reduce or eliminate risk, including insurance, diversification, hedging, and de-leveraging (Berk and DeMarzo, 2014: 985-1025).

Following Banks (2004: 8-9), risk treatment approaches can be divided into three groups (Figure 14), including loss control, loss financing, and risk reduction:

Figure 14. Approaches for Risk Treatment

Notes: This figure illustrates the three main groups of risk treatment following the classification by Banks (2004: 8-9).

- ***Loss Control*** includes approaches to taking safety measures, implementing controls, or taking precautions to reduce a potential risk (e.g., installing a fire protection system in a factory plant). Loss control techniques typically involve upfront investment and ongoing cost.

- ***Loss Financing*** refers to the retention, transfer, or hedging of a risk exposure. When retaining risk, a company actively maintains a certain portion of risk. Buying insurance against a certain risk transfers the risk by providing a compensation in the case that the risk creates financial loss. Through hedging, a company takes the opposite position in a specific asset to offset the risk to another party.[17]

- ***Risk Reduction*** is about withdrawing from a risky business activity or diversifying a risk exposure by pooling different risks.

[17] From a cash flow perspective, risk transfer and hedging might be identical.

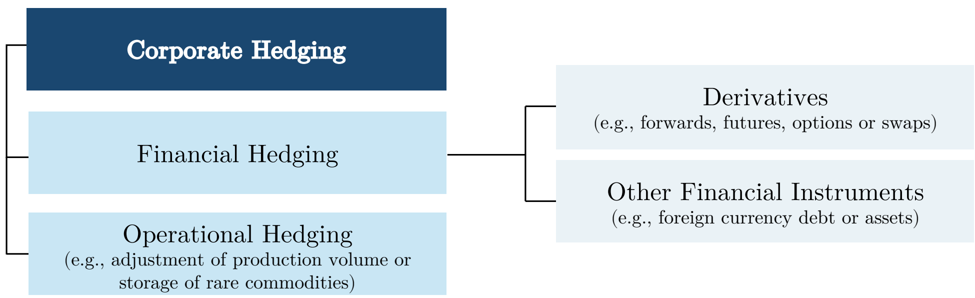
The Wharton School conducted multiple surveys concerning the risk management practices of companies (Bodnar et al., 1996, 1998). The responses of 399 U.S. non-financial companies show that risk management approaches are different across firms. This includes the goals to be achieved by risk management, but also the risk management instruments vary between firms, including derivatives hedging, diversification of the business strategy, physical storage of commodities, or cash buffers. The Wharton surveys also find that larger firms tend to manage their risk exposure more actively than smaller firms by using financial derivatives (Bodnar et al., 1996: 116; 1998: 72). In a more recent survey with 342 responses from Chief Financial Officers of companies from different countries and industries, Bodnar et al. (2019: 5004) reveal that 50% of the companies in their sample engage in risk management. When comparing public with private companies, the survey shows that 73% of the public companies employ active risk management as compared to 37% of the private companies.

3.2.3. Risk Treatment through Hedging

Hedging is one of the key means of managing risk. In general, the goal of hedging is *"to reduce risk by holding contracts or securities whose payoffs are negatively correlated with some risk exposure"* (Berk and DeMarzo, 2014: 1054). Although hedging can theoretically be used to protect against any type of risk, the most commonly hedged risk types are common market risks, including FX risk, IR risk, CP risk, and equity risk. (Bodnar et al., 1998: 73). Through hedging, a company can reduce the potential volatility of its cash flow due to unexpected changes in the market prices of FX, IR, CP, or equity.

After a company has determined its specific risk portfolio, there are different strategies for hedging the risk exposure, as illustrated in Figure 15.

Figure 15. Classification of Corporate Hedging Strategies



Notes: The figure illustrates the distinction between the different strategies of corporate risk reduction through hedging. The classification of the hedging strategies is based on Döhring (2008: 5).

Hedging can be classified as either financial or operational (Allayannis et al., 2001: 391; Kim et al., 2006: 836; Treanor et al., 2013: 66):[18]

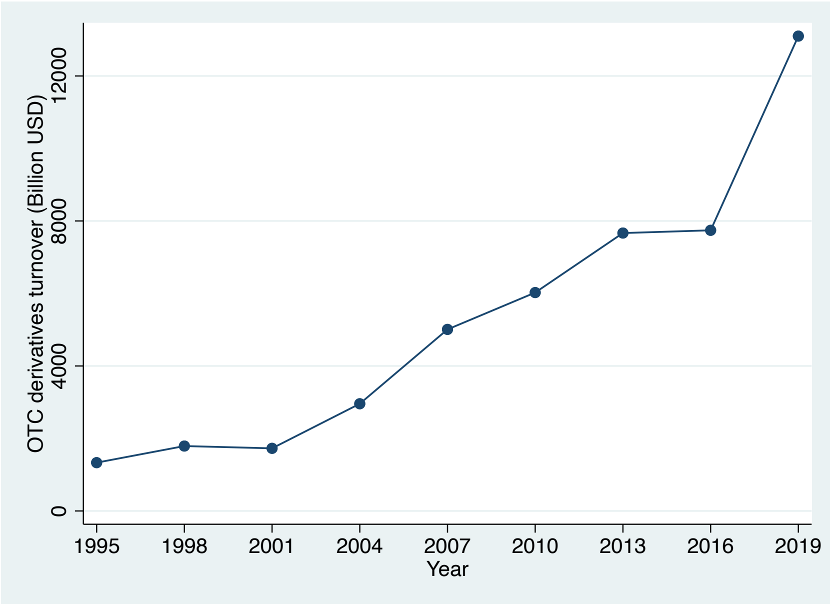
- ***Financial Hedging*** makes use of financial derivatives, such as forwards, futures, options, or swaps. Alternatively, other financial hedging instruments can be used, such as holding foreign-denominated debt in the same currency as the foreign income stream in order to match the expected cash inflows with cash outflows for debt payments (Géczy et al., 1997: 1325) or adopting conservative financial policies (Tufano, 1996: 1112).

- ***Operational Hedging*** refers to hedges by means of non-financial instruments, including the use of various real options within a company, such as flexibility in the adjustment of production or operational diversification strategies (Boyabatli and Toktay, 2004: 10-11). An example of an operational hedge would be to locate a production plant in a country with sales revenues in the foreign currency to avoid negative effects through unexpected FX changes. Operational hedging can also be achieved through production flexibility that enables a firm to adjust costs in times of volatile CP.

Financial derivatives have gained enormous importance as financial hedging instruments. Today, there are numerous financial derivatives available to hedge FX, IR, and CP risks (Giambona et al., 2018: 31 and 35). The rapid growth of international financial markets has boosted the availability of derivatives and made them a standard instrument for financial risk management. Figure 16 depicts the development of the daily average turnover of over-the-counter (OTC) FX and IR derivatives as reported by the Bank for International Settlements (BIS). The growth of the daily average turnover of FX and IR derivatives from around 2 trillion U.S. dollars in 1995 to almost 13 trillion U.S. dollars in 2019 underlines the increasing global demand for derivative instruments (BIS, 2019a, b). In addition to the analysis of aggregated data of derivatives trading, numerous studies directly survey companies from different countries and industries on their use of derivative instruments. These studies confirm the growing trend by specifically examining the volumes, types, and maturities of derivatives held by non-financials firms (among others, Alkebäck and Hagelin, 1999; Bodnar and Gebhardt, 1999; Bodnar et al., 1996, 1998; Dhanani et al., 2007; Giambona et al., 2018; Loderer and Pichler, 2000; Phillips, 1995; Prevost et al., 2000).

[18] An alternative classification divides hedging into internal instruments (e.g., matching cash inflows and cash outflows in foreign currencies) and external instruments (e.g., using derivatives to offset risk) (Buckley, 2004: 212 and 224). Furthermore, hedging methods can be classified into either on-balance sheet items (e.g., holding liquid assets to overcome cash shortfalls) or off-balance sheet items (e.g., holding derivatives for risk management) (Fok et al., 1997: 573).

Figure 16. Daily Average Derivatives Turnover



Notes: The data is based on the triennial survey by the Bank for International Settlements (BIS) among 1,300 banks and other derivative traders from 52 countries (BIS, 2016a: 3). It captures more than 90% of global activities in OTC derivatives markets. The graph shows the daily average trading volume in billions of U.S. dollars. The data points are calculated as the sum of the daily turnover on net-gross basis in the FX and IR markets (BIS, 2019a, b).

Even more difficult than collecting financial hedging data is obtaining information on operational hedging activities. Thus, it can be observed that a long line of studies defines hedgers as derivatives users (among others, Allayannis and Weston, 2001: 249; Knopf et al., 2002: 808; Lin et al., 2008: 1576). However, a simple example shows that this narrow definition of hedging can be problematic and might have a large impact on the conclusions to be drawn from the empirical results. For instance, suppose there are two firms having FX exposures from overseas sales. While the first holds a currency future contract to hedge the risk from FX changes, the second opens a production subsidiary in the same country from which it receives the sales. Hence, the latter can use cash inflows from sales to finance its foreign production plant. Both companies evidently engage in corporate hedging to reduce their risk exposure, but when defining hedgers as derivatives users, only the first firm is classified as a hedger. Accordingly, the group of derivatives users is not necessarily equal to the group of hedgers when hedging is understood as a broader concept beyond the use of financial derivatives (Berk and DeMarzo, 2014: 985).

3.2.4. Firm Value Creation through Hedging

Following the shareholder value maximization principle, the primary goal of a firm is to create value for its shareholders by maximizing the value of the firm (Lazonick and O'Sullivan, 2000: 14). The firm value, FV, can be defined as the sum of the expected free cash flow to the firm,

$FCFF$, discounted back to the present by using a discount rate κ, where κ is the firm's weighted average cost of capital (Vernimmen et al., 2009: 650):

$$FV = \sum_{t=0}^{\infty} \frac{FCFF_t}{(1+\kappa)^t}. \tag{10}$$

Risk introduces unexpected cash flow variability, which can have a negative impact on the value of a firm. Risk might also impose volatility on the equity return and the default risk. Thereby, the firm shares risk with its shareholders and debtholders through its capital structure (Berk and DeMarzo, 2014: 479-583). Both the firm and its investors, which are the shareholders and the debtholders, can reduce their exposure to risk. A firm can mitigate risk through insurance, hedging, and other risk treatment approaches as described in Section 3.2.3. Shareholders and debtholders can hold well-diversified portfolios to hedge their unsystematic risk.[19] In the extreme case, with a full hedge of all of its risk, a firm would be equivalent to a Treasury bond paying the risk-free rate (Bartram, 2019: 12). However, investors can already buy Treasury bonds, and therefore they actively invest in non-financial firms to gain higher returns while taking risk (Bartram, 2019: 12). The question arising is which risks a firm should take and which risks should not be taken to maximize the firm value?

An answer to this question is provided by traditional finance theory. According to the portfolio theory of Markowitz (1952), investors can eliminate idiosyncratic risk, which is the risk associated with individual assets, by holding well-diversified portfolios of many assets. Under the condition of optimal diversification at the shareholder level, a firm should not invest resources in risk management, because investors *"do not care about the firm-specific risk"* (Christoffersen, 2003: 2). In addition, the irrelevance theorem of Modigliani and Miller (1958) implies that under the assumption of a 'perfect capital market', there is no value to the firm from any financial transaction, including risk management and hedging. In a frictionless capital market, individual investors can hedge at the same price on their own instead of the company hedging on their behalf, because they have access to the same information and the same hedging instruments like the company itself (Aretz and Bartram, 2010: 319). In such a world of perfect capital markets, where the firm value is independent of a firm's risk structure, companies should focus all resources on the maximization of their expected profits (Brealy et al., 2011: 646).

However, the strict conditions of the Markowitz portfolio theory and the Modigliani-Miller theorem are routinely violated in practice (Allayannis and Weston, 2001: 247). In the case that

[19] As an alternative to portfolio diversification, shareholders and debtholders can use hedging instruments like a protective put to hedge against the unsystematic risk at the individual investor level (Berk and DeMarzo, 2014: 717).

the assumptions of a perfect capital market are not fulfilled, there are several channels through which hedging at the firm level may affect shareholder value and create a hedging premium.

Asymmetric Information. DeMarzo and Duffie (1991, 1995) show that a firm can hedge more effectively than its shareholders if mangers have proprietary knowledge about the firm's risk exposure. Such informational asymmetries may result from high expenses of disseminating necessary information to the shareholders, whereby the costs increase with firm complexity (Dolde and Mishra, 2007: 19). Moreover, companies may even want to maintain information asymmetry to prevent information sharing with competitors (Marshall and Weetman, 2007: 709). As corporate hedging lowers the influence of outside factors, such as unexpected price changes in FX or IR, it can be used as an instrument to overcome informational asymmetries between shareholders and managers regarding the corporate risk exposure.

Financial Distress Costs. A firm under financial distress cannot or not fully meet its financial obligations. Assuming a perfect capital market, financially distressed firms do not encounter additional costs. However, under real-world conditions, direct and indirect costs of financial distress arise (Froot et al., 1993: 1633-1634). Direct expenses occur, for example, from legal fees or increasing costs of borrowing, indirect expenses arise, for example, from employee turnover or rejecting profitable investments due to the lack of financial resources (Wruck, 1990: 436-438). Smith and Stulz (1985: 395-398) suggest that by reducing the variability of the firm value, corporate hedging can lower the likelihood of bankruptcy and thus also the expected costs associated with financial distress. Moreover, hedging can also be valuable, since a lower default probability improves the access to external debt, for example, through better credit rating or the avoidance of debt covenant[20] violations (Beatty et al., 2012: 701; Chidambaran et al., 2001: 489). Additionally, more stabilized cash flows through hedging increase the corporate debt capacity and thereby firms can profit from the tax shield of debt, which means a greater interest deduction in their tax statements (Leland, 1998: 1213).

Underinvestment and Coordination of Investment and Financing. Firms may hedge in response to agency conflicts between debtholders, who receive fixed payments for offering their capital to the firm, and shareholders, who own the residual firm value after paying off fixed obligations to the bondholders. Myers (1977: 149) states that managers acting in line with shareholders' interests have an incentive to reject projects even if they have a positive net

[20] Debt covenants are specific contract agreements and include, for example, the maintenance of a certain liquidity ratio (Chava and Roberts, 2008: 2085). Their violation can end up in penalty fees or the immediate repayment of the obligation.

present value (underinvestment), given that a firm is highly leveraged, and thus investment returns mainly accrue to bondholders. This leads to an increase in external financing costs if bondholders anticipating this conflict claim restrictive debt covenants or higher interest rates (Mayers and Smith, 1982: 287-288). Hedging can alleviate the underinvestment problem by increasing the number of future states in which shareholders receive returns from profitable investments after serving the fixed payments to the debtholders (Bessembinder, 1991: 520). A further motive for underinvestment is described by Froot et al. (1993). Their model shows that firms forgo profitable investment opportunities if no internal funds are available and external financing is more expensive than internal financing. In this case, hedging can be used to better match cash inflows and outflows to avoid costly external financing or underinvestment due to a lack of internal funds.

Tax Benefits. Another rationale for corporate hedging is provided by the tax argument introduced by Mayers and Smith (1982: 289-293) and Smith and Stulz (1985: 392-395). If a firm’s tax payments increase more than proportionally with its earnings (convex tax function), corporate hedging can smooth cash flows, such that the taxable income falls less often into the progressive region of the tax schedule, where the marginal tax rate is greater than the average tax rate paid by the firm.

In summary, in an imperfect capital market, cash flow volatility can be costly due to financial distress (Smith and Stulz, 1985: 395-399), information asymmetry between the firm and its shareholders (DeMarzo and Duffie, 1991: 264-269), external financing (Froot et al., 1993: 1638-1642), or convex tax functions (Smith and Stulz, 1985: 392-395). As hedging is an instrument to improve cash flow stability, it can reduce the costs of these market frictions, which might have a positive impact on firm value and therefore also on shareholder value.

Considering that in capital markets are neither perfect nor fully imperfect, the question remains as to which types of risk a firm should hedge. According to finance theory, companies should hedge only those risk that they can hedge at lower costs than their shareholders, assuming that the risk can actually be hedged.[21] Bartram (2019: 12) suggests that firms should take operating risks where they have competitive advantages, and hedge risks where they do not have a competitive advantage. For example, a manufacturing firm might have a competitive advantage in developing new technologies or predicting trends and consumer demand and therefore should take risk in this area. In contrast, manufacturing companies may not have a competitive advantage in predicting and speculating on market prices of FX, IR, or

[21] Not all types of corporate risk are (fully) hedgeable at the company level, such as the bankruptcy risk.

CP, and therefore they should hedge their exposure to these types of market risks if the cash flow variability through these risks entails costs to the firm.

In contrast to the various theoretical arguments on how the relaxation of the assumptions of a perfect capital market creates opportunities for corporate hedging to be beneficial for shareholders, scholars also put forward arguments that hedging could be associated with a discount in firm value. For example, MacKay and Moeller (2007: 1410) highlight that hedging is not costless and if hedging costs outweigh its benefits, it might not be valuable. In a similar vein, hedging can be associated with agency costs and monitoring problems for shareholders if managers make selective hedging decisions to protect their individual interests or to increase the risk exposure for speculative purposes (Brown et al., 2006: 2926; Smith and Stulz, 1985: 309-403; Tufano, 1996: 1109-1111). Traditional hedging theories rely on the assumption that firms engage in hedging to reduce their risk exposure (Faulkender, 2005: 935; Hentschel and Kothari, 2001: 93). However, there may also exist incentives for managers to actively time derivatives transactions based on their market view, a practice called 'market timing' or 'selective hedging' (Adam et al., 2017: 269; Brown et al., 2006: 2926-2928). Assuming efficient capital markets without informational asymmetries, the gains from market timing should be small or even negative when accounting for transaction costs arising from hedging, especially from buying or selling derivatives. Stulz (1996) sees two criteria as necessary prerequisites for selective hedging to be valuable. First, firms need private information about future market development and this information should not be available to other competitors (informational advantage) (Stulz, 1996: 15). Second, the benefits from selective hedging must exceed shareholders' expected rate of return for alternative investments with comparable risk (economic profit) (Stulz, 1996: 23). Due to the different goals that can be achieved through financial derivatives (hedging vs. speculation), it is an important question in finance research whether empirical studies on corporate hedging measure risk reduction or market timing.

3.3. Related Empirical Literature

While the previous section discussed the corporate hedging theories, this section reviews the findings of the previous empirical literature testing these theories in both primary and secondary research. As the meta-analysis to be presented in the subsequent sections of this chapter represents a comprehensive quantitative review of the empirical literature in the field, I focus in this section on a qualitative summary of some of the key findings from the previous empirical studies.

3.3.1. Primary Research on Hedging and Firm Value

In their seminal article, Allayannis and Weston (2001) analyze a sample of 720 U.S. firms and find that the average increase in firm value measured by Tobin's Q[22], i.e. the hedging premium, is about 5% for firms that use FX derivatives as compared to non-users. Beginning with this first empirical evidence, a comprehensive literature stream on the value impact of hedging developed. Many succeeding studies build on the study design of Allayannis and Weston (2001) but change one or more aspects in terms of data and applied methods.

A first aspect of the differences between the previous studies relates to the examined industry and connected with that the type of risk exposure that is investigated. For example, for U.S. high-tech firms, Gleason et al. (2005) show that financial hedging of FX risk is value-increasing. Similarly, Carter et al. (2006) observe that financial derivatives usage for jet fuel hedging enhances firm value in the airline industry by more than 10%. For pharmaceutical and biotech firms, Choi et al. (2013) find that the use of financial derivatives for hedging is related to greater firm value and that the value increase is larger for firms that are subject to greater information asymmetry and better growth opportunities. In contrast to the sector-specific findings for a positive hedging premium, Jin and Jorion (2006) provide empirical evidence that hedging is not rewarded by higher valuations in the oil and gas industry.

Another dimension by which the primary literature can be clustered is the country analyzed in the data set. Allayannis et al. (2012) present an international study across 39 countries with a sample of firms that have a significant FX exposure. The authors report strong evidence for both internal firm-level and country-level governance factors being associated with significant hedging premiums. In single country studies, Clark and Judge (2009), Gómez-González et al. (2012), Búa et al. (2013), Jankensgård (2015a), and Bae et al. (2018) reveal a value-enhancing effect of corporate financial hedging for firms in the United Kingdom, Colombia, Spain, Sweden, and South Korea, respectively. The opposite effect, i.e., a negative or zero value impact of hedging, is reported by Nguyen and Faff (2007), Khediri (2010), Li et al. (2014), Ayturk et al. (2016), and dos Santos et al. (2017) for companies in Australia, France, New Zealand, Turkey, and Brazil.

The prior literature also addresses various methodological challenges occurring in the empirical analysis of the hedging and firm value nexus. Much discussion evolved as both variables, corporate hedging and firm value, are endogenous (Aretz and Bartram, 2010: 362). In other words, hedging can affect firm value, but firm value can also affect hedging. To control for the issue of endogeneity, Bartram et al. (2011) match derivatives users and non-users

[22] Tobin's Q is defined by the ratio of the market value of financial claims and the replacement cost of the firm's assets (Kaldor, 1966: 317).

depending on their estimated propensity to use derivatives. They document a positive effect of derivatives use on firm value, which is, however, sensitive to endogeneity and omitted-variable concerns. In another study, Pérez-González and Yun (2013) exploit the introduction of weather derivatives as an exogenous shock to the firm’s ability to hedge weather risk. They apply this natural experiment to control for endogeneity and conclude that derivatives usage leads to higher firm valuations. Other studies implement an IV approach where they ‘instrument’ the use of hedging with variables that are likely to affect hedging but have no effect on firm value (among others, Allayannis et al., 2012; Fauver and Naranjo, 2010; Kwong, 2016; Phan et al., 2014).

Another aspect that is intensively discussed in the literature is the differentiation between hedging and derivatives use. Many studies in the hedging literature assume that the use of derivatives, as a strategy of financial hedging, is equivalent to hedging in general. More recent research considers the existence of other hedging mechanisms beyond financial hedging, such as operational hedging strategies (Almeida et al., 2017; Chod et al., 2010; Hankins, 2009; Hoberg and Moon, 2017). For example, Allayannis et al. (2001) find that shareholders benefit from operational hedging strategies only when used in combination with financial hedging strategies. In a similar vein, Kim et al. (2014) report that both operational and financial hedging activities are valuable for non-family firms, but do not create value in family firms.

3.3.2. Secondary Research on Hedging and Firm Value

There is a group of related meta-analyses that synthesizes empirical studies examining the determinants of corporate hedging decisions (Arnold et al., 2014; Geyer-Klingeberg et al., 2019; Geyer-Klingeberg et al., 2018a). In contrast to these studies that examine why companies hedge (rationales for hedging), this chapter focuses on the question if hedging is actually valuable for firms (financial impact of hedging).

In addition to the prior meta-analyses on the hedging determinants, a series of authors present qualitative reviews of the corporate hedging literature and its impact on firm value (Ammon, 1998; Aretz and Bartram, 2010; Judge, 2007; Krause and Tse, 2016; Ramlall, 2010; Sahoo, 2015; Triki, 2005). These studies aggregate the available research record using a descriptive procedure but without presenting a statistical integration of the empirical findings reported in the primary studies. Despite the detailed and critical discussions in these studies, such narrative reviews are often subject to several limitations because their conclusions depend on the authors' judgments rather than on a statistical integration of the effect size estimates reported in the primary studies (Stanley, 2001: 144). Accordingly, narrative reviews are widely considered not to be the preferred method of research synthesis (Borenstein et al., 2009: 14).

The study closest to this chapter is by Bessler et al. (2019), who present a meta-analysis of 47 primary studies. The authors use the traditional meta-analysis approach to aggregate and compare correlation coefficients between corporate hedging behavior and Tobin's Q. As a key result, they find a statistically significant, but small mean correlation coefficient of 0.044. In their heterogeneity analysis, hedging of FX risk is found to be consistently associated with higher shareholder value, while there is no clear evidence for hedging of IR and CP risks.

The meta-analysis approach and the implied results presented in this chapter extend the study by Bessler et al. (2019), especially with regard to three aspects. First, the effect size to be aggregated in this meta-analysis is the actual hedging premium. A key advantage of the hedging premium over correlation coefficients, as used by Bessler et al. (2019), is that it allows for interpreting the economic magnitude of the accumulated effects and not only the statistical relation between the two variables.[23] Second, Bessler et al. (2019) follow the meta-analysis method by Hunter and Schmidt (2004) to calculate weighted mean correlations for different subgroups depending on the data and methods used in the primary studies. In other words, this approach spotlights single contingency factors of the hedging and firm value link. In contrast, in this chapter I use meta-regression analysis to simultaneously model the impact of a wide set of moderator variables in a multiple regression framework that accounts for the interrelations among the variables. Meta-regression also has the advantage of explicitly controlling and correcting for publication selection. A third distinction stems from the variables analyzed as moderators of the hedging premium. As an extension of the work by Bessler et al. (2019), I explore several new directions, especially regarding the estimation and model characteristics of the primary studies, as well as the macroeconomic differences between the countries examined in the primary data samples.

Another related study was published by Hang et al. (2021a). The authors analyze the interaction between capital structure decisions and risk management decisions, as well as the channels through which they add value to firms using the MASEM methodology. By incorporating 6,312 reported empirical results from 411 primary studies and testing competing theories in a meta-path model, they find that there is no significant impact of the corporate hedging decision on firm value. Although this review is based on advanced meta-analysis methods, I see several limitations in this study. First, building on weighted averages across a broad set of articles requires that the observed effects be comparable among these studies in order to avoid mixing up apples with oranges (Card, 2012: 25). However, research differs in terms of the examined countries, risk exposures, and econometric estimation methods. The

[23] See also Section 2.5 for a discussion on economic as compared to statistical effect sizes.

diversity among the studies is underlined by the statistically significant results from the heterogeneity test reported by Hang et al. (2021a: 4911).[24] Pooling up estimates from such a heterogeneous sample without controlling for the differences in study design is hard to justify. In contrast, the meta-regression approach applied in this chapter explicitly models the sources of heterogeneity among the primary studies, thereby providing evidence on why individual studies reach different conclusions. Another issue is that the study by Hang et al. (2021a) uses effect sizes estimated from the univariate empirical results reported in the primary studies. However, the relation between hedging and firm value could be determined by the interactions among firm value and hedging determinants, as well as many other control variables. In contrast, I collect estimates from the multivariate analysis sections of the primary studies. Thus, the derived effect sizes measure the relationship between hedging and firm value while controlling for other important variables. Another limitation of the meta-analysis by Hang et al. (2021a) appears as a result of their focus on the impact of certain firm characteristics on the hedging decision. Accordingly, they aggregate results from models where the dependent variable is a hedging dummy variable indicating whether a specific firm hedges or not. However, literature shows that beyond the decision to engage in hedging, the extent of hedging is also an important aspect of the corporate risk management policy (Allayannis and Ofek, 2001: 289; Haushalter, 2000: 137). In this chapter, I consider both aspects, the decision to hedge and the extent of hedging.

The objective of this chapter is to overcome the identified problems of prior reviews and to use the power of meta-regression analysis to explain the sources of heterogeneity among the previously published literature as well as to derive the mean hedging premium implied by the existing research record.

3.4. Sample Selection

This section describes the collection and preparation of the meta-analysis data set. Moreover, the extraction of the relevant information from the primary studies and the calculation of the hedging premium from the observable primary study results are presented.

3.4.1. Quality Assurance

The literature search and the subsequent analyses are in line with the guidelines laid out by MAER-Net (Havranek et al., 2020; Stanley et al., 2013). One deviation from the recommendations in the guidelines is that the coding of the primary studies was performed by

[24] The study rejects the null hypothesis of homogenous effect sizes from the Cochrane's Q-test at any conventional level of statistical significance. Moreover, the applied random effects models show a significant proportion of between-study variance.

only one person, rather than two, as suggested in the guidelines. However, the coding protocol was developed with a second researcher and ambiguities that occurred in the coding process or the calculation of the effect sizes were discussed and solved together. Because the majority of coded data for this meta-analysis are directly evident from the results sections of the primary studies and no subjective judgments are required, coding errors should be random and therefore be absorbed by the error term of the meta-regression model.

3.4.2. Inclusion and Exclusion Criteria

Before collecting a sample of primary studies, the inclusion and exclusion criteria must be defined. These criteria are important to decide which of the identified publications should be added to the final database for the meta-analysis. In this step, it is also necessary to ensure that the collected effect size estimates are comparable within and between studies in such a way that differences can be coded by the moderator variables.

A primary study must meet the following criteria to be added to the data sample:

Design of the Empirical Analysis. To be included in the sample, effect size estimates from a regression analysis must be reported. In the regression model, the dependent variable needs to be a measure of firm value and one of the explanatory variables needs to be a measure of corporate hedging. Studies are excluded where the explanatory variable is a measure for the general presence of an ERM program. Firms with ERM activities are not necessary hedgers, as ERM can also be achieved by other risk treatment methods (Section 3.2.2, Figure 14).

Non-Financial Firms. The data sample of the primary study must refer to non-financial companies. This is a common procedure in the hedging literature, which is validated by the fact that most financial firms, especially banks, are also market makers in derivatives markets (Allayannis and Weston, 2001: 248). Hence, their rationales for holding derivatives are usually different from those of non-financial firms.

Reported Empirical Results. Studies must provide the statistical information required for meta-regression analysis. This includes regression coefficients, sample sizes, and a precision measure of the regression estimates such as standard errors, *t*-statistics, or *p*-values. Information on the study characteristics, data sample, and applied methods must be reported in a way that allows for the precise coding of the moderator variables.

Publication Outlet. To avoid ex-ante selection bias, for example by focusing only on journal articles, primary studies are included independent of their publication status. The meta-regression explicitly controls for variation in study quality. Therefore, the included primary studies can either be articles published in referred journals or unpublished manuscripts and conference papers. The inclusion of unpublished work is preferable in several ways. First, it allows to capture recent studies that have not yet gone through the refereeing process. Thus, their exclusion often reduces the observed time span in the meta-analysis, which may affect the possibility of identifying whether the heterogeneity among the hedging premiums is driven by time trends. Second, primary studies might be unpublished because their results are not in accordance with mainstream literature or since they fail to find statistical significance. Some researchers even suggest that the inclusion of grey literature can alleviate the problem of biases from publication selection (Duran et al., 2015: 1245). However, other meta-analysts argue that in the case that journals prefer certain results, authors will adopt this selection preference already when preparing their studies for submission and sharing manuscripts as working papers (Rusnak et al., 2013: 40). Due to the opposing arguments, this chapter includes unpublished work but also controls for the impact of publication selection in the meta-regression model.[25] A third reason for the inclusion of unpublished work is that focusing only on top journals would lead to a significantly smaller data set and a reduction of variation in the collected estimates for the hedging premium. Such variation, however, is necessary to unfold the power of meta-regression to identify the sources of heterogeneity.

Single Country Studies. Most primary studies in the hedging and firm value literature examines firms in a specific country. If an estimate for the hedging premium refers to one specific country, I can clearly assign a value of the country-level factors analyzed as drivers of heterogeneity to respective the hedging premium. Therefore, I drop observations in the primary studies referring to multi-country data samples in order to better explore the real differences between countries. For multi-country data samples, especially studies covering world samples, I could only assign average values for the country-level variables to the hedging premiums. However, such average values could neglect important differences between the countries.

3.4.3. Search Strategy

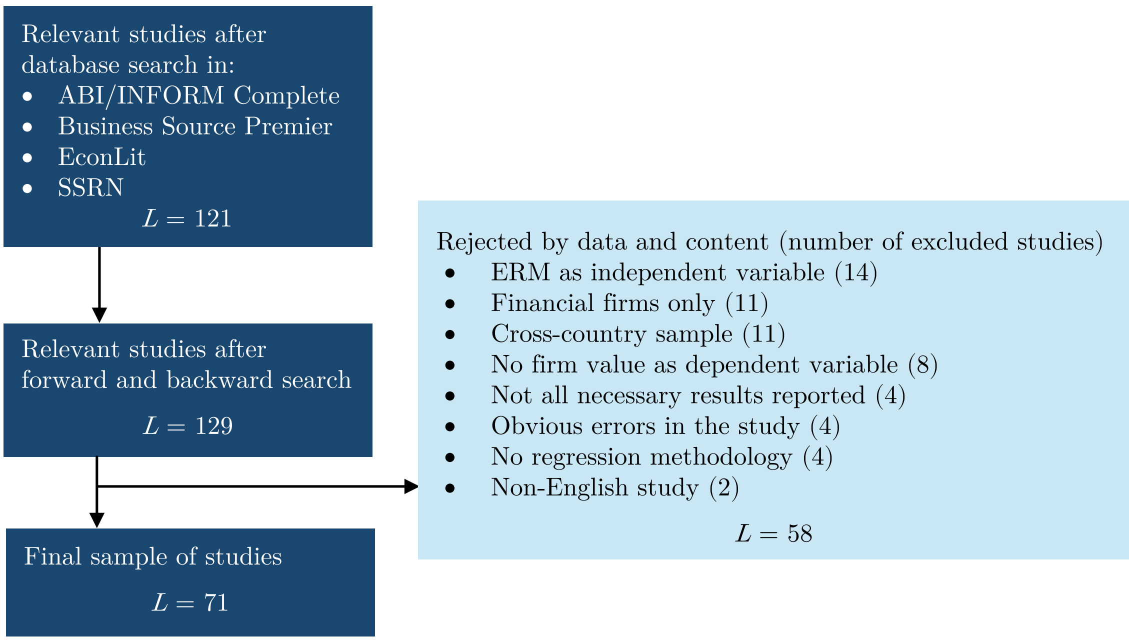
To identify the empirical studies analyzing the impact of hedging on firm value, I performed a keyword search in the following electronic databases: ABI/INFORM Complete via Proquest, Business Source Premier via EBSCOhost, EconLit via EBSCOhost, and the working paper database of the Social Science Research Network. The search string consists of keywords linking important terms for hedging (‘hedging’, ‘hedger’, ‘risk management’, ‘derivatives’, ‘option’,

[25] One exception to the inclusion of grey literature is that bachelor’s and master’s theses are not considered in the sample.

'swap', 'forward', 'future') and firm value ('firm value', 'premium', 'Tobin's Q', 'market-to-book ratio'). During the database search, I manually reviewed 912 database hits, resulting in the selection of 121 potentially relevant studies after reading their titles, abstracts, and results sections and filtering them against the selection criteria defined in the previous section.

To find studies that do not contain the required keywords, a snowballing technique was conducted after the database search. This strategy involves a backward search in the reference lists of the 121 previously identified articles and a forward-tracking of all 3,068[26] studies listed as their citations in Google Scholar. The last study was added in March 2018. Figure 17 outlines the search strategy and the number of excluded studies in each step.

Figure 17. Summary of the Literature Search Process



Notes: This figure depicts the database search process and the number of studies added to the sample or rejected in each step.

After the database and backward/forward searches, the 129 relevant studies were read in detail and filtered against the inclusion and exclusion criteria defined in the previous section. The application of the selection criteria produces a final sample of 71 primary studies published between 2001 and 2018 that examine the impact of corporate financial hedging on firm value. The full list of included primary studies is reported in Appendix B.1. The 71 primary studies report 1,016 regression estimates for the hedging and firm value nexus that can be converted into hedging premiums as outlined in the next section.

[26] This number includes duplicates.

3.4.4. Effect Size Calculation

The studies in my sample use several variants of the following baseline regression model to measure the impact of corporate hedging on firm valuations:

$$Q_{ht} = \alpha_0 + \alpha_1 H_{ht} + \sum_{c=1}^{C} \gamma_c X_{c,ht} + \eta_h + \zeta_t + u_{ht}, \tag{11}$$

where $h = 1, \dots, N$ and $t = 1, \dots, T$ are firm and time subscripts, Q is a measure of firm value, H is a measure of corporate hedging activities, X represents a set of $c = 1, \dots, C$ control variables, and u is the error term. The coefficients η_h and ζ_t capture firm-specific and time-specific effects.[27]

The parameter of interest in this meta-analysis is $\hat{\alpha}_1$, which is the estimated regression coefficient of the hedging variable and can be interpreted as the firm value premium or discount of hedging. As authors use different functional forms and model specifications of Eq. (11), the collected estimates from the set of primary studies are not directly comparable. A first distinction relates to the dependent variable. There are primary studies estimating the model in a level-level specification as shown above. However, the majority of studies apply a log-level specification, i.e., they use the natural logarithm of Q as the dependent variable. A second distinction regards the definition of the hedging variable H, which can be a 'dummy' variable (denoted HD). This variable is equal to one for hedging firms and zero for non-hedgers. Alternatively, the hedging variable can be a 'continuous' measure quantifying the actual hedging volume (denoted HC).

As meta-analysis requires the underlying effect sizes to be comparable, I transform the reported estimates, $\hat{\alpha}_1$, such that they measure the hedging premium, HP. The hedging premium quantifies the average percentage markup or discount in firm value for hedging firms compared to non-hedgers (if H is a dummy variable HD) or the percentage markup or discount of a firm with an average hedge ratio compared to non-hedgers (if H is a continuous variable HC). For the calculation of the hedging premiums, I use the estimated regression coefficients from Eq. (11) and the descriptive statistics reported in the primary studies. To account for the fact that the hedging premium is estimated with error, I calculate the corresponding standard error, $SE(HP)$, to capture precision.

From the sample of collected primary studies, I extract the marginal effects of hedging on firm value as well as the corresponding measure of precision, which are mostly standard errors or *t*-statistics. In addition, I observe the type of the hedging variable (HD or HC) and the specification of the model (level-level or log-level). Finally, sample mean values of the firm

[27] Cross-country studies may also include country fixed effects.

value, $\bar{Q}$, and the continuous hedging variable, $\overline{HC}$, are obtained from the descriptive statistics reported in the primary studies. Table 5 summarizes the calculation of the hedging premiums.

Table 5. Computation of Hedging Premiums

	Q	$ln(Q)$	Observations
$HD \in \{0,1\}$	(1) $HP = \hat{\alpha}_1/\bar{Q}_{NH}$	(2) $HP = exp(\hat{\alpha}_1) - 1$	628
$HC \in [0,1]$	(3) $HP = \hat{\alpha}_1(\overline{HC}_H/\bar{Q}_{NH})$	(4) $HP = \exp(\hat{\alpha}_1 \times \overline{HC}_H) - 1$	259
Observations	302	585	887

Notes: HD = Hedging dummy variable, HC = Continuous hedging variable, HP = Hedging premium, $\hat{\alpha}_1$ = Estimated marginal effect of hedging on firm value, Q = Measure of firm value, H = Group of hedging firms, NH = Group of non-hedgers.

The hedging premium quantifies the difference between the average firm value of the group of hedgers and non-hedgers in relation to the firm value of non-hedgers: $(\bar{Q}_H - \bar{Q}_{NH})/\bar{Q}_{NH}$. In the first case in Table 5 (HD, Q), the observed regression coefficient, $\hat{\alpha}_1$, measures the firm value differences between hedgers and non-hedgers. Hence, $\hat{\alpha}_1$ must be divided by the sample mean of the non-hedgers group, $\bar{Q}_{NH}$, to receive the percentage value increase through hedging. If the sample mean was not reported in the primary study, I requested it from the authors. Otherwise, I use the full sample mean of firm value (hedgers and non-hedgers) as a proxy. For the second case (HD, $ln(Q)$), the estimated regression coefficient quantifies the percentage markup in a logarithmic scale. Thus, the exponential value of the reported regression coefficient minus one equals the hedging premium. For case three (HC, Q) and case four (HC, $ln(Q)$), I follow Carter et al. (2006: 74) as well as Phan et al. (2014: 348) and evaluate the value premium for an average hedging firm by multiplying the primary study regression coefficients with the sample mean of the continuous hedging variable for the group of hedgers, $\overline{HC}_H$. Afterwards, I conduct the same transformations as for case one and case two.

The standard errors of the hedging premiums are calculated using the t-statistics reported in the primary studies:

$$SE(HP) = \frac{HP}{t}, \tag{12}$$

where t is the reported t-statistic of $\hat{\alpha}_1$.

As an extension of Eq. (11), about 10% of the estimates $\hat{\alpha}_1$ represent interactions of the hedging variable with other firm characteristics, such as the corporate capital expenditures. For these estimates, I follow Havranek et al. (2016: 136) and evaluate the interaction term at the sample mean of the interacting variable to calculate the hedging premium. A regression model with one interaction can be described as:

$$Q_{ht} = \alpha_0 + \alpha_1 H_{ht} + \delta_1 H_{ht} INT_{ht} + \sum_{c=1}^{C} \gamma_c X_{c,ht} + \eta_h + \zeta_t + u_{ht}, \tag{13}$$

where INT denotes an interaction variable. Other variables and subscripts are the same as in Eq. (11). In the case of interaction terms, the hedging premiums are evaluated at the sample mean of the interacting variable. If the sample means of the interacting variables are unreported, I asked the primary study authors to provide them. Otherwise, these effects are not considered in the sample. The calculation of the hedging premiums for the four cases with one interacting variable is summarized in Table 6.[28] The corresponding standard errors for the hedging premiums are approximated using the delta method (Papke and Wooldridge, 2005: 415-416; Valentine, 1979: 364).

Table 6. Computation of Hedging Premiums in Models with Interaction Terms

	Q	$ln(Q)$	Observations
$HD \times INT$	(1) $HP = \hat{\alpha}_1/\bar{Q}_{NH} + \hat{\delta}_1(\overline{INT}/\bar{Q}_{NH})$	(3) $HP = exp(\hat{\alpha}_1) - 1 + exp\left(\hat{\delta}_1 \times \overline{INT}\right) - 1$	62
$HC \times INT$	(2) $HP = \hat{\alpha}_1(\overline{HC}_H/\bar{Q}_{NH}) + \hat{\delta}_1(\overline{HC}_H \times \overline{INT}/\bar{Q}_{NH})$	(4) $HP = exp(\hat{\alpha}_1 \times \overline{HC}_H) - 1 + exp\left(\hat{\delta}_1 \times \overline{HC}_H \times \overline{INT}\right) - 1$	67
Observations	51	78	129

Notes: HD = Hedging dummy variable, HC = Continuous hedging variable, HP = Hedging premium, $\hat{\alpha}_1$ = Estimated marginal effect of hedging on firm value, Q = Measure of firm value, H = Group of hedging firms, NH = Group of non-hedgers, INT = Interaction variable. The observation counts refer to the total number of observations with interaction terms for this group, i.e., also observations with more than one interaction term are included.

Studies routinely report multiple estimates for the firm value effects of hedging, for example, for different risk exposures, alternative model specifications, or other robustness tests. I adhere to a common best practice in meta-analysis research and sample all available estimates (among others, Feld et al., 2013: 2851-2852; Kysucky and Norden, 2016: 93; Rusnak et al., 2013: 41). Such multiple sampling approach increases the power of meta-analysis tests and enhances the accuracy of estimates due to the larger underlying sample. Bijmolt and Pieters (2001: 162-166) show in a Monte Carlo simulation that meta-analysis approaches that include the complete set of estimates per study outperform procedures which represent a study by a single estimate. Moreover, selecting just one estimate per study requires objective selection rules to decide

[28] The calculation of the hedging premiums is analogously performed for models with more than one interaction term (33 observations in total).

which estimate to favor. In addition, sampling only one estimate per study leads to a loss of information regarding within-study variation (Stanley and Doucouliagos, 2012: 32).

As the inclusion of multiple estimates per study violates the assumption of statistical independence, I explicitly control for data clustering at the study level and the between-study level in the estimation of the meta-regression model (Section 3.5.2). Moreover, I follow Kysucky and Norden (2016: 94) and 'winsorize' extreme observations of the effect size estimates at the 1% and 99% quantiles to reduce an undue influence of outliers of the hedging premium.[29]

3.5. Methodology

This section describes the applied meta-analysis methods and model specifications for the analysis of publication selection bias, heterogeneity analysis, and the estimation of the mean hedging premium implied by the empirical literature.

3.5.1. Meta-Regression Model

I employ meta-regression analysis to derive generalizations about the hedging premiums, to detect distorting biases in the literature, and to explore the conditional factors that drive variation within and between studies. The methodology follows that presented in Section 2.2.2.

Analysis of Publication Selection Bias.[30] Publication selection bias arises when researchers, editors, and reviewers omit undesired results from publication (Begg and Berlin, 1988: 419; Rothstein et al., 2005: 1-3; Stanley, 2005: 310-311). Undesirable outcomes might be statistically insignificant effects, results without support of the ex-ante hypothesis, results that are inconsistent with theoretical predictions, or results that do not agree with what was found in the previous empirical literature. If uncontrolled, an active selection of preferred results can lead to biases in the aggregated effects inferred by meta-analysis (Stanley, 2008: 104).

Without publication selection, the estimates of the hedging premium and their standard errors should be independent quantities (Egger et al., 1997: 629-630). However, if authors seek to reach conventional significance levels (commonly, t-ratios > 2), estimates could remain unreported if they are not large enough to offset standard errors (commonly, estimates lower than twice their standard errors). If publication selection is present in the literature, estimates for the hedging premium and their standard errors will be correlated.

[29] There are other procedures to detect and correct outliers, such as DFBETAS (Bollen and Jackman, 1985: 518-519), that might outperform the winsorizing approach.
[30] See also Section 2.5.

Following Card and Krueger (1995: 239), a statistical test to investigate the presence of a relation between the effect size estimates and their standard errors is given by:

$$HP_{ij} = \beta_0 + \beta_1 SE(HP_{ij}) + e_{ij}, \text{ with } e_{ij} \sim N(0, SE(HP_{ij})^2), \tag{14}$$

where i and j are study and estimate subscripts, HP_{ij} are the estimates of the hedging premium calculated from the empirical results reported in the primary studies, $SE(HP_{ij})$ is the standard error of the respective estimate for the hedging premium, which is included to control for publication selection, and e_{ij} is the error term.

According to Egger et al. (1997: 629-630), the rejection of the null hypothesis, H_0: $\beta_1 = 0$, tests for the presence of publication selection bias. The estimate for the intercept, $\hat{\beta}_0$, is the mean hedging premium assuming that $SE(HP)$ is close to zero, $SE(HP) \to 0, E(HP) \to \beta_0$. Thus, rejecting the null hypothesis, H_0: $\beta_0 = 0$, is a test for the existence of a genuine effect beyond publication selection (Stanley, 2008: 108).

Heterogeneity Analysis. For the heterogeneity analysis, the estimates of the hedging premium are regressed on a set of variables measuring differences in methods, data, and other aspects of study design (Stanley et al., 2008: 283):

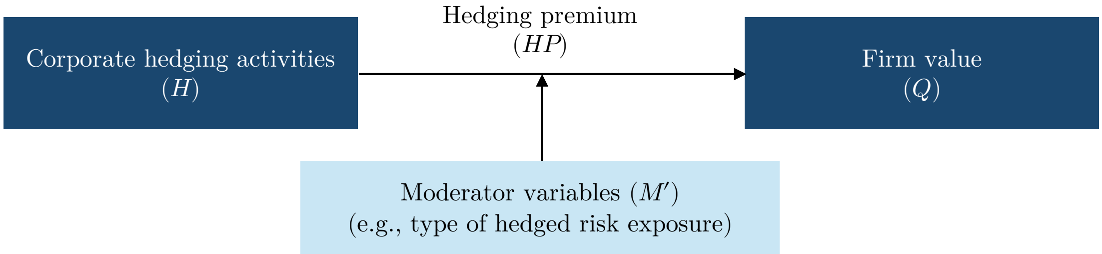
$$\begin{gathered} HP_{ij} = \beta_0 + \beta_1 SE(HP_{ij}) + \sum\nolimits_{l_1=2}^{P_1+1} \beta_{l_1} M_{l_1,ij} + \sum\nolimits_{l_2=\mathrm{P}_1+2}^{P_2+P_1+1} \beta_{l_2} M'_{l_2,ij} + e_{ij}, \\ \text{with } e_{ij} \sim N(0, SE(HP_{ij})^2), \end{gathered} \tag{15}$$

where, in addition to Eq. (14), $M_{l_1,ij}$ describes the $P_1 + 1$ dummy variables measuring the presence or absence of important control variables in the primary regression (the X variables in Eq. 11). $M'_{l_2,ij}$ represents a collection of $P_2 + P_1 + 1$ meta-explanatory variables that capture relevant study characteristics and explain the variation between the effect size estimates related to differences in data and methods.[31]

The estimated meta-regression coefficients, $\hat{\beta}_{l_1}$, measure the sensitivity of the examined effect to changes in the variables of the primary regression model and the omission of important control variables (misspecification bias). $\hat{\beta}_{l_2}$ reflect the average effect of a given study characteristic on the effect size estimates. Accordingly, the explanatory variables $M'_{l_2,ij}$ can be interpreted as moderators of the effect sizes HP_{ij}, as illustrated in Figure 18.

[31] Some of the independent meta-variables might be study-invariant (like the publication year of the study). In this case, the values of these moderators are constant in the same study.

Figure 18. The Moderating Effect of Meta-Regressors



Notes: The figure illustrates the impact of the moderator variables (M') on the effect size, which is the hedging premium (HP).

The intercept, β_0, represents the mean hedging premium conditional on $SE = 0$ and the vectors $\boldsymbol{M} = 0$ and $\boldsymbol{M'} = 0$. As this is the implied hedging premium for a rather specific scenario where all meta-regressors are equal to zero, I substitute more plausible 'best practice' values for the meta-regressors in Section 3.7.5. This creates an average hedging premium implied by the specification of a synthetic best practice study (Stanley and Doucouliagos, 2012: 93). However, it should be noted that such an approach is inevitably subjective, depending on the definition of a best practice study.

3.5.2. Model Specification

For the application of the meta-regressions described in Eqs. (14) and (15), I consider the following model specifications.

Heteroscedasticity. Meta-regression models commonly exhibit heteroscedasticity, as the estimates' standard errors depend on the sample size and the sample size usually varies from study to study. It is an established approach in meta-regression research to use WLS regression to obtain efficient estimates in the presence of heteroscedastic errors (Stanley and Doucouliagos, 2012: 110). The optimal weight is the reciprocal of the squared standard errors of the effect size estimates. This implies that more precise and thus more reliable statistical estimates (those with lower standard errors) receive a larger weight in the regression.

Indeed, there is a debate about inverse variance weighting (Zigraiova and Havranek, 2016: 971-974), as it indirectly puts larger weights on studies reporting more estimates. To avoid such unintentional weighting of studies with many estimates, I also employ the number of estimates per study as well as the interaction between the inverse of the number of estimates per study and the inverse of the estimates' variance as alternative weights. The two alternative weights assign equal importance to studies independently of the number of reported estimates. Chapter 4 investigates the weighting schemes in meta-regression analysis in more detail and compares the statistical properties of the three weighting schemes in a Monte Carlo simulation.

Within-Study Dependency. To maximize data availability and to avoid biases arising from a subjective sample selection, I include all estimates for the hedging premium calculated from the results reported in each of the 71 primary studies. By implication, standard errors are likely to be inflated in a pooled cross-study regression because of the dependence structure, especially at the study level. This point is equivalent to the issue of correlated errors in panel data regressions across multiple firms and time periods (Petersen, 2008: 435-436). To control for within-study dependency, I follow a common best practice in meta-regression research (Stanley and Doucouliagos, 2012: 71) and adopt cluster-robust standard errors (Froot, 1989) with clusters at the level of the individual studies.

Between-Country Dependency. The clustering at the study level presumes the clusters themselves to be independent. As the data samples used in different primary studies may overlap, the assumption of cluster-independence is potentially violated. To consider between-study dependencies, I treat data sets from different studies as similar when they examine the same country, for example, when two studies analyze the firm value implications through hedging in a group of U.S. firms, e.g., the S&P 500 companies. Besides the study-level clustering, standard errors are additionally clustered at the country level following the two-way clustering approach by Cameron et al. (2011).[32]

3.6. Moderator Variables

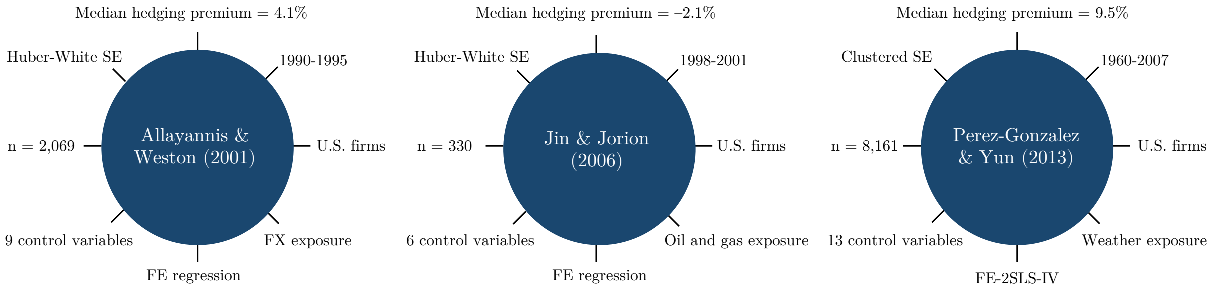
This section is devoted to the selection of the moderator variables that are used as explanatory variables in the meta-regression model (vectors $\boldsymbol{M}$ and $\boldsymbol{M'}$ in Eq. 15). The analysis of the moderator variables allows the identification of the sources of heterogeneity among the hedging premiums and the quantification of the direction and magnitude of the impact of these variables on the hedging premiums.

3.6.1. Heterogeneity in the Hedging and Firm Value Literature

Figure 19 shows the study characteristics of three seminal articles examining the impact of corporate financial hedging on corporate firm value. To illustrate the wide heterogeneity that is present in the literature, a selection of the main study characteristics is compared across these three primary studies. The analyzed study characteristics include the main results in terms of the median hedging premium per study, as well as key study design and data characteristics.

[32] For a recent application of two-way clustering in meta-analysis, see also Havranek and Irsova (2017).

Figure 19. Illustration of Heterogeneity in the Corporate Hedging Literature



Notes: This figure illustrates three seminal primary studies examining the impact of corporate financial hedging on firm value. The median hedging premium is the median of all hedging premiums per study that are recomputed from the reported results. The other aspects capture some of the key data and method choices by the study authors. n = Sample size, FE = Fixed effects regression, 2SLS = Two-stage least squares, IV = Instrumental variable, SE = Standard errors.

It becomes readily apparent that the median of the hedging premiums reported per study, i.e., the percentage surplus in firm value of hedging companies against non-hedgers, ranges from a negative hedging discount of -2.1% to a larger positive hedging premium of 9.5%. Hence, it seems that the literature does not agree on the general firm value impact of hedging. The wide range of hedging premiums coincides with diverse study characteristics. For example, Allayannis and Weston (2001) focus on hedgers of FX volatility, while Pérez-González and Yun (2013) investigate the firm value effects arising from hedging of weather risk exposures. Moreover, the sample size, and hence the power of the study to detect the true effect, varies between 330 observations and 8,161 observations. The econometric approaches are also diverse. Pérez-González and Yun (2013) control for endogeneity by a two-stage least squares IV approach, while Allayannis and Weston (2001) and Jin and Jorion (2006) apply fixed effects panel regression. In summary, just three of the 71 primary studies in the sample illustrate the large heterogeneity in empirical results and study designs. Such large heterogeneity tends to be the norm in many research areas in economics and finance (Stanley and Doucouliagos, 2012: 80-82).

The sources of heterogeneity can be divided into two main groups (Havranek and Irsova, 2011: 238; Rusnak et al., 2013: 54). First, estimates might vary due to the applied methodology and specific data characteristics, which is referred to as 'methodological and data-related heterogeneity' (Section 3.6.2). The second group of moderators in this study covers real differences among the hedging premiums. For example, the corporate hedging behavior of U.S. firms might differ from that of European firms. This group of heterogeneity is known as 'structural heterogeneity' (Section 3.6.3). A wide range of moderator variables were manually coded to quantify the effects of both types of heterogeneity. The selection of the variables is driven by data availability from the primary studies and by discussions in the literature.

3.6.2. Data and Method Choices

Table 7 provides an overview of the methodological aspects and data characteristics of the primary studies that I expect to influence the size and direction of the estimates for the hedging premium.

Table 7. Description of Data and Method Choices as Moderator Variables

Variable	Description	Mean	Std. Dev.
Journal Quality*			
Top journal	= 1 if a study is published in a journal with a Scimago Journal Ranking above 1.00, 0 otherwise	0.22	0.42
Geographical Region*			
North America**	= 1 if a sample refers to firms from North America, 0 otherwise	0.52	0.50
Europe	= 1 if a sample refers to firms from Europe, 0 otherwise	0.25	0.43
East Asia & Pacific	= 1 if a sample refers to firms from East Asia & Pacific, 0 otherwise	0.09	0.28
South Asia	= 1 if a sample refers to firms from South Asia, 0 otherwise	0.03	0.18
Latin America	= 1 if a sample refers to firms from Latin America, 0 otherwise	0.11	0.31
Sample Year*			
After 2001	= 1 if the average year of sample data is after 2001, 0 otherwise	0.51	0.50
Measurement of Hedging*			
FX hedgers**	= 1 if the estimate refers to FX hedgers only or hedgers of mixed exposures including FX, 0 otherwise	0.42	0.49
IR hedgers	= 1 if the estimate refers to IR hedgers only, 0 otherwise	0.10	0.30
CP hedgers	= 1 if the estimate refers to CP hedgers only, 0 otherwise	0.29	0.45
Hedging dummy variable	= 1 if a dummy variable is used as a hedging measure, 0 if a continuous variable is used	0.68	0.47
Derivatives users	= 1 if hedgers are defined as derivatives users, 0 if the hedging definition includes other financial and operational hedging strategies	0.94	0.23
Control for ex-ante exposure	= 1 if the estimate refers to a sample of firms with an ex-ante risk exposure, 0 otherwise	0.44	0.50
Focus on specific instruments	= 1 if the estimate refers to a specific group of derivative instruments (e.g., options or futures), 0 otherwise	0.10	0.31
Measurement of Firm Value*			
Market-to-book ratio	= 1 if Tobin's Q is measured by the market-to-book ratio, 0 if Tobin's Q is measured by an advanced approach	0.66	0.47
Estimation Characteristics*			
Control for firm fixed effects	= 1 if the estimation controls for firm fixed effects, 0 otherwise	0.21	0.41
Control for endogeneity	= 1 if the estimation controls for reverse causality between the firm value measure and the hedging measure, 0 otherwise	0.20	0.40
Control for sample selection	= 1 if the estimation controls for sample selection bias, 0 otherwise	0.03	0.17
Robust errors	= 1 if heteroscedasticity-robust and/or cluster-robust standard errors are reported, 0 otherwise	0.45	0.50
Interaction term	= 1 if the hedging variable enters the primary regression model in interaction with other variables, 0 otherwise	0.13	0.33

(Continued on next page)

*Control Variables******			
Control for other risk exposures	= 1 if the primary regression includes two or more estimates for different risk exposures, 0 otherwise	0.30	0.46
Control for operational hedging	= 1 if the primary regression includes a control variable for operational hedging, 0 otherwise	0.52	0.50
Control for managerial ownership	= 1 if the primary regression includes a control variable for managerial ownership, 0 otherwise	0.17	0.37
Control for liquidity	= 1 if the primary regression includes a control variable for liquidity, 0 otherwise	0.34	0.47
Control for leverage	= 1 if the primary regression includes a control variable for the corporate debt ratio, 0 otherwise	0.91	0.29
Control for dividend policy	= 1 if the primary regression includes a control variable for dividend policy, 0 otherwise	0.74	0.44

Notes: This table presents the definition and summary statistics of the variables measuring methodological and data-related heterogeneity. All variables are manually collected from the primary studies estimating the firm value effects of corporate hedging. * These variables are denoted $\boldsymbol{M'}$ in equation Eq. (15). ** Marks the omitted base category in the meta-regression analysis. *** These variables are denoted $\boldsymbol{M}$ in equation Eq. (15).

Journal Quality. To consider quality differences not captured by the subsequent method and data characteristics, I define a dummy variable (*Top journal*) that indicates whether hedging premiums are calculated from the results of a study published in one of the most influential finance journals. A journal is assigned to this group if its Scimago Journal Ranking (SJR) exceeds the score of 1.00.[33] I expect lower hedging premiums in higher ranked journals. This assumption is based on the hypothesis of a higher quality regarding the data and methods of studies published in those journals. High quality in top journals is assured by a rigorous review process with multiple referees and revision rounds. Moreover, studies published in leading journals often examine larger samples, which implies that standard errors are usually lower and thus even small hedging premiums might produce statistical significance. Hence, there might be no need to find large premiums to offset large standard errors, which could reduce the risk of publication selection.

Geographical Region. The countries examined in the primary studies are clustered into five geographical areas using regional dummy variables. The regional dummy variables indicate whether a specific observation belongs to a certain region or not. The baseline category is North America, which is the most examined region, accounting for 52% of the estimates in the sample. Following discussions about country-level differences in corporate hedging behavior (Allayannis et al., 2012; Bartram et al., 2009), I predict better access to hedging instruments in more developed world regions, especially North America and Europe. These regions have more

[33] The journals in this category are *The Journal of Finance*, *Journal of Corporate Finance*, *Review of Financial Studies*, *Review of Finance*, *Journal of International Economics*, and *Energy Economics*.

mature derivatives markets providing better access to hedging instruments, which facilitates hedging at the shareholder level and could therefore lead to lower hedging premiums.

Sample Year. Following previous meta-analyses (among others, Hanousek and Kočenda, 2011: 316), I capture structural changes in the hedging and firm value nexus over time. A dummy variable (*After 2001)* is coded that is equal to one for studies with an average sample year after 2001. This breakpoint represents the mid of the average sample years across all studies in the sample.[34] Another motivation for this breakpoint is the sharp increase in the derivatives market turnover starting after 2001 (Section 3.2.3, Figure 16). This inflection point in derivatives turnover could reflect a higher demand for corporate hedging instruments, which might be rewarded by a higher hedging premium.

Measurement of Hedging. I collect information on whether an estimate of the hedging premium refers to FX, IR, or CP hedging by coding a dummy variable for each risk category. Some primary studies analyze mixed exposures, which always include FX hedgers. Thus, I use both the FX hedgers and the mixed exposure group as a baseline category (*FX hedgers*). In line with previous literature (Allayannis et al., 2012; Carter et al., 2006; Jin and Jorion, 2006), I predict systematic differences in the hedging premiums for the three exposure types. FX risk often arises from complex foreign activities and currency streams across various countries and world regions. Accordingly, FX risk may be difficult for external shareholders to monitor and hedge. Consequently, companies might hedge these risks more effectively because of superior information about the company's FX exposure. In contrast, hedging IR risk (*IR hedgers*) and hedging CP risk (*CP hedgers*) could be associated with lower information asymmetry for shareholders. Moreover, previous literature argues, with regard to CP risk, that investors might prefer to leave them unhedged, as they may want to actively invest in the risk exposure of, for example, gold mines, oil and gas companies, or airlines (Jin and Jorion, 2006: 895).

I also consider whether an observed estimate of the hedging premium comes from a model where the independent variable is a hedging dummy variable (*Hedging dummy variable*) or a continuous hedging measure. While continuous measures convey more detailed information than a simple dummy variable, the disclosed hedging volume often depends on accounting rules, which are different across countries and regions. The continuous hedging measures used in the 71 primary studies are notional values of derivatives obtained from annual reports (48%),

[34] I prefer a breakpoint dummy over the average sample year, as the average sample year is highly correlated with other variables exhibiting time trends, especially the macro variables to be explained in the next section.

actual hedge ratios[35] obtained from internal company information (42%), fair values of derivatives obtained from annual reports (1%), and other continuous measures like the number of different contracts used for hedging (9%). In summary, it is not clear whether to expect a larger or smaller hedging premium for dummy hedging measures as compared to continuous hedging measures.

Risk reduction can be achieved by different financial and operational strategies (among others, Allayannis et al., 2001; Hoberg and Moon, 2017; Treanor et al., 2013). Equating hedgers with derivatives users bears the risk of incorrectly defining firms without derivative holdings as non-hedging firms, although they could just employ alternative strategies for hedging. To control for the impact of the hedging definition on the hedging premium, a dummy moderator is introduced (*Derivatives users*) that is equal to one if a study defines hedgers as derivatives users, and zero if studies also consider alternative strategies beyond derivatives use in their hedging definition. Following previous literature on the complementary relation between operational and financial hedging (among others, Allayannis et al., 2001; Kim et al., 2006), I predict that studies accounting for alternative hedging strategies report larger hedging premiums.

If a firm does not hedge, this might either be driven by its explicit decision against hedging or by the absence of a risk exposure. To avoid biased inferences, several authors propose the exclusion of firms without ex-ante exposure from the primary study's data sample (among others, Allayannis and Weston, 2001: 392; Magee, 2013: 64). I define a dummy moderator variable (*Control for ex-ante exposure*) for studies focusing on firms or industries with an ex-ante risk exposure and expect those studies to find larger hedging premiums due to the more precise distinction of the treatment group (hedgers) and the control group (non-hedgers). Another dummy variable (*Focus on specific instruments*) denotes whether a study reports estimates for the firm value implications of single derivative instruments, such as options, futures, or swaps, as opposed to multiple instruments being used for hedging. As companies usually employ different instruments to hedge their risk exposures (Bodnar et al., 1998: 87-91), the analysis of a single derivative instrument might not cover the full hedging strategy. Hence, I expect lower hedging premiums for estimates focusing on single instruments.

Measurement of Firm Value. The primary studies in the meta-sample quantify the firm value by a measure of Tobin's Q. However, the calculation of Tobin's Q requires information about the market value of the long-term debt and the replacement costs of fixed assets, which

[35] This is typically measured by the actual quantity that is hedged (e.g., the volume of oil production hedged) divided by the actual risk exposure (e.g., the total oil production).

is usually not easy to obtain. Therefore, most studies apply approximations of Tobin's Q through the market-to-book ratio. I control for differences in the definition of the firm value measure by a dummy variable (*Market-to-book ratio*) denoting whether firm value is measured by the market-to-book ratio or alternatively by more advanced measures of Tobin's Q as proposed by Chung and Pruitt (1994), Perfect and Wiles (1994), or Lewellen and Badrinath (1997). The advanced measure is the base category. Following Allayannis and Weston (2001: 266), who find small differences in the hedging premium using alternative measures of firm value, I have no clear ex-ante expectation for the impact of the firm value definition on the size of the reported premium.

Estimation Characteristics. Unobserved heterogeneity across groups of firms might introduce biases in the primary regression estimates. If these unobserved group factors are correlated with the variables of interest, omitted variable bias might infect the estimated parameters. In regressions with a firm value proxy as dependent variable, firm fixed effects have been shown to be an important factor (Gormley and Matsa, 2014: 640-642). To control for this, I define a dummy variable (*Control for firm fixed effects*) that is equal to one for hedging premiums obtained from primary regression models estimated with firm fixed effects and zero otherwise. In line with recent econometric literature suggesting that, in the presence of unobserved group heterogeneity, the fixed effects estimator is consistent for models with Tobin's Q as dependent variable (Gormley and Matsa, 2014), I expect studies including firm fixed effects in their models to report lower hedging premiums.

Another major threat for the validity of empirical studies arises from endogenous relations among the regression variables and the error term. A common source of endogeneity is reverse causality, which arises when firms with higher values tend to hedge rather than hedging causing higher firm values.[36] Especially more recent studies address the endogeneity problem by using IV approaches. A dummy variable (*Control for endogeneity*) signals the cases where researchers explicitly address the endogeneity problem in their models by IV approaches or other related methods.[37] I hypothesize that estimates obtained from models that do not control for endogeneity report biased inferences of the hedging premium.

Empirical research on corporate hedging has shown that hedgers exhibit different firm characteristics than non-hedgers (among many others, Géczy et al., 1997; Mian, 1996; Nance et al., 1993). Hence, firms do not randomize their hedging activities, which might generate a

[36] As OLS estimation of Eq. (11) relies on the exogeneity assumption of the regressors, hedging premiums without accounting for endogeneity might be biased.

[37] 69% of the models with endogeneity correction use IV approaches via two-stage/three-stage least squares regression, 23% use a generalized method of moment (GMM) estimation, the remaining 9% use other methods.

selection bias in the estimation of the value effects of hedging. To control for this bias, several authors apply a two-step Heckman regression (Heckman, 1979). A dummy moderator variable (*Control for sample selection*) is coded to be one if an estimate is observed from a model that applies such a two-stage procedure. I expect lower premiums from models accounting for selection bias, as these models reduce the threat of an overestimation of the hedging premium.

As a common remedy to avoid biased standard errors due to non-independent observations or non-constant variances of the error term, authors apply robust estimations of the standard errors (Petersen, 2008). I consider this error correction through a dummy variable (*Robust errors*) and expect lower hedging premiums for estimates collected from regression models accounting for robust errors in their estimation. Moreover, some estimates of the hedging premium are obtained from models with interaction terms between the independent hedging measure and other corporate variables (e.g., the size of the risk exposure). I capture these estimates with interaction terms through a corresponding dummy variable (*Interaction term*).

Control Variables. While examining the impact of a certain risk exposure (e.g., FX risk), some studies control for the impact of hedging other potentially correlated risk exposures (e.g., IR risk). To consider the control for other risk exposures, a binary variable (*Control for other risk exposures*) is defined that tracks whether a primary regression model includes more than one variable for different hedging exposures in the same regression model. I expect the premium of a specific exposure to be lower if the model considers controls for other exposures in the same primary regression.

The primary studies in the sample include a wide set of controls to filter out the influence of other variables that might impact firm value. Common controls are measures for firm size, liquidity, leverage, dividend policy, operational hedging,[38] and managerial ownership. As the firm size is included in over 97% of the models, I do not explicitly control for it. For the other variables, corresponding dummy variables are coded (*Control for operational hedging*, *Control for managerial ownership*, *Control for liquidity*, *Control for leverage*, *Control for dividend policy*). If estimates are collected from regressions including the respective control variables, the corresponding dummy is equal to one, and zero otherwise. Following the reasoning outlined before, I predict a positive sign for operational hedging, i.e., studies controlling for operational hedging estimate larger premiums. The expected sign of the other variables depends on the correlation between the control variable and the corporate hedging variable as well as the firm value variable. A recent meta-analysis of 175 studies on the determinants of corporate hedging

[38] Following Allayannis et al. (2001: 392), studies including a measure for geographical diversification are regarded as controlling for operational hedging.

by Geyer-Klingeberg et al. (2019) shows that the previous literature finds the following average impact of the controls on the 'hedging variable': liquidity (negative), leverage (positive), dividend policy (positive), and managerial ownership (mixed). Another meta-analysis by Hang et al. (2021b) reports the following average impact of the controls on 'firm value': liquidity (positive), leverage (negative) and dividend policy (negative). For managerial ownership they do not report results. Taking the previous evidence together, I predict larger hedging premiums for studies controlling for liquidity and smaller premiums for studies controlling for leverage, dividend policy, and managerial ownership.

3.6.3. Country-Level Differences

The previous literature points out that country-level factors, such as measures for external governance, are important determinants for the value implications of hedging (Allayannis et al., 2012). The meta-analysis in this chapter contributes to the literature by analyzing which country-level moderators promote or inhibit the impact of hedging on firm value. With meta-analysis, I can take advantage of the full set of countries examined across all studies. Based on the country and sample period being reported, I construct average values of several country-level variables that are assumed to condition the degree of market friction and thus, based on the positive theory of corporate hedging, also have an impact on the value generation through hedging.[39] If not stated otherwise, the country-level data is obtained from World Bank Open Data (World Bank, 2017a, b, c, d). The country-specific moderators are reported in Table 8.

Financial and Economic Development. Fundamental hedging theory assumes hedging to be costless (Smith and Stulz, 1985: 392-393). In practice, however, hedging entails transaction costs. Following Bartram et al. (2009: 190), I use the natural logarithm of the average daily derivatives trading volume in FX and IR markets (scaled by the country's GDP) to measure the size and liquidity of markets for financial hedging instruments (*Derivatives market volume*).[40] Since the development of derivatives markets and other financial markets is highly correlated, I also consider the logarithm of the country's average stock trading volume scaled by the GDP as a further proxy for the maturity of financial markets (*Stock trading volume*).

Moreover, I follow Bartram et al. (2009: 191) and include the trade magnitude as another measure of economic development (*Trade magnitude*). The trade magnitude is defined as the sum of imports and exports scaled by GDP. Finally, I define a dummy variable classifying

[39] If the sample period of the data in the primary studies does not exactly correspond to the data available from the external sources, I follow Kysucky and Norden (2016: 93) and use the closest available country-year observation.
[40] This data is obtained from the Bank of International Settlements (BIS). As the information is only available on a triennial basis (starting in 1995), I estimate the missing annual values by linear interpolation.

whether the countries examined in the primary studies are members of the Organisation for Economic Co-operation and Development (OECD) or not (*OECD member*). I predict that the hedging premium should be higher in countries with less developed markets, i.e., the non-OECD members, where transaction costs might be higher and thus access to derivative instruments for hedging, especially for outside investors, might be constrained.

Table 8. Description of Country-Level Moderators

Variable	Description	Mean	Std. Dev.
Financial and Economic Development			
Derivatives market volume	Logarithm of the country's average daily derivatives trading volume in FX and IR markets scaled by GDP (BIS, 2016b, c)	0.13	0.21
Stock trading volume	Logarithm of the country's average stock trading volume scaled by GDP (World Bank, 2017c)	4.46	0.77
Trade magnitude	Logarithm of the country's average sum of exports and imports scaled by GDP (World Bank, 2017b)	3.49	0.48
OECD member	= 1 if a country is a member of the OECD, 0 otherwise	0.82	0.39
Legality and Governance			
Rule-of-law	The country's average rule-of-law index (World Bank, 2017d)	1.18	0.75
Shareholder rights	The country's average aggregate index of shareholder rights protection (World Bank, 2017a)	5.46	1.99
Creditor rights	The country's average aggregate index of creditor rights protection (World Bank, 2017a)	8.20	3.43
Ownership concentration	The country's average measure of ownership concentration (Dahlquist et al., 2003)	0.24	0.25
Financial Distress and Taxes			
Time to resolve insolvency	Logarithm of the country's average time (in years) between filing for insolvency in court until the resolution of distressed assets (World Bank, 2017a)	0.88	0.34
Financial risk*	Logarithm of the country's average International Country Risk index for financial risk (PRS Group, 2015)	3.60	0.09
Composite risk*	Logarithm of the country's average International Country Risk composite risk index (PRS Group, 2015)	4.32	0.07
Tax rate	Logarithm of the country's average corporate tax rate (World Bank, 2017a)	3.80	0.24

Notes: This table presents the definition and summary statistics of the variables measuring country-level heterogeneity. Values of the country-level variables are assigned according to the country and observation period (sample year) examined in the primary studies' samples. * Higher values indicate a less risky environment.

Legality and Governance. In line with Allayannis et al. (2012: 68), I analyze different variables capturing a country's governance mechanisms and legal environment. I predict that in countries with higher values for the rule-of-law index (*Rule-of-law*), hedging premiums should be lower because companies encounter lower transaction costs to enter complex financial contracts like derivatives (Bartram et al., 2009: 191). The rule-of-law index measures the effectiveness of the legal system in a country (Kaufmann et al., 2010: 4). As another measure of agency costs, I examine the indices of shareholder rights and creditor rights (*Shareholder* *rights, Creditor rights*). This is based on the hypothesis that in countries with significant rights for shareholders and creditors, agency costs for reducing information asymmetries should be lower and thus according to shareholder value maximization theory, value creation through hedging might be lower as well. Finally, I analyze the ownership concentration measure (*Ownership concentration*) by Dahlquist et al. (2003: 88-89) and hypothesize that tighter concentration of market capitalization of closely held shares suggests lower shareholder diversification. Less diversification at the shareholder level can lead to higher value creation through hedging, as less diversified shareholders have more incentives to hedge at the firm level (Smith and Stulz, 1985: 403).

Financial Distress and Taxes. To quantify differences in financial distress costs, I use the logarithm of the number of years between fillings for insolvency in court until resolution of the distressed assets (*Time to resolve insolvency*). As a long period of resolution implies higher costs of financial distress, I predict a positive relation with the hedging premium because hedging can reduce the effect of financial distress costs (Froot et al., 1993: 1632). Analogously to Bartram et al. (2009: 190-191), the heterogeneity of a country's financial and overall risk factors is used to test financial distress theory. Both the financial and the composite[41] risk indices (*Financial risk*, *Composite risk*) as laid out by the International Country Risk Guide (PRS Group, 2015) are included as moderator variables. I predict larger hedging premiums in countries with higher risk.[42] Finally, the country's tax rate (*Tax rate*) serves as a proxy for the tax smoothing effect of hedging. The hypothesis is that countries with higher tax rates should have larger hedging premiums because the tax reduction effect of hedging is likely to increase with the tax rate. It should be noted, however, that this may be an inaccurate measure because the tax impact of hedging depends on the convexity of the tax schedule (Smith and Stulz, 1985: 393) and not on the actual tax rate.

3.7. Results

The results section is divided into six parts. Section 3.7.1 presents summary statistics of the primary studies and the hedging premiums calculated from the results reported in these studies. Section 3.7.2 is dedicated to the analysis of publication selection bias. Section 3.7.3 presents the results of heterogeneity analysis, especially the impact of method and data choices on the hedging premiums. Section 3.7.4 shows the findings for the country-level factors as drivers of

[41] This is the sum of economic, financial, and political risk factors.
[42] These indices are inverse measures of country risk, i.e., higher scores imply lower risk.

the hedging premium. Section 3.7.5 reports the mean hedging premium when assuming a best practice study design. Finally, Section 3.7.6 shows the results of additional analyses.

3.7.1. Descriptive Meta-Statistics

General Study Characteristics. Table 9 shows the sample composition and summary statistics for key study characteristics of the final collection of 1,016 estimates of the hedging premiums across 71 primary studies.

Table 9. Summary Statistics of Primary Studies

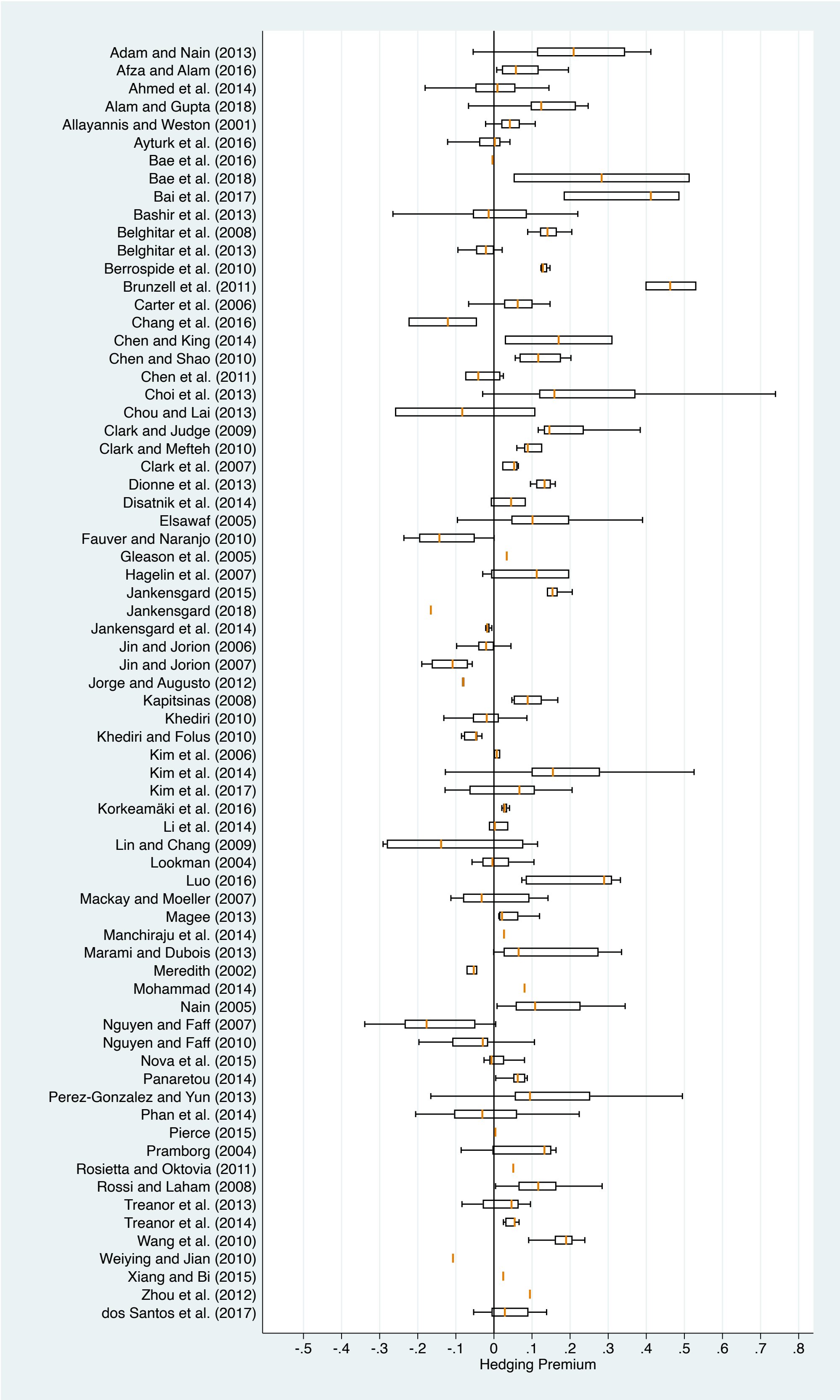
Panel A. Sample Composition					
Publication Status per Study		**Publication Year per Study**		**Region per Estimate**	
Published journal article	41	2001-2005	8	North America	530
Conference paper	14	2006-2010	22	Europe & Central Asia	253
Working paper	13	2011-2015	29	Latin America	110
Thesis	2	2016-2018	12	East Asia & Pacific	89
Book chapter	1			South Asia	34
Total	71	Total	71	Total	1,016

Panel B. Sample Characteristics					
	Mean	**Median**	**Min**	**Max**	**Std. Dev.**
Average sample year	2003	2004	1993	2012	4.83
No. of sample years	6	6	1	20	4
No. of sample firms	309	176	13	2,612	435
No. of firm-year observations	1,213	445	56	11,085	2,185
No. of citations	43	5	0	1,027	139
Journal impact factor	2.58	0.72	0.22	14.55	4.45

Notes: This table shows key characteristics of the selected primary studies. The publication status and the publication year refer to the sample of 71 primary studies. The distribution across geographical regions is calculated for the total sample of 1,016 hedging premiums that are computed from the estimates reported in the primary studies. Each study can provide one or multiple hedging premiums. All summary statistics in Panel B refer to the sample of 71 primary studies. The sample years are the time span between the start year and the end year of the data sample examined in a primary study. The firm count is the total number of unique firms included. The firm-year observations represent the total number of observations used for the estimation of the hedging effect on firm value. The number of citations is collected from Google Scholar, as of March 2018. The journal impact factor, which is measured by the SJR, refers to published journal articles only.

Most papers are published in referred academic journals (41 studies). The sample also includes grey literature such as conference papers, working papers, PhD theses, and book chapters (30 studies). The first study in the sample was published in 2001, while the most recent study was published in 2018. The publication year of the majority of papers is between 2010 and 2015 (29 studies). Consequently, the literature stream focusing on the interactions between hedging and firm value is relatively new as compared to many other traditional research fields in corporate finance, such as corporate debt policy, where the first studies were published in the 1970s (among others, Taub, 1975).

Figure 20. Boxplots of Hedging Premiums per Primary Study

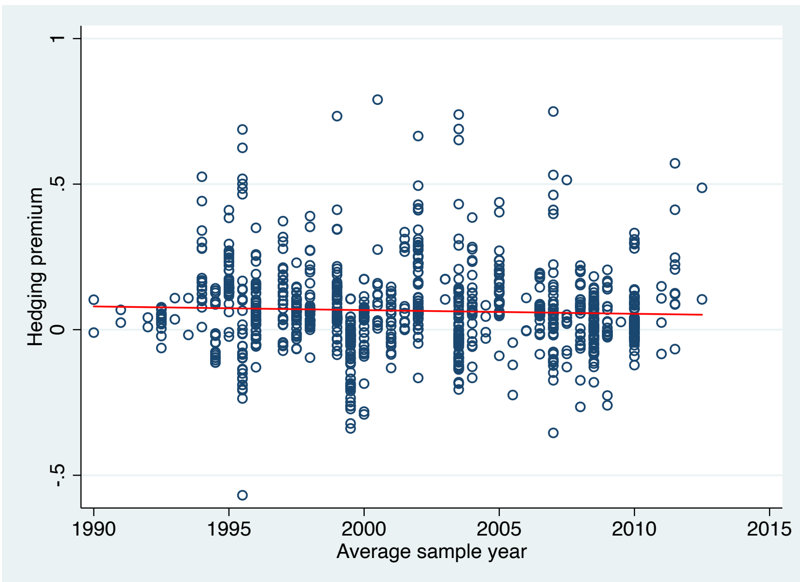


Notes: The figure shows a boxplot of the hedging premiums per primary study. Outliers are excluded from the figure but are included in all statistical tests. Results vary within studies depending on, among other factors, the number of reported robustness checks and the subsamples analyzed in the study.

Table 9 also shows that the studies in the sample analyze firm-level data across all geographical regions, providing data to explore the impact of geographical differences. However, most hedging premiums refer to companies in North America (530 estimates), followed by Europe (253 estimates), and Latin America (110 estimates). The earliest data sample in the studies begins in 1993, while the most recent sample ends in 2012. On average, the primary studies analyze 6 years of firm-level data with 309 sample firms and 1,213 firm-year observations. The most cited article, that of Allayannis and Weston (2001), counts over 1,000 citations on Google Scholar, while the average study has 43 citations.

Distribution of Hedging Premiums. Figure 20 shows the boxplots of the hedging premiums per study. Following Tukey (1977), the median premium per study is highlighted in each box. The edges of the boxes show the interquartile range between the 25% and 75% percentiles. Whiskers cover the 25% (75%) percentile +/− 1.5 times the interquartile range, if such estimates exist. The boxplots show substantial heterogeneity among the hedging premiums, both within and between studies. Between the studies, the premiums cover a range spanning from less than –30% (Nguyen and Faff, 2007) to over +70% (Choi et al., 2013). The premiums within the studies also cover a wide range. Some studies even report both large negative and large positive premiums for different subsamples and robustness analyses. Much of this substantial heterogeneity within and between studies may stem from differences in data, methods, and country characteristics described in the Sections 3.6.2 and 3.6.3.

Figure 21. Hedging Premiums over Time



Notes: The figure shows the hedging premiums over time. The horizontal axis represents the average sample year to which the respective hedging premium refers. Many studies report multiple findings for varying sample years. The red line represents the time trend.

Figure 21 illustrates the development of the hedging premiums over time. The graph depicts the premiums for each study on the vertical axis and the average year of the respective study sample on the horizontal axis. The red-colored trend line shows a downward trend that is, however, only significant at the 10% level ($\hat{\beta} = -0.0008$, $t = -1.73$). Accordingly, the hedging premium decreases by –0.08% per average sample year. Furthermore, from the decreasing dispersion of premiums it becomes apparent that the firm value effects of hedging seem to converge slightly over time, but without reaching a clear consensus in more recent studies.

Mean Hedging Premium. Table 10 reports the meta-analyzed means of the hedging premiums. Panel A presents simple average effects, while Panel B shows weighted means using different meta-methods and weighting schemes as described in Section 2.2.1.

Table 10. Means of the Hedging Premium across the Literature

	Mean	95% CIs		Weights	Clustered SEs	No. of studies	No. of obs.
Panel A. Unweighted Means							
Arithmetic mean	0.064	0.056	0.072	Unweighted	No	71	1,016
Arithmetic mean	0.064	0.048	0.079	Unweighted	Yes	71	1,016
Panel B. Weighted Means							
FEM	0.016	0.014	0.017	$1/SE(HP)^2$	No	71	1,016
REM	0.037	0.032	0.041	$1/(SE(HP)^2 + \hat{\tau}^2)$	No	71	1,016
WLS	0.016	0.007	0.025	$1/SE(HP)^2$	Yes	71	1,016
WLS	0.057	0.035	0.079	$1/m$	Yes	71	1,016
WLS	0.017	0.012	0.022	$1/(SE(HP)^2 \times m)$	Yes	71	1,016

Notes: The unweighted average effect is calculated as the arithmetic mean across all estimates of the hedging premiums. FEM uses the inverse of the effect size estimates' variances as weights, while the REM also considers the between-study variation in the weighting (Borenstein et al., 2009: 69-74). The WLS approach follows Stanley and Doucouliagos (2015). Confidence intervals (CIs) with robust standard errors (SEs) are clustered at the study level and country level. m = The number of estimates for the hedging premium reported per primary study.

According to the results reported in Panel A, the average hedging premium across all 1,016 estimates for the hedging premium is 0.063, which is statistically significant at the 5% level using standard CIs, but also for CIs with standard errors clustered at the individual study level and country level. Clustering of the standard errors is applied to account for the effects of non-independent premiums calculated from the estimates reported in the same study, or when different studies use overlapping data sets from firms in the same country. Accordingly, when I consider the full set of primary studies and take the average of the hedging premiums calculated from their reported results, I find that hedging firms exhibit, on average, an excess in firm value of 6.3% as compared to non-hedging firms.

Panel B reports the mean effects when the hedging premiums calculated from the primary studies are weighted by either the inverse variance of the hedging premiums using the FEM or WLS approach, the inverse of the sampling variance plus the between-study variance using the REM approach, the inverse of the number of estimates for the hedging premium collected per study using WLS, or the interaction between the inverse variance and the inverse number of estimates per study using WLS. FEM and REM are calculated as described in Eqs. (1), (2), and (3) in Section 2.2.1. As the unrestricted WLS approach has been shown to be superior to the conventional FEM and REM if publication selection bias and heterogeneity are present in the sample (Stanley and Doucouliagos, 2015), I prefer the WLS means over the FEM and REM means.

Across all models and weighting schemes, I find that the mean hedging premium is statistically significant at the 5% level, but the size of the mean effects varies between 1.6% for the FEM and WLS approaches using the inverse variance as weights and 5.7% when assigning equal weights to all estimates reported in the same study.

In summary, the analysis of the mean effects shows that when pooling all previous results from the literature, there seems to be a tendency towards an overall positive hedging premium,. Accordingly, hedging seems to create value for firms on average. However, the actual size of the premium is quite different depending on the method and weights used to calculate the average impact. Nevertheless, these findings can only be preliminary because the mean effects neither account for the influence of publication seleciton nor consider the impact of hetereogeneity. Therefore, the next sections continue with the analysis of publication selection bias and heterogeneity.

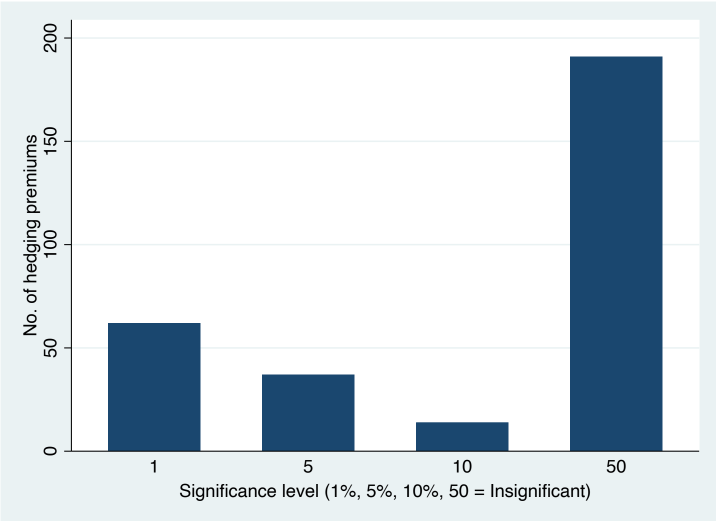
3.7.2. Analysis of Publication Selection Bias

The analysis of publication selection bias begins with graphical overviews, followed by statistical analysis in the second part of the section.

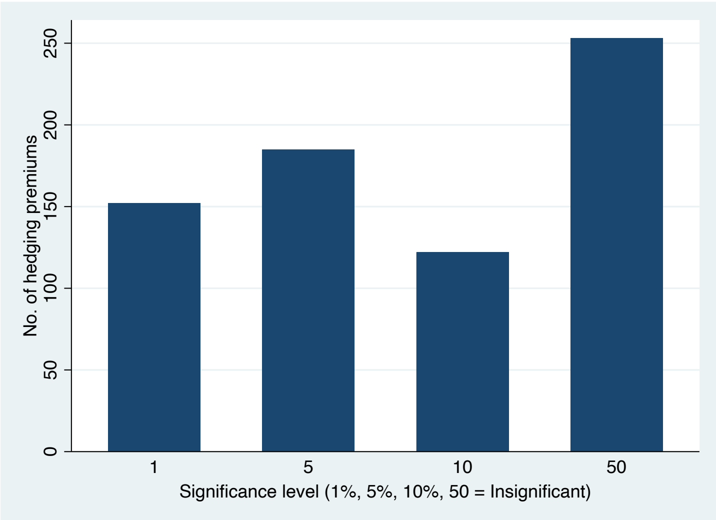
Graphical Analysis. Figure 22 shows the distribution of significance levels of the estimates for the hedging premium, categorized by the 1%, 5%, and 10% thresholds. Hedging premiums that are not statistically significant are clustered in the right bar (50). The two graphs are split by the sign of the premium: negative premiums are displayed in Figure 22(a) and positive premiums in Figure 22(b). Comparing the number of observations per category shows that significantly positive hedging premiums occur more often than significantly negative premiums. If negative premiums are found to be significant, they tend to be highly significant, as the 1% level is the dominant group. In contrast, for the positive premiums, the 5% level is the category into which most of the significant values fall. The skewness in the significance levels gives a first indication of a tendency in the literature towards the selection of significant findings. However, splitting the results into groups according to significance levels assigns the estimates to categorical groups but does not take into account the actual size of the significance level.

Figure 22. Significance Levels of Hedging Premiums

(a) Negative Hedging Premiums



(b) Positive Hedging Premiums



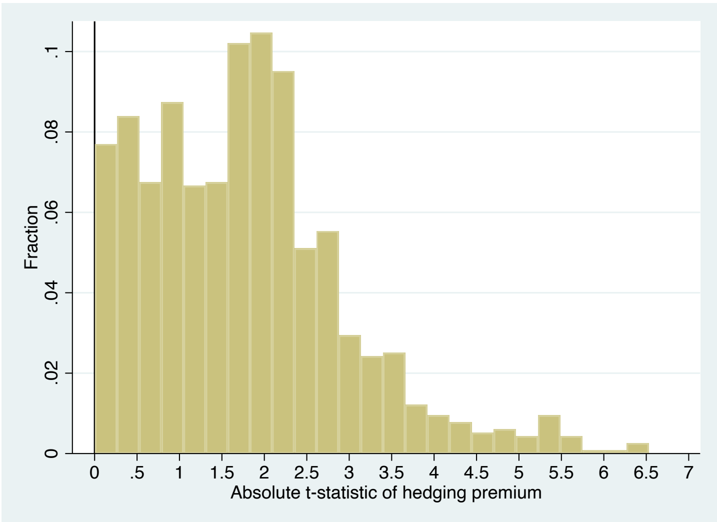
Notes: The figure shows the significance levels of the estimates for the hedging premium clustered by the 1%, 5%, and 10% significance levels and a group of premiums that are not significant (indicated by 50).

As an extension to the previous graphs, Figure 23 depicts the histogram of the absolute t-statistics of the estimated hedging premiums. Without an active selection for statistical significance, I would expect a monotonically decreasing function in Figure 23(a). However, the graph provides evidence that t-statistics at the margin of common critical values are overreported, especially around $t = 2$, which indicates statistical significance at the 5% level.

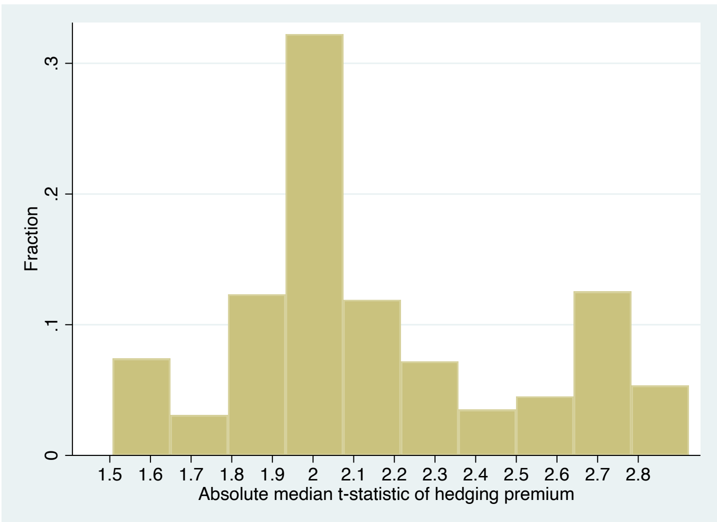
To control for the fact that some studies report many similar t-statistics, a second histogram in Figure 23(b) depicts the median of the t-statistics per study around the critical margin of the 5% significance level. This histogram shows that the fraction of estimates is about three times higher when t-statistics are around $t = 2$. A similar, but less extensive pattern can be found for the critical values at the 1% significance level, which is around $t = 2.7$. The graphical analysis of the t-statistics amplifies the indication that researchers might select estimates of the hedging premium based on their statistical significance. A similar finding was exposed by Harvey et al. (2016: 46) for factor studies on the cross-section of expected returns.

Figure 23. Distribution of t-statistics of Hedging Premiums

(a) All Premiums



(b) Median t-statistics between 1.5 and 2.9

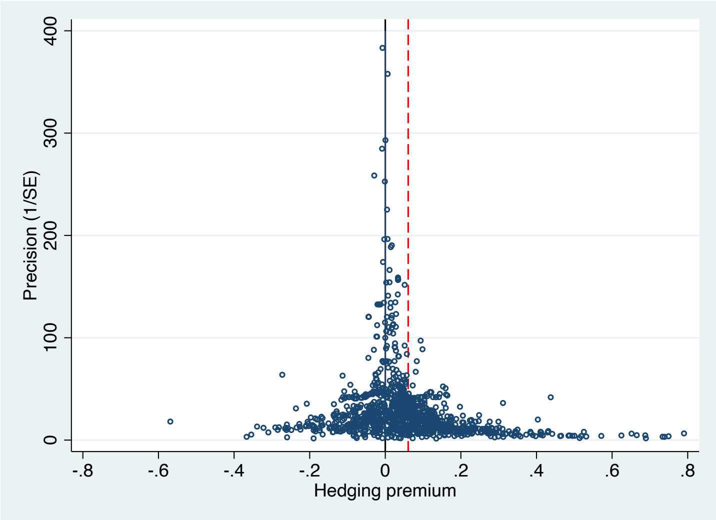


Notes: The figures show (a) the t-statistics of the estimates for the hedging premium, and (b) the median t-statistics per study around $t = 2$. Estimates with large t-statistics > 7 are excluded from the graph, but not from the subsequent statistical analysis.

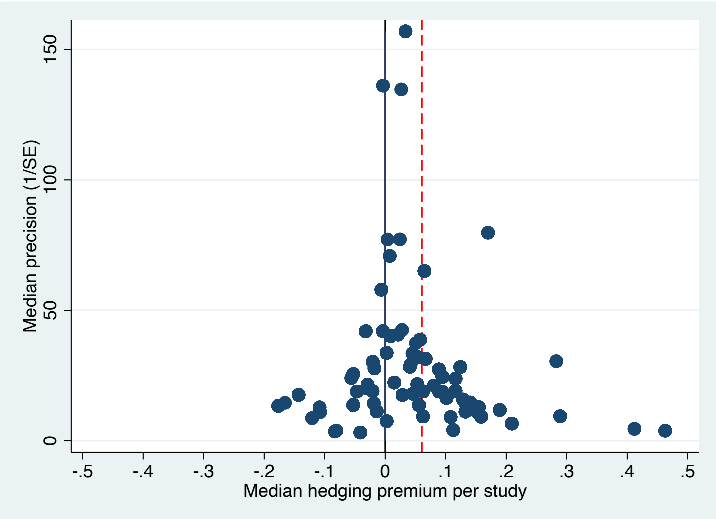
To further explore selective reporting of statistically significant and positive hedging premiums, Figure 24 shows the funnel plots, which is the common graphical tool for the inspection of publication selection (Egger et al., 1997: 632). A funnel plot depicts the magnitude of the estimated hedging premiums on the horizontal axis and the estimates' precision (the inverse standard error) on the vertical axis. Without publication selection, the plot should resemble a symmetric inverted funnel with the most precise estimates concentrated at the top and imprecise estimates at the bottom with a wider dispersion due to larger sampling error. Selective reporting of a particular effect leads to skewness, and the graph becomes asymmetric.

Figure 24. Funnels Plots of Hedging Premiums

(a) All Premiums



(b) Median Premium per Study



Notes: The funnel plot in (a) shows the hedging premiums and their precision, which is the inverse of the estimates' standard errors. In the absence of publication selection, the funnels should be symmetrically distributed around the most precise estimates, which are clustered around the top of the funnel. The dashed lines in red show the sample means. The funnel plot in (b) refers to the median hedging premium and the median precision per study.

Figure 24(a) presents the funnel plot of all estimates for the hedging premium. The most precise estimates, which are shown at the top of the graph, are close to zero. In terms of asymmetry, there is some indication that the right tail of the funnel is heavier. Accordingly, negative hedging premiums appear less frequently than positive hedging premiums. This is an indication for publication selection in favor of positive firm value effects through hedging. In addition, Figure 24(b) plots the median hedging premium and the median precision per study. Again, the most precise estimates are close to zero and a larger proportion of estimates fall into the positive region of the graph.

In summary, the graphical analysis indicates two sources of publication selection: selection for positive premiums and selection for statistically significant premiums. To test whether the graphical implications are also supported by statistical analysis, I employ meta-regression analysis for publication selection bias testing in the next step.

Statistical Analysis. I use the meta-regression model for the analysis of publication selection bias as defined in Eq. (14). This meta-regression can be seen as a test for the asymmetry of the funnel plot resulting from rotating the axes of the plot by 90° and inverting the values on the new horizontal axis. The slope coefficient, $\hat{\beta}_1$, measures publication selection and the constant, $\hat{\beta}_0$, captures the genuine hedging premium beyond publication selection. Simulation studies have shown that $\hat{\beta}_0$ tends to underestimate the true mean effect in cases where a non-zero effect exists (Stanley, 2017: 582). In this case, using a non-linear term by replacing $SE(HP)$ with $SE(HP)^2$ in Eq. (14) reduces the bias of the constant and yields a better estimate of the genuine effect corrected for publication selection (Stanley and Doucouliagos, 2014: 68):

$$HP_{ij} = \lambda_0 + \lambda_1 SE(HP_{ij})^2 + v_{ij}, \text{ with } v_{ij} \sim N(0, SE(HP_{ij})^2), \tag{16}$$

where variable notations are the same as in Eq. (14). The estimate for the intercept, $\hat{\lambda}_o$, is also referred to as precision-effect estimate with standard errors (PEESE) (Stanley and Doucouliagos, 2014: 64).

The results for the analysis of publication selection bias are reported in Table 11. In Models (1), (4), (5), and (6), I apply WLS estimation using the inverse variance of the estimates for the hedging premium as weights to account for heteroscedasticity. Accordingly, these models put more emphasis on the top of the funnel plot (Figure 24a), which are the most precise and thus the most reliable findings in the literature. This is the base model, which is also the established standard in the meta-analysis community in economics (Stanley and Doucouliagos, 2012: 61). Given the ongoing discussion about inverse variance weighting and its caveat that it gives larger weights to studies just because they report more estimates[43] (Zigraiova and Havranek, 2016: 971-974), I use two alternative weighting schemes for robustness analysis. Model (2) uses the inverse number of estimates per study, which assigns the equal weight to 'small' and 'large' studies in terms of the number of reported estimates. Accordingly, the alternative weight considers the top of the funnel and the less reliable and widespread bottom area as equally important. Applying the inverse number of estimates per study as weights in the meta-regression leads also to a weighted regression, but compared to inverse variance weighting, it is not necessarily a WLS estimator. As a compromise, Model (3) employs the interaction between the inverse variance and the inverse number of estimates reported per study as weights, which accounts for differences in the precision within the studies but also assigns equal weight across studies to avoid an undue influence of studies with many estimates.

Table 11. Analysis of Publication Selection Bias

	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) FX Hedgers	(5) IR Hedgers	(6) CP Hedgers
Weights	$\frac{1}{SE(HP)^2}$	$\frac{1}{m}$	$\frac{1}{SE(HP)^2 \times m}$	$\frac{1}{SE(HP)^2}$	$\frac{1}{SE(HP)^2}$	$\frac{1}{SE(HP)^2}$
Panel A. Linear Term						
Constant: $\hat{\beta}_0$	0.001	0.013	0.008*	0.006	0.001	-0.028***
	(0.11)	(1.34)	(1.87)	(0.84)	(0.01)	(-5.79)
SE: $\hat{\beta}_1$	0.931***	0.558***	0.646***	1.421***	0.155	1.053***
	(5.13)	(3.71)	(2.73)	(4.98)	(0.66)	(7.61)
Panel B. Non-Linear Term						
Constant: $\hat{\lambda}_0$	0.014***	0.046***	0.016***	0.025***	0.001	-0.010***
	(3.07)	(4.67)	(6.10)	(3.45)	(0.08)	(-2.87)
SE^2: $\hat{\lambda}_1$	3.879***	0.739***	2.037***	5.483***	3.726**	3.770***
	(4.22)	(2.15)	(3.42)	(2.61)	(2.10)	(8.04)
No. of studies	71	71	71	37	17	24
No. of obs.	1,016	1,016	1,016	424	98	292

Notes: This table reports the results of $HP_{ij} = \beta_0 + \beta_1 SE(HP_{ij}) + e_{ij}$ in Panel A and $HP_{ij} = \lambda_0 + \lambda_1 SE(HP_{ij})^2 + v_{ij}$ in Panel B, where HP_{ij} is the ith estimate of the hedging premium derived from the results reported in the jth study. I use the variance of the estimated hedging premiums as a regressor in Panel B as it has been shown to yield a better correction for publication selection than using a linear term. $\hat{\beta}_1$ and $\hat{\lambda}_1$ measure the presence and magnitude of publication selection. $\hat{\beta}_0$ and $\hat{\lambda}_0$ capture the mean hedging premium beyond publication selection. Models (1), (4), (5), and (6) are estimated by WLS using the inverse of the estimates' squared standard errors as weights. Model (2) uses the inverse number of estimates collected from each study as weights. Model (3) applies the inverse of the estimates' squared standard errors times the inverse of the number of estimates per study as weights. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

[43] On average, a primary study in the sample reports 14 estimates, the minimum is 1, and the maximum is 82.

In addition to the different weighting schemes used in the meta-regression, I report findings for subsamples focusing on the three risk exposure types: FX hedgers (Model 4), IR hedgers (Model 5), and CP hedgers (Model 6). Panel A reports the findings for the linear model (Eq. 14) and Panel B reports the findings for the model with squared standard errors (Eq. 16). In each model, the standard errors of the regression coefficients are clustered at the study level and country level to account for non-independence of the effect size estimates.

In all models, except Model (5), H_0: $\hat{\beta}_1 = 0$ can be rejected at the 1% level. The positive sign of $\hat{\beta}_1$ in these models reveals that the funnel plot is asymmetric and positive premiums are systematically overrepresented in the literature. To interpret the magnitude of the bias, I follow the guidance of Doucouliagos and Stanley (2013: 320-321). They classify the magnitude of publication selection bias as 'little to modest' if the bias coefficient is statistically insignificant or $\left|\hat{\beta}_1\right| < 1$, 'substantial' if the bias coefficient is statistically significant and $1 \leq \left|\hat{\beta}_1\right| \leq 2$, and 'severe' if the bias coefficient is statistically significant and $\left|\hat{\beta}_1\right| > 2$. This classification of publication selection bias refers to the WLS meta-regression with inverse variance weighting shown in the Models (1), (4), (5), and (6). Given these categories, I can confirm a modest to substantial publication selection bias in my sample, as the slope coefficient, which is 0.931 in Model (1), is close to one and statistically significant at the 1% level.

When splitting the sample by the examined risk exposure type, publication selection is substantial for both FX hedging premiums (Model 4) and CP hedging premiums (Model 5). No evidence for publication selection can be found for IR hedgers (Model 6). A reason for the strong bias in the subsample of FX premiums might be that the sample average of reported estimates for FX hedgers, which is 9.7%, is by far the highest as compared to IR and CP hedgers. This might also be driven by the seminal study of Allayannis and Weston (2001: 268), who disclose an average FX premium of 4.87%. Many of the succeeding studies in the field use this large positive premium as the benchmark for the value implications of FX hedging and might actively search for large premiums in their samples to confirm the results of the seminal study of Allayannis and Weston (2001).

As there is weak evidence for some genuine effect beyond publication selection, I replace the standard error with the variance of the hedging premiums (Panel B). The results suggest a positive overall hedging premium across the literature, i.e., taking all 1,016 estimates collected from the 71 primary studies together implies a mean hedging premium of 1.4%. For the alternative weights, the overall premium is even larger, with 4.6% in Model (2) and 1.6% in Model (3). Therefore, assigning larger weights to more precise results leads to a much smaller mean value of the premiums than treating all studies equally in the estimation. When breaking down the mean effects by the type of risk exposure that is being hedged, FX hedgers have a statistically significant firm value mark-up of 2.5% over non-hedgers. For IR hedgers, the effect is 0.01%, which is not statistically significant. For CP hedgers, the mean premium in firm value is -1.0% and statistically significant at the 1% level.

Table 12 reports additional results to identify the mean hedging premium beyond publication selection. Panel A shows the arithmetic mean for the total sample and each exposure type. The distortion due to selective reporting of positive hedging premiums can be illustrated by comparing the corrected mean effects (Panel B) to the uncorrected average premiums (Panel A). Panel B shows additional non-linear and recent techniques to correct for publication selection. The first model is the PEESE estimator, which is already reported in the previous Table 11. The second meta-estimator is the result of the selection model by Andrews and Kasy (2019). Their model assumes that the publication probability is a step function that alters abruptly after reaching prespecified thresholds of the *t*-statistic. I follow Matousek et al. (2022: 332) and use 1.96 and 2.33 as cut-off values of the *t*-statistics in the model. As a third model, I apply the *p*-uniform* method of van Aert and van Assen (2020). This approach follows the idea that *p*-values should have a uniform distribution at the mean underlying effect. The method recalculates the *p*-values for different values of the underlying effect and compares this distribution to the uniform distribution until finding a coefficient that leads to a uniform distribution of the *p*-values. As a further method to calculate mean effects while considering the impact of publication selection, I use the approach of Stanley et al. (2010) and focus only on the 10% of the most precise estimates, i.e., the hedging premiums at the top of the funnel plot. Stanley et al. (2010: 70) suggest that *"discarding 90% of the published findings greatly reduces publication bias and is often more efficient than conventional summary statistics."*

Finally, I apply the Weighted Average of Adequately Powered (WAAP) model by Ioannidis et al. (2017). This approach focuses on effect size estimates that are adequately powered as low power causes high Type II errors. According to Ioannidis et al. (2017: F239), "*a study is adequately powered when there is a high likelihood that it will detect a genuine empirical effect*". The WAAP method is especially useful when there are studies with several high-powered und low-powered estimates, which is not given in the sample data of this chapter. Since the sample of adequately powered estimates is rather small, I follow Bajzik (2020: 18) and reduce the denominator of 2.8 suggested by Ioannidis et al. (2017: F239) to 1.84. This is not the optimal approach as it lowers the boundary for a study to be considered as adequately powered. Therefore, the results of the WAAP should be interpreted with caution.

Table 12. Mean Hedging Premium Beyond Publication Selection

	(1) Full Sample	(2) FX Hedgers	(3) IR Hedgers	(4) CP Hedgers
Panel A. No Correction for Publication Selection				
Arithmetic mean	0.064***	0.097***	0.021	0.033***
	(8.07)	(12.68)	(1.50)	(8.78)
Panel B. Correction for Publication Selection				
PEESE	0.014***	0.025***	0.001	-0.010***
(Stanley and Doucouliagos, 2014)	(3.07)	(3.45)	(0.08)	(-2.87)
Selection model	0.035***	0.055***	-0.005	0.037**
(Andrews and Kasy, 2019)	(3.89)	(4.58)	(-0.33)	(2.18)
p-uniform*	0.042***	0.030***	0.009	0.005
(van Aert and van Assen, 2020)	[<0.001]	[0.004]	[0.522]	[0.635]
Top 10%	0.008	0.020***	0.009	-0.029***
(Stanley et al., 2010)	(1.59)	(3.50)	(1.02)	(-7.01)
WAAP	0.004	0.015***	-	-
(Ioannidis et al., 2017)	(1.54)	(2.61)		
No. of studies	71	37	17	24
No. of obs.	1,016	424	98	292

Notes: Panel A reports the arithmetic mean of the estimates for the hedging premium. Panel B reports five non-linear estimation techniques. PEESE stands for Precision-Effect Estimate with Standard Errors. WAAP stands for Weighted Average of Adequately Powered. I follow Bajzik (2020: 18) and use 1.84 as constant denominator for the WAAP. In the models with no results for the WAAP estimator, there are no adequately powered estimates in the sample. The *t*-statistics of the parameters are reported in parentheses. For PEESE, Top 10%, and WAAP, the reported *t*-statistics are based on robust errors, clustered at the study level and country level. For the *p*-uniform* test, *p*-values are reported in brackets.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

The average hedging premium in the sample is 6.4%. For the subsample of FX hedgers (Panel A), the premium even increases to 9.7%. For IR hedgers and CP hedgers the average premium is much smaller (2.1% and 3.3%) and statistical significance can only be detected for CP hedgers. The distortion due to publication selection bias can be illustrated by comparing the simple average premiums with the methods that correct for selective reporting (Panel B). Across all five models, I find that there is a positive and statistically significant premium for FX hedgers in a range between 1.5% and 5.5%, which is much smaller than the 9.7% of the uncorrected mean. Accordingly, when I do not account for publication selection, the mean effect is exaggerated by a factor of 1.8 to 6.5.

For IR hedgers, I do not find evidence for an overall firm value effect. For CP hedgers and in the full sample, the results are mixed. In the full sample, all models show a positive premium between 0.4% and 4.2%. However, the Top 10% and the WAAP, which restrict the sample to the most powerful estimate, yield rather small mean effects that are not statistically significant at conventional levels. For the CP hedgers, the PEESE and the Top 10% model reveal a negative mean effect between -2.9% and -1.0%, while the selection model shows a positive and significant premium.

In summary, selective reporting of positive and significant hedging premiums distorts the view of the true underlying effect. If the simple average across the literature, which does not correct for publication selection, represents the common view on the impact of hedging on firm value, the meta-analysis uncovers that this view is highly exaggerated.

To further investigate the patterns of publication selection, I interact the standard error of the hedging premiums with dummy variables of four study characteristics: top journal publications with an SJR score above 1.0, FX hedgers, endogeneity control in the primary regression model, and the number of control variables included to capture model misspecification bias. I expect less publication selection in top journals due to the more rigorous review process. In addition, I hypothesize that selective reporting is more pronounced for FX hedgers, as the majority of the literature analyzes this risk exposure (61%) and many previous authors have found large positive premiums for FX hedging (among others, Allayannis and Weston, 2001; Jankensgård, 2015a; Panaretou, 2014). In contrast, some prominent studies find no effects or even negative premiums for IR and CP hedgers (Carter et al., 2006; Jin and Jorion, 2006; Phan et al., 2014). Therefore, I expect that researchers might tend to report positive estimates for FX exposures. I do not have a clear ex-ante expectation as to whether controlling for endogeneity moderates the degree of publication selection. Instrumental variable (IV) estimation, which is the most common method in the sample to control for endogeneity, often yields lower precision but might also find systematically different effects if endogeneity is present. Finally, I hypothesize that publication selection decreases with an increasing number of important control variables because authors have fewer degrees of freedom to modify the model specification.

The findings for the moderators of publication selection are reported in Table 13. The table presents only the results for the baseline model with inverse variance weighting. The results for the alternative weighting schemes are report in Appendix B.2. The outcomes imply that publication selection is significantly lower in top-ranked journals. Similar to the results reported in Table 11 (Models 4, 5, and 6), the studies that focus on the hedging of FX exposures show greater correlation with selective reporting of positive premiums. In contrast, controlling for endogeneity shows no effect, whereas the number of control variables included reduces the magnitude of the publication selection, which confirms the ex-ante hypothesis. In summary, the results of the moderator analysis suggest that publication selection is not equally distributed across all studies and their methodological designs but it is rather driven by specific study characteristics.

Table 13. Moderating Factors of Publication Selection

	(1) Journal Quality	(2) Focus on FX Exposure	(3) Correction for Endogeneity	(4) No. of Control Variables	(5) All Variables
Intercept: $\hat{\beta}_0$	0.001	0.001	0.001	0.001	0.002
	(0.22)	(0.26)	(0.16)	(0.41)	(0.59)
SE: $\hat{\beta}_1$	1.109***	0.190	0.999***	1.659***	1.105***
	(5.96)	(1.42)	(5.04)	(3.73)	(2.44)
SE × Top journal: $\hat{\delta}_1$	-0.860**				-0.594**
	(-4.12)				(-3.14)
SE × FX hedgers: $\hat{\delta}_2$		1.163***			0.928***
		(6.54)			(8.25)
SE × Control for endogeneity: $\hat{\delta}_3$			-0.378		0.097
			(-1.89)		(0.50)
SE × No. of control variables: $\hat{\delta}_4$				-0.074***	-0.067**
				(-2.58)	(-2.33)
No. of studies	71	71	71	71	71
No. of obs.	1,016	1,016	1,016	1,016	1,016

Notes: This table reports the results of Eq. (14) including additional interaction variables: $HP_{ij} = \beta_0 + \beta_1 SE(HP_{ij}) + \delta_k SE(HP_{ij}) M_{ij} + v_{ij}$, where HP_{ij} is the ith estimate of the hedging premium reported in the jth study. M_{ij} is a moderator variable for publication selection. All models are estimated by WLS estimation using the inverse of the estimates' squared standard errors as weights. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level. Top journal = Hedging premium is calculated from the estimates reported in a journal with a SJR score larger than 1.00. FX hedgers = Hedging premium refers to hedgers of FX risk. Control for endogeneity = Primary study explicitly controls for endogeneity. No. of control variables = Total number of variables included as controls in the primary study regression.

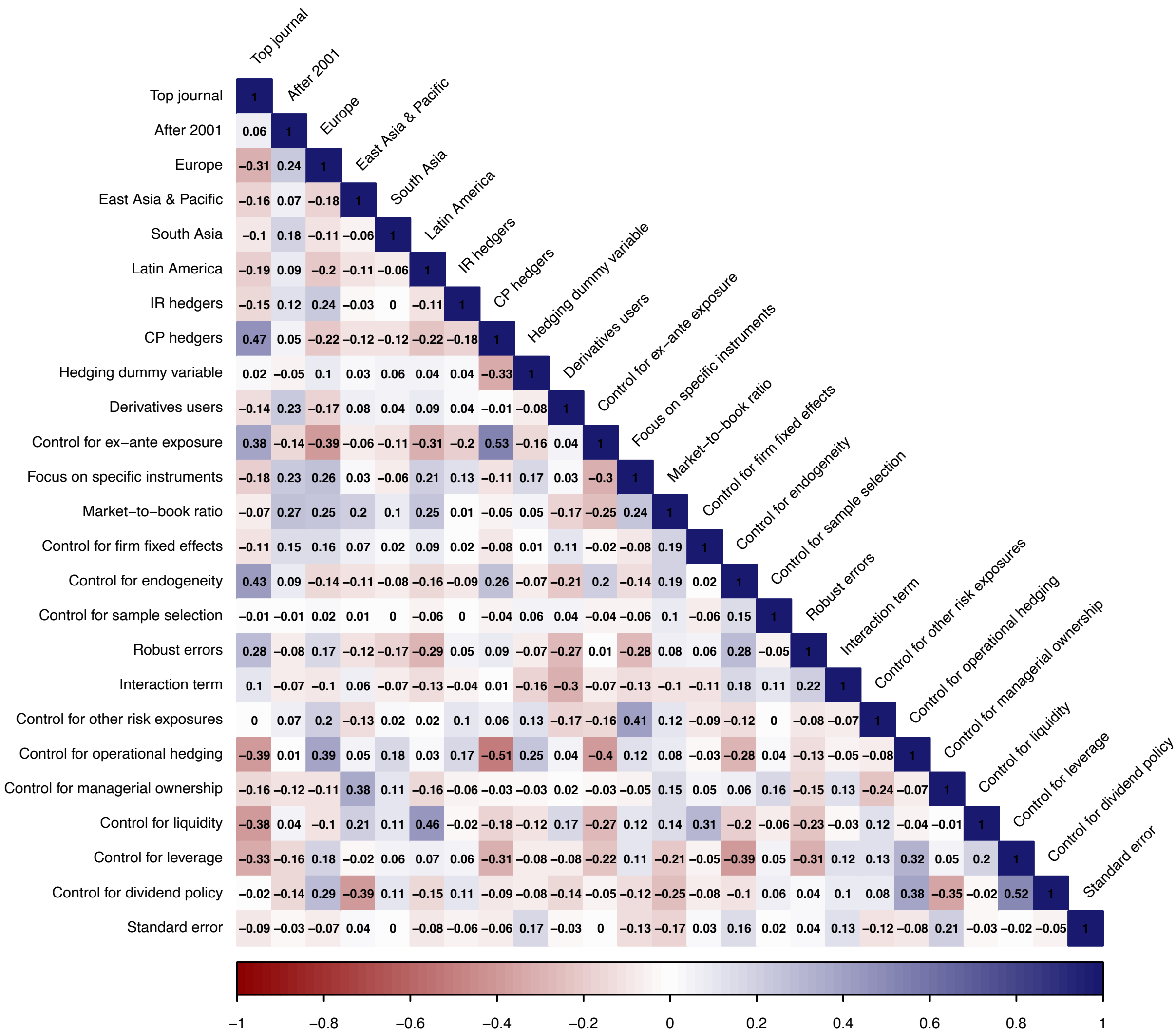
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

3.7.3. Data and Method Choices as Drivers of Heterogeneity

Correlation Analysis. Figure 25 shows the Pearson correlation coefficients for the explanatory variables described in Table 7. The largest correlation among the explanatory variables is 0.53. Moreover, all variance inflation factors are below 10. This indicates that there is no big risk of multicollinearity in the meta-regression model. The correlations also reveal some interesting patterns among the study characteristics. For example, CP hedging as compared to IR and FX hedging tends to be investigated more often in top journals, and articles in top journals tend to control for endogeneity more frequently. In addition, studies on CP hedging often control for firms with an ex-ante risk exposure but tend not to control for operational hedging. Also, studies controlling for dividend policy usually also control for leverage.

Heterogeneity Analysis. Table 14 reports the results for the meta-regression model (Eq. 15), including the variables for data and method choices from Table 7 as moderator variables. Since most explanatory variables are dummy variables, their estimated meta-regression coefficients reflect the average impact on the hedging premiums if the study design deviates from the base group in that specific aspect, while holding all other things equal. Significant variables in the meta-regression indicate that the respective variable indeed affects the hedging premiums.

Figure 25. Correlation Matrix of Explanatory Variables for Data and Method Choices



Notes: The figure shows the Pearson correlation coefficients of the explanatory variables described in Table 7.

Model (1) shows the baseline meta-regression using the inverse of the hedging premiums' variance as weights in the WLS model to put larger emphasis on the more precise and therefore more reliable findings in the literature. Model (2) adds regional cluster as moderator variables. Model (3) uses the number of estimates per study ($1/m$) as alternative weights, and Model (4) weights each premium in the regression by the interaction between the hedging premiums' variance and the number of estimates reported per study ($1/(SE(HP)^2 \times m)$). Model (5) is a reduced model using the general-to-specific approach to derive a parsimonious model. In this model, I sequentially omitted the least significant explanatory variable until all remaining factors are significant at least at the 10% level.[44] This is a common procedure to derive parsimonious models in meta-analysis research (Doucouliagos and Stanley, 2009: 418-420; Valickova et al., 2015: 518). The standard errors in all models are clustered at the study level and country level to control for the non-independence of the estimates.

Table 14. Data and Method Choices as Drivers of Hedging Premiums

	(1) Baseline Model	(2) Regional Differences	(3) Alternative Weights I	(4) Alternative Weights II	(5) Reduced Model
Weights	$\frac{1}{SE(HP)^2}$	$\frac{1}{SE(HP)^2}$	$\frac{1}{m}$	$\frac{1}{SE(HP)^2 \times m}$	$\frac{1}{SE(HP)^2}$
Journal Quality					
Top journal	-0.020*** (-3.61)	-0.025*** (-5.51)	-0.025 (-0.93)	-0.025*** (-3.07)	-0.021*** (-7.63)
Sample Year					
After 2001	0.025** (2.51)	0.031*** (2.86)	0.029 (1.37)	0.040*** (4.85)	0.021** (2.15)
Geographical Region					
Europe vs. North America		0.002 (0.11)	0.067** (2.45)	-0.007 (-0.56)	-0.020 (-1.53)
East Asia & Pacific vs. North America		-0.024* (-1.70)	-0.002 (0.28)	-0.020*** (-2.97)	-0.021*** (-2.86)
South Asia vs. North America		-0.002 (-0.12)	0.089** (2.49)	-0.014 (-0.80)	-0.012 (-0.73)
Latin America vs. North America		0.039** (2.31)	0.114*** (3.59)	0.044*** (3.65)	0.012** (2.21)
Measurement of Hedging					
IR vs. FX hedgers	-0.026*** (-6.62)	-0.026*** (-8.08)	0.012 (0.45)	-0.014*** (-2.58)	-0.024*** (-5.47)
CP vs. FX hedgers	-0.023*** (-3.35)	-0.021*** (-2.60)	-0.062** (-2.45)	-0.034*** (-4.26)	-0.023** (-2.53)
Hedging dummy variable	0.009** (2.53)	0.012** (2.54)	-0.001 (-0.02)	0.022*** (3.43)	0.010* (1.76)
Derivatives users	-0.051*** (-3.78)	-0.043** (-2.12)	0.047 (0.53)	-0.033** (-2.14)	-0.036*** (-3.08)
Control for ex-ante exposure	0.007 (0.81)	0.011 (0.94)	0.078*** (2.67)	0.030*** (2.83)	
Focus on specific instruments	-0.031** (-2.34)	-0.026 (-1.33)	-0.055 (-1.55)	-0.029** (-2.00)	-0.019*** (-2.13)
Measurement of Firm Value					
Market-to-book ratio	-0.006 (-0.38)	-0.017 (-1.52)	-0.014 (-0.94)	-0.030*** (-3.30)	

(Continued on next page)

[44] For the regional cluster variables (North America, Europe, East Asia & Pacific, South Asia, and Latin America), I keep all interrelated moderator variables in the model if one of the regional variables is significant at least at the 10% level.

Estimation Characteristics					
Control for firm fixed effects	-0.023**	-0.027*	-0.066***	-0.012	
	(-2.07)	(-1.85)	(-2.74)	(-1.35)	
Control for endogeneity	-0.013**	-0.011	-0.001	-0.009	-0.013*
	(-2.11)	(-1.54)	(-0.03)	(-1.08)	(-1.82)
Control for sample selection bias	-0.001	0.003	0.140***	0.159**	
	(-0.03)	(0.14)	(7.10)	(2.43)	
Robust errors	0.002	0.016**	0.042**	0.008	
	(0.21)	(2.38)	(2.18)	(0.79)	
Interaction term	-0.021*	-0.008	-0.038	-0.003	-0.020***
	(-1.75)	(-1.36)	(-0.98)	(-0.86)	(-3.14)
Control Variables					
Control for other risk exposures	-0.021***	-0.018***	-0.002	-0.002	-0.017***
	(-5.07)	(-4.75)	(-0.09)	(-0.39)	(-4.33)
Control for operational hedging	0.011***	0.017***	-0.020	0.012*	
	(2.65)	(3.35)	-0.54	(1.83)	
Control for managerial ownership	-0.032***	-0.014**	-0.011	-0.006	-0.021***
	(-2.60)	(-1.98)	(-0.34)	(-0.48)	(-2.82)
Control for liquidity	0.010	0.009	0.006	0.009	
	(1.12)	(0.88)	(0.34)	(0.98)	
Control for leverage	-0.019	-0.015	-0.012	-0.012	
	(-1.28)	(-1.06)	(-0.19)	(-0.94)	
Control for dividend policy	-0.012	-0.020***	-0.025	-0.011*	
	(-1.21)	(-3.03)	(-1.36)	(-1.72)	
Publication Selection					
SE	0.898***	0.809***	0.548***	0.336	0.847***
	(6.00)	(4.78)	(3.40)	(1.59)	(4.45)
Constant	0.086***	0.068*	-0.034	0.060*	0.057***
	(3.48)	(1.93)	(-0.26)	(1.87)	(5.03)
No. of studies	71	71	71	71	71
No. of observations	1,016	1,016	1,016	1,016	1,016

Notes: This table reports the meta-regression results of Eq. (15). Reported coefficients reflect the average impact on the hedging premiums if the study design deviates from the base group of this specific aspect, all other things being equal. Definitions of the explanatory variable can be found in Table 7. Models (1), (2), (5) are estimated by WLS using the inverse of the estimates' squared standard errors as weights. Model (3) uses the number of estimates per study as weights. Model (4) uses the inverse of the estimates' squared standard errors times the inverse number of estimates per study as weights. Model (5) is a reduced model based on a general-to-specific approach. The *t*-statistics in parentheses are based on robust errors, clustered at the study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Several key findings can be derived from Table 14. These results provide new explanations for the sources of heterogeneity among the hedging premiums calculated from the existing research record, but also confirm several aspects of earlier discussions in the literature. For the interpretation of the results, I refer to inverse variance weighting (Models 1, 2, 5) as the baseline approach. The other weighting schemes are robustness tests. Appendix B.3 reports the results of the reduced model (Model 5) using the two alternative weights in the meta-regression.

Journal Quality. Even after controlling for all other observable aspects regarding the data and methods used in the primary studies, the meta-regression results indicate that hedging premiums found in the highest ranked journals are systematically different from the empirical findings reported in other journals and unpublished work. All other things being equal, the studies in the top journals report a significantly lower impact of hedging on firm value, by 2.0% on average. Accordingly, journal quality matters for the size of the reported hedging premium, which is in line with the ex-ante hypothesis that higher ranked journals report lower premiums due to a more rigorous review process.

North America vs. Other Regions. The estimated meta-regression coefficients for the world regions indicate the average difference in hedging premiums between the respective region and North America. The results show that the hedging premiums are not the same all over the world. According to Model (2), the markup in firm value through corporate hedging is 3.9% larger for firms located in Latin America as compared to firms located in North America. The meta-regression coefficient for East Asia indicates that corporate hedging is less valuable as compared to North America (-2.4% on average). In contrast, there is no systematic difference in the hedging premiums of European or South Asian firms as compared to North American companies. The result for Europe could be explained by the similarity of the European and North American markets in terms of both economic maturity and accessibility to hedging instruments. The lack of evidence for a systematic difference in the hedging premiums of South Asian firms may be due to the low number of observations in the sample, as only 3% of the estimates refer to South Asia. Nevertheless, regional clusters are only rough estimates of the real drivers for country-level differences in hedging premiums. Therefore, I extend the analysis in the next section by replacing the regional dummies with variables that measure the macroeconomic differences across the countries in the sample.

Risk Exposure Type. A breakdown of the hedging premiums by the three risk exposure types reveals significant differences. On average, IR hedgers and CP hedgers have a lower firm value of 2.6% and 2.3%, respectively, as compared to FX hedgers (Model 1). This notable spread in the hedging premium, especially the larger effect for FX hedging, is consistent with common sense in the literature (Allayannis et al., 2012; Jin and Jorion, 2006).

Firm Fixed Effects and Endogeneity Control. Including firm fixed effects in the primary regression models leads to a significant reduction in the size of hedging premiums by -2.3% on average. Accordingly, unobserved heterogeneity at the individual firm level creates an upward bias in the estimation of hedging premiums. Another source of bias in the primary estimation comes from endogeneity. When primary studies control for endogeneity in their estimation by using IV approaches or other related methods, they tend to report lower hedging premiums of -1.3% on average. This is especially interesting in light of the intensive discussion surrounding endogeneity issues and reverse causality between firm value and corporate hedging (Aretz and Bartram, 2010: 363-365). The results provide an indication that not accounting for endogeneity leads to an overestimation of the hedging premium. However, the statistical significance is not stable across the models and goes down to the 10% level in the reduced model (Model 5).

Derivatives Usage and Broader Hedging Definition. The results suggest that studies that equate hedgers with derivatives users find, on average, lower premiums of -5.1%. This implies that the firm value premium of using derivatives for hedging is not equal to the broader concept of a hedging premium including non-derivatives hedging methods as well.[45] In addition to an advanced measurement of the dependent hedging variable in the primary regression, some studies also include other hedging methods as control variables in their models. Omitting the control variable for operational hedging in the primary regression causes a downward bias in the size of reported premiums by -1.1% on average. However, this finding disappears in the reduced model (Model 5). For both variables, accounting for other hedging methods in the definition of the dependent variable and adding an independent variable for operational hedging as a control, hedging premiums increase when operational hedging is considered. Therefore, it appears that financial hedging in combination with operational hedging has a greater effect on firm value than financial hedging only. A similar finding is reported by Allayannis et al. (2001: 393), suggesting that financial and operational hedging are rather complements than substitutes, as a comprehensive corporate risk management approach may require multiple hedging strategies. The result that the effects of hedging are different when other hedging strategies are accounted for is also the subject of an ongoing discussion in the literature (among others, Amberg and Friberg, 2016: 86; Hoberg and Moon, 2017: 229-234).

After 2001. The breakpoint variable for the time period after 2001 indicates a distinct trend in hedging premiums over time. The time effect is reflected by a 2.5% larger premium for samples with an average sample year after 2001. This finding might be explained by the strong growth in international trade and export ratios, as well as increased global market volatility since the 2000s, especially during the financial crisis in 2007/2008 and the European debt crisis since 2009. Such enlarged uncertainty also affects the corporate risk exposures and thus might require more active risk management at the firm level, which may be rewarded by higher firm values.

[45] It should be noted that the number of observations including a broader hedging definition is rather small (5% of the total sample). Consequently, the results for this variable can only be interpreted as a first indication of differences between derivatives usage and other hedging strategies.

Further Results. Regarding the hedging measure, the results suggest significantly larger premiums of 0.9% when a hedging dummy variable is used compared to a continuous hedging measure. A driver for this result could be that the hedging dummy variable is less precise as it classifies firms as hedgers independently of their actual hedging volume. Moreover, it should be noted that the hedging premiums collected from primary models with a continuous hedging variable are evaluated at the sample mean hedging volume, i.e., they show the hedging premium that the study implies for an average hedger. In contrast, the hedging premiums for the dummy variable refer to the difference between hedgers and non-hedgers. Thus, the evaluation of the continuous variable at the sample mean might drive the systematically lower premiums as compared to models using a hedging dummy variable. As for the control variables, it can be observed that models that account for other risk exposures report lower premiums on average. This indicates that the different exposure types are correlated and omitting other exposures from the primary regression may bias the estimated hedging premium. In addition, the negative sign of the coefficient for managerial ownership indicates that adding this control variable to the primary regression reduces the hedging premium by -3.2% on average.

Publication Selection Bias. Even after controlling for various aspects of study design, the significant coefficient of the primary standard error indicates publication selection bias. Without publication selection, the hedging premiums' standard errors and the estimates for the premium should be independent. But if primary study authors actively change their data and methods to find hedging premiums to be large enough to offset large standard errors, i.e., to reach statistical significance, correlation between the estimates' standard errors and the hedging premiums occurs (Ashenfelter et al., 1999: 460; Stanley, 2005: 320-323). The sign of the coefficient is positive in all five models, confirming the results of the previous section that positive hedge premiums are systematically overrepresented.

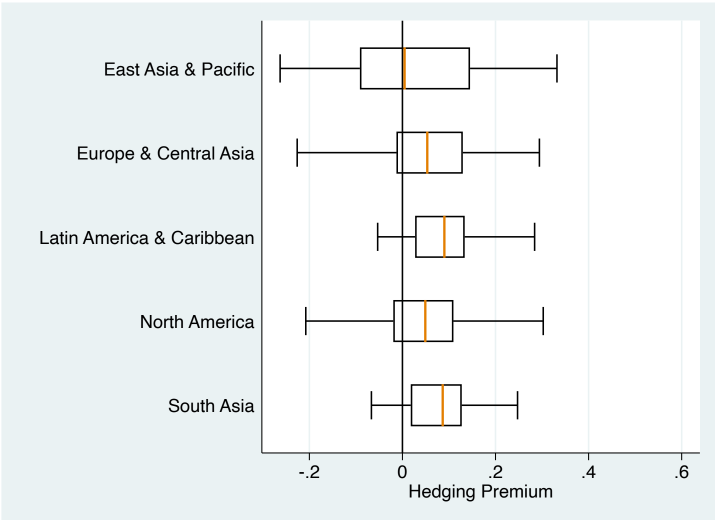
Alternative Weighting. The Models (3) and (4) use the alternative weighting schemes to estimate the meta-regression coefficients. When weighting all estimates reported in a study equally (Model 3), I can confirm the findings for the variables that control for fixed effects and CP hedgers. However, I cannot confirm the other results detected in the baseline model. In contrast, the impact of the regional clusters becomes stronger, showing larger premiums in Europe, South Asia, and Latin America than in North America. Using the combined weight of the inverse variance and the inverse number of estimates (Model 4) yields very similar results as in the baseline model. In contrast, the results for firm fixed effects and controlling for endogeneity disappear, as do the effects of the moderators measuring the presence of the control variables. However, using the alternative weights in Model (4), I find a systematically lower premium for studies using the simple market-to-book ratio to measure firm value as compared to studies using a more precise measure of Tobin's Q. Moreover, controlling for sample selection bias turns out to have a positive influence on the reported hedging premiums.

3.7.4. Country-Level Factors as Drivers of Heterogeneity

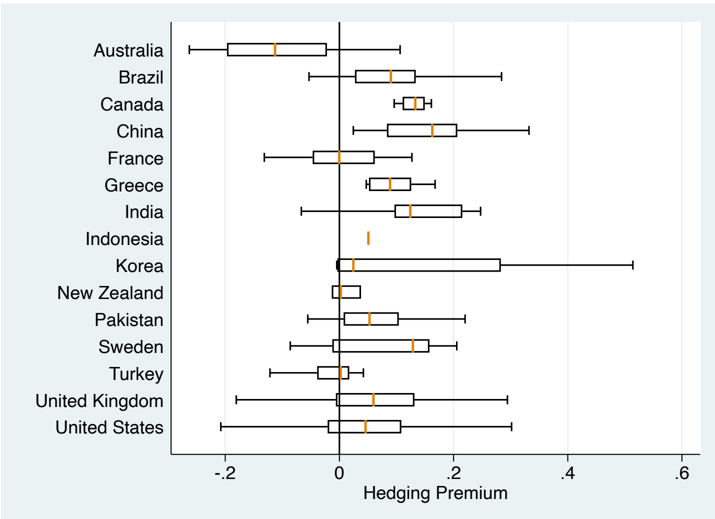
Figure 26 depicts the boxplots of the distribution of hedging premiums for each of the five geographical regions (Figure 26a) and for each of the 15 countries examined in the primary studies (Figure 26b).

Figure 26. Boxplots of the Regional Distribution of Hedging Premiums

(a) Hedging Premium per Region



(b) Hedging Premium per Country



Notes: The figures show (a) the boxplots of the hedging premiums observed for each geographical region and (b) the boxplots for each country examined in the primary studies. Outliers are excluded from the figure but included in all statistical tests. Results vary across studies, depending on the number of robustness checks reported and the subsamples analyzed.

The boxplots reveal differences in the hedging premiums across the world regions (Figure 26a). The widest range of premiums is reported for East Asia & Pacific, reaching from below -20% to over +30%. The median premium is lowest in East Asia & Pacific (0.4%) and highest in Latin America & the Caribbean (9%).

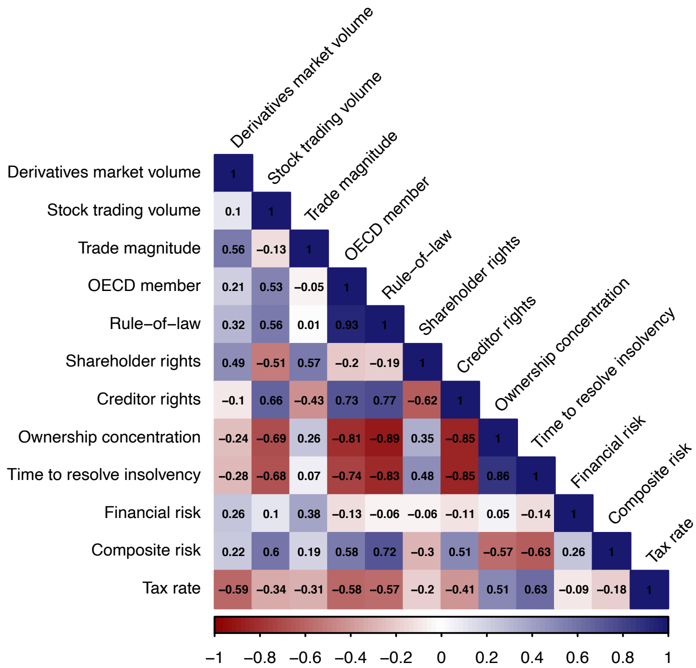
When breaking down the regional clusters into individual countries (Figure 26b), the heterogeneity increases. In Australia, where the examined risk exposures are mostly CP, the premiums are almost all negative. For Europe, the results are different. While the median hedging premium in France is only 1.4%, Sweden (9.6%) and the United Kingdom (6.5%) show large positive median premiums. In the U.S., the median premium is 6.2% and for China it is even 15.8%. Accordingly, there are important differences among the effects of corporate hedging on the firm value depending on the country and world region of the companies analyzed.

However, the regional clusters are only rough estimates of the geographical differences because they cannot quantify the actual macroeconomic and institutional dissimilarities among countries. An important strength of meta-analysis is that it allows adding new information from external sources to the original primary study based on the country and time period examined in the primary studies. For the hedging literature, this is especially valuable as many of the existing primary studies refer to data from a single country. Thus, by construction of their samples, they cannot analyze cross-country differences. With meta-analysis, I can build an international sample of hedging premiums from the 15 countries analyzed in the previous literature. Thereby, the regional dummies are replaced with the country-level variables defined in Table 8. The country-specific variables are assigned to the primary studies as average values of the time series corresponding to the country and time period examined in a study.

Figure 27 shows the Pearson correlation coefficients among the country-level variables introduced in Table 8. The heatmap of correlation coefficients among these variables reveals large interdependencies among the variables, with correlations (below) above (-)0.80. This is not surprising as many of the variables are complementary and measure a similar macroeconomic dimension. The correlation analysis also reveals some interesting findings. For example, ownership tends to be less concentrated in countries with a high rule-of-law index, the time required to resolve insolvency appears to be shorter in countries with strong creditor rights, and a country with a high derivatives market volume tends to have lower tax rates.

Given the large correlation coefficients among most of the macro variables, the meta-regression results are estimated separately for each country-level factor to avoid biases. Table 15 reports the corresponding results. The estimation is based on the baseline specification (Table 14, Model 1) but the regional clusters are replaced by the country-level variables (Table 8). Results with alternative weights are reported in Appendices B.4 and B.5.

Figure 27. Correlation Matrix of Country-Level Factors



Notes: The figure shows the Pearson correlation coefficients of the explanatory variables defined in Table 8.

Financial and Economic Development. The results show evidence that the access to derivative instruments for hedging, measured by the local market's derivatives trading volume, moderates how hedging impacts firm value. Accordingly, hedging premiums are larger in countries with less liquid derivatives markets. In those markets with lower derivatives market volumes, investors encounter additional costs and hedging at the firm level might be preferable to hedging at the individual shareholder level. If we see market liquidity as a proxy for transaction costs, this finding shows that, in contrast to basic hedging theory where hedging is often assumed to be costless (Smith and Stulz, 1985: 392-393), the costs of hedging are an important conditional factor for whether hedging is valuable or not.

This finding is supported by the results for the influence of the country's stock trading volume, which can be seen as a proxy for the development of financial markets in general. As further proxies for economic development, the country's trade magnitude and OECD membership confirm that economic and financial development drive how hedging influences firm value. Hedging premiums decrease with stronger economic and financial development. This might be driven by the fact that market inefficiencies, which provide rationales for corporate hedging, are lower in more developed countries.

Table 15. Country-Level Drivers of Hedging Premiums

	(1)	(2)	(3)	(4)
Panel A. Financial and Economic Development				
Derivatives market volume	-0.028**			
	(-2.43)			
Stock trading volume		-0.019**		
		(-2.19)		
Trade magnitude			-0.024***	
			(-2.66)	
OECD member				-0.025***
				(-2.65)
Constant	0.086***	0.153***	0.154***	0.102***
	(3.31)	(4.87)	(3.09)	(3.97)
	(5)	(6)	(7)	(8)
Panel B. Legality and Governance				
Rule-of-law	-0.008*			
	(-1.65)			
Shareholder rights		0.002		
		(0.62)		
Creditor rights			-0.003	
			(-1.18)	
Ownership concentration				0.020
				(0.81)
Constant	0.090***	0.077**	0.109***	0.077**
	(3.45)	(2.32)	(6.35)	(2.39)
	(9)	(10)	(11)	(12)
Panel C. Financial Distress and Taxes				
Time to resolve insolvency	0.018			
	(1.12)			
Financial risk		-0.060		
		(-1.32)		
Composite risk			-0.061	
			(-1.29)	
Tax rate				0.049**
				(2.46)
Constant	0.065*	0.300*	0.344	-0.106
	(1.84)	(1.73)	(1.64)	(-1.11)
Other controls from Table 14 included	Yes	Yes	Yes	Yes
No. of studies	71	71	71	71
No. of observations	1,016	1,016	1,016	1,016

Notes: This table presents the results of the same meta-regression model as shown in Table 14 (Model 1), but the dummy variables for geographical regions are substituted by the country-level variables defined in Table 8. The country-level variables are assigned to the hedging premiums based on the sample year and country of the underlying samples in the primary studies. Meta-regressions are estimated by WLS using the inverse of the estimates' squared standard errors as weights. The *t*-statistics in parentheses are based on robust errors, clustered at study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Legality and Governance. Legality and governance factors do not moderate the size of the hedging premium. Accordingly, I cannot confirm previous evidence by Allayannis et al. (2012) that external governance factors explain differences in the hedging premium. An explanation for this finding could be that, rather than country-level differences in governance and legality (examined in this study), the firm-level differences in governance (examined by Allayannis et al. 2012) could matter more for the firm value premium of hedging. However, a meta-analysis without access to the original primary data can only compare the between-study and between-country differences in the reported hedging premiums, but not the individual firms analyzed in the primary data sets, as this would require access to the full underlying data sets of the primary studies.

Financial Distress and Taxes. I find evidence that hedging premiums are, on average, larger in countries with higher tax rates. Following hedging theory (Mayers and Smith, 1982: 289-293; Smith and Stulz, 1985: 392-394), corporate hedging can smooth cash flows under a convex tax schedule. This smoothing effect means that taxable income is less likely to fall into the progressive range of the tax schedule, i.e., the range in which the marginal tax rate is higher than the average tax rate paid by a company. Although the actual tax rate is not an ideal proxy for the convexity in the tax schedule, it gives some indication of the potential for value creation due to lower tax payments induced by a more stabilized cash flow through hedging.

Overall, the results for the country-level determinants propose that beyond the common firm-level channels, country-specific differences also condition whether corporate hedging is valued by investors. So far, the theoretical explanations of corporate hedging largely abstract from country-specific features, thereby assuming that the theories are equally important for firms from different countries. However, the outcomes of this meta-analysis challenge this assumption and give rise to the proposition that the explanations for firm value creation through hedging should be extended by a further dimension covering the country-specific surroundings, especially a country's financial market conditions and regional tax rate schemes.

3.7.5. Best Practice Estimates

The heterogeneity analysis presented in Tables 14 and 15 reveal that the hedging premium is conditional on various study-specific and country-specific variables. To consider these contingency factors also in the estimation of the mean hedging premium, I take advantage of meta-analysis to create a synthetic study and predict the hedging premium by substituting 'best practice' values for the explanatory variables.

Best practices values are predicted for two different scenarios: (i) using the estimated regression coefficients of the baseline model in Table 14 (Model 1) to explore the impact of differences in the methodological variables, and (ii) using the estimated regression coefficients from the analysis of the country-level factors in Table 15.

Although it should be noted that any specification of such a benchmark study remains subjective, I define the best practice case as follows. First, I filter out publication selection and set the impact of the standard error to zero. Next, I insert a value of one for the breakpoint variable to model a study examining recent sample data. Moreover, I prefer a study controlling for ex-ante risk exposure and using an advanced statistical approach with robust standard errors. The best practice model also controls for self-selection bias, endogeneity, and firm fixed effects. For the firm value proxy, I assume an advanced measurement of Tobin's Q. Moreover, I estimate the hedging premium for a top journal article because the meta-regression results have shown a systematic difference in hedging premiums from higher ranked journals. To minimize the impact of misspecification bias, a model is selected that includes all of the analyzed control variables. As there is no clear preference for the remaining variables, I insert the sample mean for these variables. For example, I set the sample mean for the measurement of the hedging variable (dummy vs. continuous variable) because both measures come with caveats and there is no clear preference as to which is the better proxy. Due to the fundamental differences in the value effects for the different risk exposure types, all estimates are computed separately for FX, IR, and CP hedgers.

To quantify how the mean of the hedging premium changes due to model misspecification in the primary studies, Table 16 reports the predicted values for the best practice model without controlling for important aspects in the model specification (Panel A). In addition, I add the derivatives markets volume and the tax rate as important findings from the analysis of country-level heterogeneity. The model is estimated with the same configuration for the aforementioned variables. The results (Panel B) are reported for a scenario when substituting the sample maximum (high) or, alternatively, the sample minimum value for the country-level variables (low).

The basic predictions imply a hedging premium of 1.8% for FX hedgers, -0.8% for IR hedgers, and -0.6% for CP hedgers. These values can be interpreted as the mean estimates for the hedging premium implied by the entire body of empirical literature, while assuming a study with best practice research design after controlling for publication selection and misspecification biases. Accordingly, taking all available estimates in the literature and assuming an 'ideal' study design, I find the values reported in Table 16.

Panel A shows the impact of hedging on firm value when deviating from the best practice study design. If firm fixed effects are omitted, I find larger and positive premiums for all three exposure types. When leaving out the control for endogeneity as well, the resulting bias increases the mean premiums even further. Finally, omitting relevant control variables for operational hedging and managerial ownership, which have been shown to be important controls in the previous meta-regression (Table 14), creates an additional upward bias in the hedging premium. This experiment illustrates the strong impact of model specification on the estimated results. By changing the model design, hedging premiums change heavily. It can be concluded that the method chosen by the authors of the primary study has a strong impact on the final results to be reported.

Table 16. Best Practice Estimates of the Hedging Premium

	(1) FX Hedgers	(2) IR Hedgers	(3) CP Hedgers
Overall best practice	1.8%	-0.8%	-0.6%
Panel A. Data and Model Choices			
Without control for firm fixed effects	4.0%	1.4%	1.7%
Without control for firm fixed effects and endogeneity	5.3%	2.8%	3.0%
Without control for firm fixed effects, endogeneity, operational hedging, and managerial ownership	7.4%	4.8%	5.1%
Panel B. Country Level Factors			
Derivatives market volume (high)	0.7%	-1.9%	-1.5%
Derivatives market volume (low)	2.7%	0.1%	0.6%
Tax rate (high)	6.7%	4.0%	4.6%
Tax rate (low)	0.9%	-1.9%	-1.2%

Notes: This table reports predicted hedging premiums obtained by substituting best practice values for the variables in the meta-regression. The overall best practice and Panel A results are estimated from meta-regression results in Table 14 (Model 1). Panel B is based on the results from Table 15. Meta-regressions use the inverse of the estimates' squared standard errors as weights.

The best practice estimates for the important country-level moderators (Panel B) uncover the implied hedging premium when accounting for differences in derivatives market development and country-specific tax schemes. When derivatives volumes are low, predicted premiums are positive for the three exposure types and show values of up to 2.7% for FX hedgers. In contrast, for regions like the U.S. with highly developed derivatives markets, the predicted hedging premium is only 0.7% for FX hedgers and negative for the other two risk exposures (-1.9% for IR hedgers and -1.5% for CP hedgers). Finally, when inserting the sample maximum of the country's tax rate, I find large positive effects of hedging and low or even negative premiums for countries with lower tax rates.

3.7.6. Further Analyses

The main results from the previous sections are complemented by several additional analyses to further explore the drivers of heterogeneity.

Reverse Causality. The relation between hedging and firm value might be affected by reverse causality, as larger and multinational firms might be the better-run firms that understand hedging better than other firms and therefore implement better strategies (Pérez-González and Yun, 2013: 2143-2144). To address endogeneity problems, the meta-regression includes a moderator variable that marks all studies that account for endogeneity in the estimation of the primary regression model. The estimated coefficients for this moderator variable indicate whether studies controlling for endogeneity find systematically different hedging premiums compared to studies that do not control for an endogenous relation between hedging and firm value. Since I cannot observe the hedging behavior for the underlying sample of firms that are examined in the primary studies but rather extract the aggregated results of the hedging-firm value relation, single firms cannot be matched based on their hedging behavior, but only the aggregated estimates from the primary studies can be matched.

Table 17. FX Hedging Premiums after Matching on IR and CP Hedging

	(1) Interaction Term	(2) Subsample
FX hedgers × Control for other risk exposures	0.005	
	(0.68)	
FX hedgers	0.017^{***}	0.011^{***}
	(3.05)	(2.81)
Control for other risk exposures	-0.021^{***}	
	(-5.29)	
Constant	0.047	0.310^{***}
	(1.60)	(5.83)
Other controls from Tab. 14 included	Yes	Yes
No. of studies	71	18
No. of observations	1,016	307

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2). Unreported control variables are identical to those in Table 14. Meta-regressions are estimated by WLS using the inverse of the estimates' squared standard errors as weights. Model (1) includes an interaction term between the dummy variable for FX hedgers and another dummy variable indicating whether multiple risk exposures are estimated in the same primary regression (suggesting that the reported hedging premiums do not suffer from a bias due to the omission of other hedging exposures). Model (2) is based on a reduced sample of all estimates observed from models with multiple risk exposures estimated in the same primary regression. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

An additional analysis (Table 17) investigates whether hedging premiums for FX hedgers are still different when FX premiums are jointly estimated with IR and/or CP hedging. This is the case when multiple risk exposure types are analyzed in the same regression. Therefore, an interaction term between the FX hedging dummy variable and the control variable for multiple risk exposures (*Control for other risk exposures*) is added, indicating primary regressions with multiple exposures tested in the same model. In a second test, I reduce the sample to all estimates from models testing several exposures simultaneously.

The significant and positive coefficients for FX hedgers in the additional tests in Table 17 (Models 1 and 2) show that the FX hedging premium is still significantly larger than IR and CP hedging premiums, even when observed from models that correct for the impact of other risk exposures. The insignificant meta-regression coefficient of the interaction term in Model (1) indicates that the results for a larger premium of FX hedgers are not biased by the effects that the hedging of other exposure types has on firm value.

Continuous Hedging Variable. I further investigate the measurement of the hedging variable. The heterogeneity analysis (Table 14, Model 1) revealed that studies measuring hedging by a dummy variable as compared to a continuous variable report larger hedging premiums on average. However, the impact of the continuous hedging measure depends on the accounting standards under which the hedging volume that the primary studies analyze was disclosed in the companies' annual reports. In an additional analysis reported in Table 18, I examine the impact of major changes in international accounting standards (Models 1 and 2). Appendices B.7 and B.8 show the results using the alternative weighting schemes. The models refer to the subsample of estimates for the continuous hedging variable with breakpoint variables for the issuance and effective dates of the Statements of Financial Accounting Standards (FAS) No. 133, Accounting for Derivative Instruments and Hedging Activities in 1998, and the introduction of the International Accounting Standard (IAS) 39, Financial Instruments: Recognition and Measurement in 2003.[46] As the latest average sample year is 2012.5, the most recent accounting changes, especially the introduction of the International Financial Reporting Standard (IFRS) 9, are not applicable to the sample. For this approach it must be considered that the 71 primary studies in the sample examine 15 different countries. Most of these countries have their own country-specific implementations of the international accounting rules. However, I see the introduction of the FAS 133 and IAS 39 as major changes in the global accounting standards that are likely to be adopted by the local rules. To further explore the continuous hedging measure, I analyze how the different subgroups of the continuous hedging measure interact with firm value. Model (3) breaks down the continuous hedging variables by the type of measurement, which are either notional amounts, actual hedge ratios, fair values, or other hedging measures.

From an information asymmetry point of view, I would expect decreasing hedging premiums as stricter accounting rules require more detailed disclosure of hedging instruments, which

[46] It should be noted that most of the primary studies in the sample examine time periods covering several years of firm data (the average length of the sample period is 6.5 years). Accordingly, the reported hedging premiums that I collect refer to the corresponding sample period of the studies. Thus, I cannot exactly assign the breakpoints to a specific year, but rather to the average sample period per study.

reduces the information asymmetry between the firm and the shareholders. If shareholders have better access to information on the risk profile and the hedging strategy of the firm, they might prefer to hedge at the individual investor level rather than at the firm level. However, the results show that the continuous hedging premiums increase with each change in the accounting standards. The introduction of FAS 133 and IAS 39 raises the average hedging premium by 3.7% and 1.9% as compared to the average time period before the new accounting standard was issued (Model 1). The results are similar when using the effective dates of the new standards (Model 2). On average, the new international accounting rules come with stricter disclosure requirements for firms to report their hedging strategies and instruments to investors. However, because the accounting changes are tied to a specific sample year, the results of Models (1) and (2) may reflect the trend of increasing hedge premiums over time, as shown by the breakpoint dummy in Table 14, rather than the impact of the accounting changes.

Table 18. Impact of Accounting Standards and Measurement on Continuous Hedging

	(1) Accounting Changes (Issue Dates)	(2) Accounting Changes (Effective Dates)	(3) Measurement of Continuous Hedging Variable
Issue of FAS 133 in 1998	0.037^{***}		
	(4.92)		
Issue of IAS 39 in 2003	0.019^{***}		
	(2.67)		
Effective date of FAS 133 in 2000		0.024^{***}	
		(3.04)	
Effective date of IAS 39 in 2005		0.015^{***}	
		(2.76)	
Fair values vs. notional amounts			-0.022^{*}
			(-1.80)
Actual hedge ratios vs. notional amounts			-0.015^{*}
			(-1.88)
Other measures vs. notional amounts			0.053^{***}
			(4.05)
Constant	0.086^{***}	0.074^{***}	0.031
	(4.05)	(3.43)	(1.42)
Other controls from Tab. 14 included	Yes	Yes	Yes
No. of studies	40	40	40
No. of observations	326	326	326

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2) using the inverse of the estimates' squared standard errors as weights. Unreported control variables are identical as in Table 14. Reported coefficients refer to the alternative variables included for robustness analysis. Model (1) includes two breakpoint variables referring to the issuance year of major accounting changes relevant for the reporting of hedging instruments (FAS 133 and IAS 39). Model (2) refers to the year in which the accounting changes became effective. The omitted base category is the time period before 1998 (Model 1) and before 2000 (Model 2). The breakpoint is assigned to the hedging premiums from the primary studies based on the average sample year examined in each study. Model (3) breaks down the continuous hedging variable in the different categories of measuring the extent of hedging. The omitted base group are notional amounts of hedging instruments reported in annual reports. The *t*-statistics in parentheses are based on robust errors, clustered at the study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In addition, the results for the different measurements of the continuous hedging variable (Model 3) suggest that fair values and actual hedge ratios yield systemically lower hedging premiums of 2.2% and 1.5% as compared to notional amounts (omitted base category). However, the results show only a low level of statistical significance (at the 10% level). The *Other Measures* include the number of different contracts used for hedging and other less precise measures of hedging that are generally considered not as reliable measures for the actual corporate hedging behavior. The large surplus of 5.3% between the other hedging measures and the notional amounts give reason to suspect that the results for the other hedging measures might be biased towards reporting positive premiums.

Model Selection. In the analysis of heterogeneity (Table 14), all 25 explanatory variables are included in the same meta-regression. A problem that comes with this approach is that it ignores model uncertainty, as not all variables might be equally important (Havranek et al., 2017: 162-163). When adding too many redundant and irrelevant variables in the meta-regression, the precision of the point estimates for the effects of the explanatory variables could be diminished (Candolo et al., 2003: 165). However, we do not know from theory which of the variables should be included to obtain efficient estimates. To address this issue of model uncertainty, I excluded the redundant variables by running a stepwise meta-regression following the general-to-specific approach (Table 14, Model 5). However, the stepwise regression approach is problematic as, among other concerns, stepwise regression might identify an incorrect best set of explanatory variables and leave out important variables by accident (Smith, 2018: 2).

To address the model uncertainty challenge, I follow recent developments in meta-regression research and apply Bayesian Model Averaging (BMA) (among many others, Duan et al., 2020: 235-238; Havranek and Irsova, 2017: 381-386; Matousek et al., 2022: 343-346; Xue et al., 2020: 87-91; Zigraiova et al., 2021: 15-19). The general idea of BMA is to run regressions with different subsets of the possible combinations of explanatory variables instead of selecting just one of the possible regression specifications (Raftery et al., 1997: 180). Thus, BMA can be thought of as a robustness check with many different subsets of explanatory variables. Across this set of different regression models, BMA produces a weighted average. The weights used for averaging are assigned based on the Posterior Model Probability (PMP), which is a measure of the model fit comparable to the adjusted R^2. The weight increases with model fit and decreases with model complexity.

However, a full enumeration of all possible subsets of explanatory variables would require too much computing capacity. In this meta-analysis, I have 25 explanatory variables, which would result in 2^{25} regression models to choose from. Thus, a Markov Chain Monte Carlo algorithm[47] is applied to consider only the most promising models, which are the models with the highest posterior model probabilities. The distribution of the model parameters across the individual models is captured by the posterior means and standard deviations, which are the weighted average of the coefficients across all models using the posterior model probability for the weighting (Hinne et al., 2020: 202). The Posterior Inclusion Probability (PIP) for each explanatory variable is the sum of the posterior model probabilities across all regression specifications including this variable. The PIP denotes the probability that a variable is included in the true regression model (Amini and Parmeter, 2020: 8).

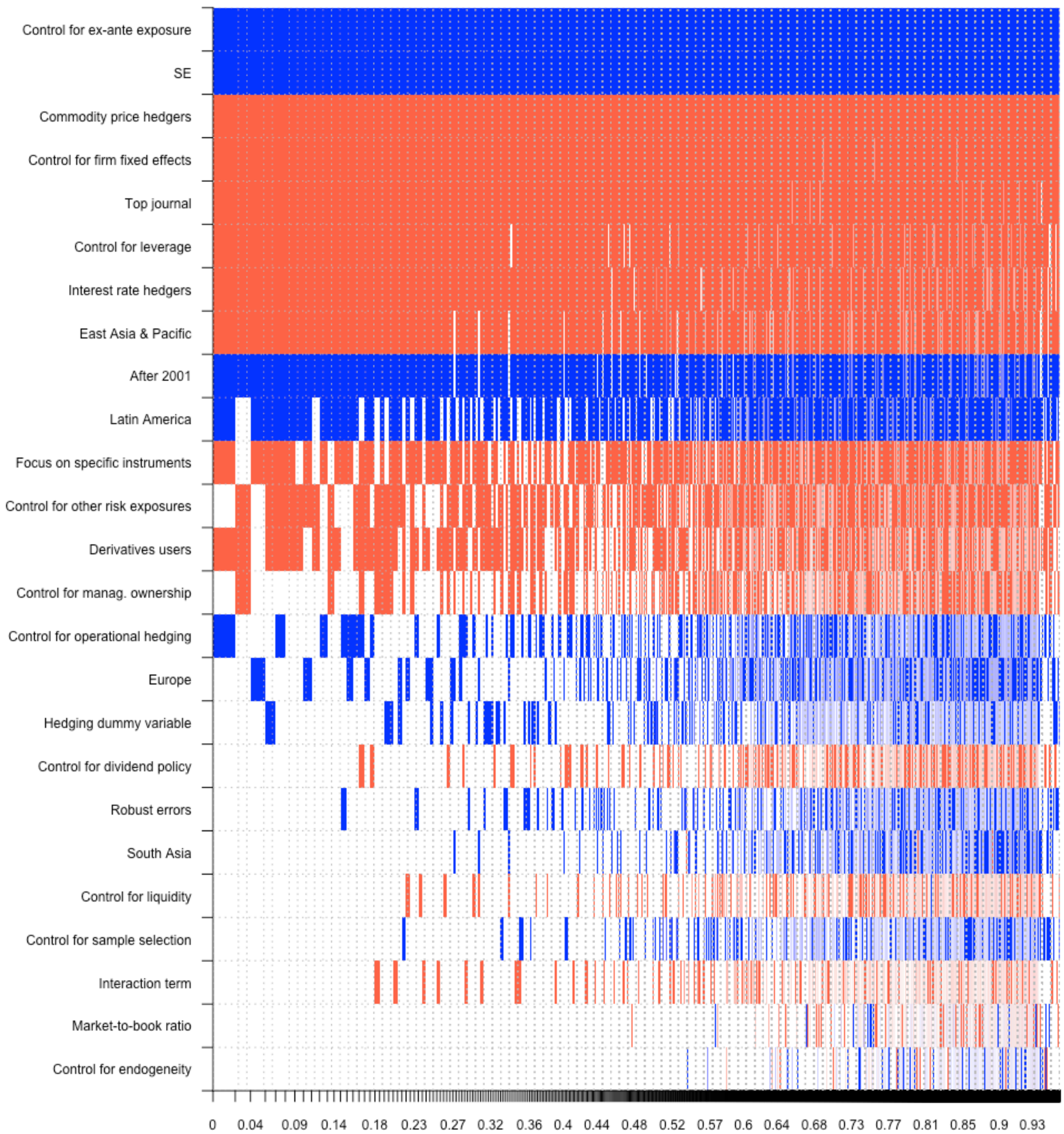
BMA is responsive to the priors regarding the model (model prior) and the coefficients (g-prior) that are used to define the expectation of the researcher. I follow Matousek et al. (2022: 21) and use the dilution prior of George (2010), which considers collinearity of the variables in the model by multiplying the model probabilities by the determinant of the correlation matrix of the variables that are added to the model. A higher collinearity of the variables drives the determinant closer to zero, and hence, the model weight decreases. For the g-prior, the unit information prior (UIP) is used, which is based on the assumption that the information contained in the prior is equal to the information in one observation (Feldkircher and Zeugner, 2009: 9).

Figure 28 visualizes the BMA results. The moderator variables are shown in the rows, with the variables with the highest PIP at the top. The columns denote the different model specifications. The width of the columns displays the PMP and thus the weight of each regression. Cumulative PMPs are shown on the horizontal axis. The blue-colored cells indicate the inclusion of a variable in the model with a positive sign and the red-colored cells indicate the inclusion of a variable with a negative sign. Uncolored cells indicate that a variable is not included in the model.

Table 19 shows the numerical results for the BMA analysis. The posterior mean, standard deviation, and the PIP are reported in Panel A. Following Havranek and Irsova (2017: 384), Panel B shows the results of a robustness check from a frequentist check using an OLS regression with standard errors clustered at study level and country level. The OLS regression includes only the variables with a PIP above 0.75. These variables are identified as substantial in the BMA exercise. Following Havranek et al. (2015a: 105), an effect is 'weak' if the PIP is between 0.5 and 0.75, 'substantial' if the PIP is between 0.75 and 0.95, 'strong' if the PIP is between 0.95 and 0.99, and 'crucial' for values of the PIP above 0.99.

[47] For the implementation, I make use of the *R* package 'bms' developed by Feldkircher and Zeugner (2009).

Figure 28. Bayesian Model Averaging



Notes: This figure presents the results for Bayesian Model Averaging (BMA) of the moderator variables. The response variable is the hedging premium. Columns represent the individual models. Variables are sorted by their posterior inclusion probability (PIP) in descending order. The cumulative posterior model probabilities are shown on the horizontal axis. The figure shows only the best 5,000 models. The estimation is based on the unit information prior (UIP) suggested by Eicher et al. (2011) and the dilution prior recommended by George (2010). The blue-colored cells indicate the inclusion of a variable with a positive sign, the red-colored cells indicate the inclusion of a variable with a negative sign, and uncolored cells indicate that a variable is not included in the model.

Table 19. Numerical Results for Bayesian Model Averaging

	Panel A. Bayesian Model Averaging			Panel B. Frequentist Check (OLS)		
	Posterior mean	Posterior Std. Dev.	PIP	Coefficient	t-statistic	p-value
Journal Quality						
Top journal	-0.060	0.016	0.989	-0.047	-6.03	<0.001
Sample Year						
After 2001	0.033	0.014	0.898	0.025	1.20	0.229
Geographical Region						
Europe vs. North America	0.013	0.022	0.306	0.044	1.69	0.091
East Asia & Pacific vs. North America	-0.063	0.026	0.915	-0.049	-1.11	0.267
South Asia vs. North America	0.009	0.026	0.149	0.346	0.99	0.323
Latin America vs. North America	0.042	0.031	0.764	0.071	3.43	0.001
Measurement of Hedging						
IR vs. FX hedgers	-0.044	0.017	0.942	-0.045	-2.99	0.003
CP vs. FX hedgers	-0.063	0.012	1.000	-0.071	-3.11	0.002
Hedging dummy variable	0.004	0.010	0.220			
Derivatives users	-0.032	0.030	0.610			
Control for ex-ante exposure	0.061	0.012	1.000	0.071	3.77	<0.001
Focus on specific instruments	-0.034	0.025	0.719	-0.058	-2.97	0.003
Measurement of Firm Value						
Market-to-book ratio	0.000	0.003	0.045			
Estimation Characteristics						
Control for firm fixed effects	-0.039	0.011	0.992	-0.048	-2.34	0.019
Control for endogeneity	0.000	0.003	0.037			
Control for sample selection bias	0.006	0.017	0.140			
Robust errors	0.004	0.009	0.180			
Interaction term	-0.003	0.010	0.139			
Control Variables						
Control for other risk exposures	-0.020	0.017	0.658			
Control for operational hedging	0.008	0.013	0.340			
Control for managerial ownership	-0.013	0.019	0.389			
Control for liquidity	-0.003	0.009	0.141			
Control for leverage	-0.072	0.022	0.963	-0.085	-2.53	0.011
Control for dividend policy	-0.006	0.016	0.193			
Publication Selection						
SE	0.740	0.065	1.000	0.760	10.23	<0.001
Constant	0.115	N.A.	1.000	0.081	1.85	0.064
No. of studies		71			71	
No. of obs.		1,016			1,016	

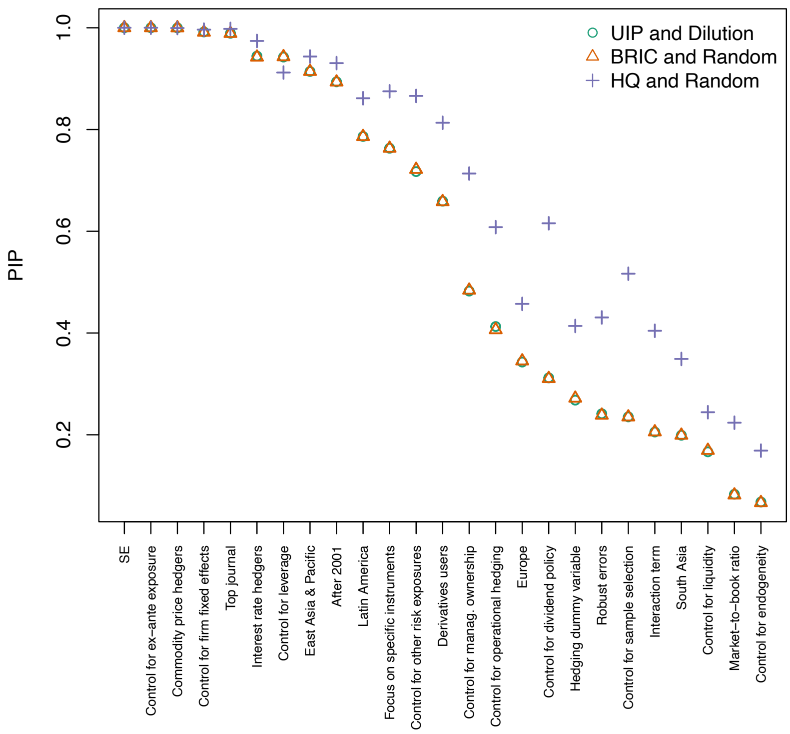
Notes: This table reports the numerical results of Bayesian Model Averaging (BMA). PIP = Posterior Inclusion Probability. The frequentist check (Panel B) includes only substantial explanatory variables with PIPs larger than 0.7. The p-values in the frequentist check are based on standard errors that are clustered at the study level and country level.

The BMA results confirm the key findings for several variables from the baseline meta-regression model (Table 14, Model 2), as they have a PIP above 0.75, which implies a substantial effect of the variable: *Top journal, After 2001, IR hedgers, CP hedgers, Control for firm fixed effects,* and *SE*. As the findings for these variables are consistent across the classical meta-regression model (Table 14) and the BMA results, I see these variables as the key drivers of heterogeneity. In contrast, I find that some variables appear less important in the BMA analysis as compared to the main results in Table 14: *Hedging dummy variable, Derivatives users, Control for endogeneity, Control for other risk exposures, Control for operational hedging,* and *Control for managerial ownership*. And some variables that are found to have no explanatory power for the heterogeneity across hedging premiums become important heterogeneity drivers in the BMA analysis: *East Asia & Pacific, Control for ex-ante exposure,* and *Control for leverage*. The results of the frequentist check indicate that the sign and the size of the regression coefficients model are consistent with the BMA results. Variables with a high PIP are in most cases statistically significant at the 5% level, except of *After 2001* and *East Asia & Pacific*.

As the priors are an important determinant of the BMA results, I perform several robustness tests with alternative priors. The UIP and dilution priors represent the baseline model, the results of which are reported above in Figure 28 and Table 19. As a robustness check, I use alternative priors. Hence, I apply the beta-binominal random model prior that assigns an equal prior probability to each model size (Ley and Steel, 2009) and combine the random model prior with the BRIC g-prior and the Hannan-Quinn g-prior (Fernández et al., 2001). Figure 29 illustrates the change in the relative importance of the explanatory variables when different priors are used for BMA.

The robustness tests with alternative priors reveal that the implied relative importance of the key explanatory variables with a PIP > 0.75 in the baseline BMA model changes only slightly when applying different priors. The alternative results show that the Hannan-Quinn and random priors lead to larger PIPs. The differences in the PIPs of the different variables increase for *Latin America* and the subsequent variables with a similar or lower PIP. For these variables, the Hannan-Quinn and random priors produce larger PIPs than the other two priors, which lead to almost identical PIPs for all variables. In summary, it can be concluded that the findings for the key variables are rather robust against changes of the priors.

Figure 29. Posterior Inclusion Probabilities for Different Priors



Notes: This figure shows the Posterior Inclusion Probabilities (PIP) for different model priors. UIP and Dilution = Priors according to Eicher et al. (2011) and George (2010). BRIC and Random = Benchmark g-prior for parameters with the beta-binominal model prior. HQ and Random = Asymptotically mimics the Hannan-Quinn criterion.

3.8. Discussion

This section is dedicated to the discussion of the main results. It compares the findings with the previous meta-analyses in the field and reflects the methodological approach in this chapter.

3.8.1. Comparison with Previous Meta-Analysis Results

As described in Section 3.3.2, there are three meta-analytical studies relevant for the comparison of the results. The study characteristics and key results of these meta-analyses are summarized in Table 20.

Study Characteristics. The effect size measures differ across the previous meta-analyses that use either Hedges' *g*[48] or (partial) correlations. Accordingly, the findings for the mean effect sizes are not comparable in terms of the actual size of the effect. From a methodological perspective, the previous studies apply 'simple' meta-methods to calculate weighted averages of the effect sizes and subgroup analysis and to compare the meta-averages for subgroups. In contrast, this chapter is the first meta-regression analysis on hedging and firm value.

[48] Hedges' *g* is an effect size that measures the difference between two groups (e.g., hedgers and non-hedgers) (Borenstein et al., 2010: 25-27).

Table 20. Comparison with Results of Previous Meta-Analyses

	Arnold et al. (2014)	Bessler et al. (2019)	Hang et al. (2020b)	This Chapter
Study Characteristics				
Effect Size	Hedges' *g*	Partial correlation	Pearson correlation	Hedging premium
Main Type of Meta-Analysis	Traditional meta-analysis (Hedges and Olkin)	Traditional meta-analysis (Hunter and Schmidt)	Meta-analytic structural equation modeling	Meta-regression analysis
No of Studies	25	47	89	71
No. of Effect Size Estimates	25	47	89	1,016
Results				
Mean Effect Size (*p*-value in parentheses)	All exposures -0.011 (>0.10)	All exposures 0.044 (<0.01) FX 0.063 (<0.01) IR 0.016 (>0.10) CP 0.004 (>0.10)	All exposures 0.019 (>0.01)	All exposures 0.014 (<0.01) FX 0.025 (<0.01) IR 0.001 (>0.1) CP -0.010 (<0.1)
Analysis of Publication Selection Bias	Not reported	Not analyzed	No publication selection bias detected	Mid to strong publication selection bias detected
Main Sources of Heterogeneity	Not analyzed	Type of risk exposure, publication status, regional differences, dummy vs. continuous hedging variable, applied regression method (OLS vs. others)	Sample year, single-industry vs. cross-industry, regional differences, macroeconomic differences (esp. legal system)	Journal quality, sample year, type of risk exposure, measurement of hedging variable, applied regression method (esp. control for fixed effects and endogeneity), macroeconomic differences (esp. derivatives market volume and tax rate)

Notes: This table reports the study characteristics and key results reported in the previous meta-analyses on the hedging-firm value relationship. The mean effect size is the weighted average according to the applied meta-analysis method. As the studies use different effect size measures, the actual size of the mean effects is not comparable. As the previous meta-analyses do not apply meta-regression analysis, I use the mean effect size for the model reported in Table 11 (Panel B) for comparison, which is the weighted mean effect size corrected for publication selection bias (PEESE). The drivers of heterogeneity for this chapter refer to the main results reported in Table 14 and Table 15.

Mean Effect Size. Comparing the mean effect sizes shows that Arnold et al. (2014: 453) find a negative overall effect of -0.011 in a sample of 25 primary studies reporting mean differences (measured by Hedges' *g*) for the market-to-book ratio of hedgers compared to non-hedgers. However, the effect is not statistically significant. Bessler et al. (2019: 227) report a mean correlation of 0.044 between derivatives hedging and firm value that is statistically significant at the 1% level. However, in terms of economic significance, this effect is rather small (Cohen, 1988: 285-287). The effect increases to 0.063 in the subsample of FX hedgers and almost disappears for IR and CP hedgers. In another meta-analysis, Hang et al. (2021b) control for the joint effect of corporate hedging and leverage. The correlation between the hedging dummy variable and their firm value measure is 0.019, but this effect is not statistically significant (Hang et al., 2021b: 4907). The results in this chapter confirm the overall positive but small effect of hedging on firm value reported by Bessler et al. (2019) and Hang et al. (2021b). In line with Bessler et al. (2019), I also find that FX hedging is associated with a consistently larger impact on firm value. The results deviate slightly for CP hedging, as this chapter reveals a negative and small effect for CP hedging that is statistically significant, whereas Bessler et al. (2019: 227) find a positive effect but no statistical significance.

Analysis of Publication Selection Bias. Among the previous three meta-analyses, only Hang et al. (2021b: 4910) present results for the analysis of publication selection and find no evidence for selective reporting in their study. In contrast, the analysis of publication selection bias in this chapter finds mid-to-strong publication selection bias. The difference in the results of this study as compared to Hang et al. (2021b) could be driven by the different effect size measures, Pearson correlations as compared to hedging premiums, as well as the difference in the samples, which is 89 direct correlations between hedging and firm value in Hang et al. (2021b: 4905) as compared to 1,016 effect size observations in this chapter.

Drivers of Heterogeneity. Two of the previous meta-analyses examine drivers of heterogeneity via subgroup analysis. Accordingly, the effect of each moderating variable is examined in a separate analysis. In this chapter, I examine the simultaneous effect of a large set of heterogeneity drives in a meta-regression analysis. Although this chapter examines more factors than the previous meta-analyses, some of the moderator variables are also investigated in the earlier studies. My findings confirm the systematic differences in the hedging-firm value relation depending on the risk exposures (Bessler et al., 2019: 227), the publication status of the primary study (Bessler et al., 2019: 227), the different geographical regions (Bessler et al., 2019: 227; Hang et al., 2021b: 4911), the definition of the hedging variable (Bessler et al., 2019: 227), the impact of the specification of the regression in the primary study (Bessler et al., 2019: 227), and the average sample year of the primary data set (Hang et al., 2021b: 4911). However, the direction of the effects differs for some of the variables. Bessler et al. (2019: 227) find that the mean correlation is larger for continuous hedging measures, while in this chapter I find that the effect is larger for a binary hedging measure. For the regional differences, Bessler et al. (2019: 227) find a smaller hedging effect in the U.S. as compared to Europe, while Hang et al. (2021b: 4911) find larger effects for Europe. In contrast, I could not detect a systematic difference for the regional clusters comparing Europe and North America (Table 14, Model 2). Moreover, in contrast to the increasing hedging premium over time as detected in this chapter, Hang et al. (2021b: 4911) find that the correlation between hedging and firm value decreases for studies with an average sample year after 2008. For the other variables, I can confirm the direction and size of the effects reported in the prior meta-analyses.

3.8.2. Challenges and Limitations

As with any other empirical method, meta-analysis comes with limitations, which are outlined together with a discussion on how they are addressed in this chapter.

Unpublished and Low-Quality Studies Should Be Excluded. The meta-sample in this chapter includes effect size estimates from studies published in top field journals, but the sample also covers studies not published in leading outlets as well as unpublished work. An alternative approach by Slavin (1986, 1995) is the 'best practice synthesis' that only considers 'good' studies. However, the main caveat of this approach is the decision regarding what a good study is. Hence, the approach comes with strong subjectivity for the selection of the primary studies. Moreover, focusing only on top journals would lead to a significantly smaller data set and a reduction of variation in the collected estimates, which is indeed necessary for the statistical identification of drivers that are responsible for the wide variation in the hedging-firm value effects. Due to these challenges of being selective, I follow Stanley and Doucouliagos (2012: 19) and include all primary studies that are in line with my inclusion criteria (Section 3.4.2). Moreover, the meta-regression model explicitly accounts for differences in research methods, model specifications, and data, which are all factors for the quality of a primary study. In addition, I also account for quality through inverse variance weighting in the WLS meta-regression estimation. This approach assigns larger weights to more precise estimates. Taking statistical precision as a measure for quality, all estimates are weighted by quality.

A Large-Sample Primary Study is More Powerful than Meta-Analysis. Meta-analysis accumulates the empirical effects on a certain phenomenon that are reported in the previous literature and uncovers the determinants of variation in existing empirical findings via statistical analysis. Thereby, it manifests several distinctive features as compared to original primary studies. Especially in the hedging literature, it is challenging to construct data samples covering many countries over a long time period due to difficulties in the manual data collection and limited availability of corporate hedging data. Even if it were possible to analyze an international data set of many companies over many years, the results would still rely on the individual study design, such as the coded hedging data, variable definitions, model specification, estimation methods, as well as other idiosyncratic aspects of the study. At the meta-level, I can control for the impact of these idiosyncratic characteristics of research design as well as the various factors that might induce biases (especially, publication selection and model misspecification). In this context, it is also important to note that detecting and controlling for publication selection can never be done on the level of an individual study, as *"publication selection is caused by the process of conducting empirical research itself"* (Stanley and Doucouliagos, 2012: 4). Moreover, bringing together a variety of primary studies from different authors minimizes random sampling error by averaging across many estimates.

The Collected Estimates are Not Independent. In contrast to the aggregation of medical trials, for which meta-analysis was originally designed, the regression results collected in economics are usually not independent. At the meta-level, I encounter non-independent observations at multiple levels because various estimates are collected from the same study (within-study dependency), authors from different studies may examine data sets of similar companies and countries (between-study dependency), or the regression models in the same study might include more than one hedging measure, e.g., for different risk exposures (within-model dependency). The sources of dependencies in a meta-study are similar to those in a primary study. For example, a global panel data set in a primary study may lead to dependencies because observations from the same country, the same time period, or multiple observations from the same company over several years are pooled in the same sample. Therefore, the problem of non-independent samples is probably no different in a meta-study compared to primary research studies. To consider dependence in the estimation, meta-regression analysis applies the same remedies as a primary study to account for different sources of non-independent observations. To accommodate the problem of correlated effect size estimates within studies and between studies, the meta-regression models are estimated with robust standard errors clustered at the study level and the country level. Moreover, I include a control variable (*Control for other risk exposures*) in the meta-regressions to account for estimates taken from the same model.

Meta-Analysis Compares Apples with Oranges. Meta-analysis in economics commonly builds on heterogenous estimates that are produced by different methods and data sets. I explicitly control for these differences among estimates for the hedging premium by the various moderator variables introduced in the meta-regression. In addition, to maximize the comparability of observations in the sample, the regression estimates from the primary studies are transformed to represent the percentage change in firm value due to hedging, i.e., the hedging premium. The hedging premium is comparable within and across studies and helps to ensure that fundamentally different effects are not mixed in the same sample.

Studies Reporting Many Estimates Dominate the Analysis. Because of the imbalance in the meta-data set, the studies that provide many different estimates are given higher weights in the baseline meta-regression, using the inverse of the variance of the effect sizes as the weights. As a robustness analysis, I also weight the meta-regressions by the inverse of the number of estimates reported in each study to avoid undue influence from studies with many estimates. At the same time, however, this approach carries some caveats because it assigns the same weight to each study. This means that, unlike inverse variance weighting, estimates are treated equally regardless of their quality. The third alternative weighting, which uses the interaction between the inverse variance and the inverse number of estimates, takes both aspects into account. Nevertheless, I refer to the results from inverse variance weighting as the baseline results, while the other weighting schemes are used for robustness analysis.

3.9. Summary

The finance literature has long debated whether financial hedging adds value to non-financial firms. Previous empirical studies show mixed effects, ranging from discounts in value to large increases in the value of hedging firms. However, little is known about the actual sources of these variations in research findings. Therefore, I contribute to the literature by statistically accumulating the empirical findings of previous research studies in the field. Using meta-analysis, I aggregate the empirical results from 71 primary studies and identify various sources of heterogeneity in the value effects of hedging. The results of the meta-analysis can be summarized as follows:

(1) The results from the heterogeneity analysis suggest that several aspects of data and method choices explain the large variation in the hedging premiums. I find that better journals (in terms of higher impact factors) show on average 2% lower hedging premiums, suggesting that journal quality is an important determinant of the size of the hedging premium. Moreover, the value effects of hedging largely depend on the type of the risk exposure to that is hedged. IR and CP hedgers are associated with systematically lower firm values than FX hedgers. In terms of estimation methods applied in the primary studies,

controlling for fixed effects and accounting for endogeneity issues is crucial for the size of the detected premium. Omitting these aspects in the primary regression creates an upward bias in the estimated hedging effects. Finally, the results suggest that considering operational hedging as an alternative hedging strategy in addition to financial hedging yields significantly higher premiums than in studies that define hedgers as derivative users only. Thus, the findings in this chapter suggest that operational hedging and financial hedging are complementary strategies.

(2) Using meta-regression, I also test the impact of new country-specific macro variables that measure the conditions under which hedging could be more or less valuable. The findings uncover that the value impact of hedging is smaller in countries with high derivatives and stock trading volumes, OECD member countries, and countries with lower tax rates. No evidence was found that a country's legal environment and governance factors explain differences in hedging premiums. The same applies to the proxies for measuring the cost of financial distress.

(3) The literature on corporate hedging suffers from both an upward bias in the level of reported hedging premiums and an overrepresentation of marginally significant results. These effects produce a biased picture of the true underlying effect. The overall mean hedging premium corrected for publication selection is about four times lower than the simple average of the collected hedging premiums. In this context, I create a synthetic best practice study using the meta-regression results and predict the mean hedging premiums, corrected for publication selection and other errors in the model specification. The hedging premium from a best practice study implies a firm value mark-up for FX hedgers of 1.8%, a negative discount of -0.8% for IR hedgers, and a negative value impact for CP hedgers of -0.6%.

As the majority of the collected estimates for the hedging premium refer to derivatives hedging, the findings of this chapter should rather be interpreted as aggregation and comparison of derivatives premiums as opposed to the broader concept of a hedging premium that also covers other hedging strategies. This highlights the limitations of meta-analysis, as any review can only summarize what has been reported in previous studies. If previous results miss out an important aspect in the definition of hedgers, meta-analysis also bears this restriction. Nevertheless, both the moderator variable capturing the definition of the hedging variable and the control of whether primary studies include variables for operational hedging in their regressions give some indication of the impact of other hedging strategies on firm value.

In a similar direction, I cannot fully rule out the impact of endogeneity on the results. Due to the endogenous nature of hedging policies, inferring causality of the channels through which hedging affects firm value is challenging. Although I explicitly examine whether primary studies account for reverse causality between hedging and firm value in their regression estimation by instrumental variables or other approaches, there is only one study in the sample examining a real natural experiment that makes use of an exogenous change in the basis risk of oil and gas companies in Canada as compared to the U.S. (Gilje and Taillard, 2017). In contrast, all other studies in the sample that correct for endogeneity do not use such a natural experiment but rather a more conventional instrumental variable regression. With more studies evolving in this research field, future meta-research could explicitly examine whether hedging premiums obtained from studies applying natural experiments are systematically different to other studies.

The results from this meta-analysis could guide future empirical research on hedging, as they show the potential sources of bias and reveal the study design characteristics that are important drivers for the empirical outcomes. Future improvements in estimation techniques and data availability could be evaluated against the benchmarks set by the meta-regression results. Moreover, Bayesian approaches relying on objective a-priori distributions could also refer to the meta-analytic findings reflecting the accumulated knowledge of the previous literature. Finally, the results for the country-level determinants point to several compelling avenues for further research. For example, unlike the country-specific variables in this study, which are derived from standard hedging theory about market frictions, other aspects such as a country's risk culture could explain differences in the value premiums of corporate hedging.

Chapter 4. Weighting in Meta-Regression: A Monte Carlo Simulation

"Certainly more research is needed to determine the optimal weighting scheme in meta-analysis with panel data."

(Zigraiova and Havranek, 2016: 973)

This chapter investigates different weighting schemes for WLS meta-regression analysis, which was introduced in Chapter 2 and applied in Chapter 3. Following the introduction (Section 4.1), the previous literature on simulations of meta-analysis methods is reviewed (Section 4.2) and the commonly applied weighting schemes for meta-regression are discussed (Section 4.3). The next section (Section 4.4) presents the design of the Monte Carlo simulation, followed by the presentation of the simulation results (Section 4.5) and a discussion of the key findings (Section 4.6). The chapter ends with a summary (Section 4.7).

4.1. Introduction

Meta-analysis has two key objectives (Stanley et al., 2008: 277). First, it summarizes effect size estimates from a set of primary studies that examine a particular phenomenon, while correcting for potential biases arising from publication selection and model misspecification. Second, meta-analysis can explore sources of heterogeneity and find out why empirical effects differ within and between primary studies. In the Chapters 2 and 3, I described how meta-regression, which is the most common method for meta-analysis in economics and finance, can be applied to estimate an overall mean effect size and explain heterogeneity.

When applying meta-regression analysis, researchers commonly observe heteroscedasticity in the effect size estimates (Stanley and Doucouliagos, 2012: 61). This means that the variance of the error term in the meta-regression is not constant because the variance of the effect size estimates depends on the sample sizes of the primary studies that estimated the effects. Because primary studies are based on different samples, the sample size, and thus the effect size variance, normally varies from study to study. In order to account for heteroscedasticity in the estimation of the meta-regression model while maintaining statistical reliability of the results, *"simple ordinary least squares (OLS) is never the preferred approach for any MRA model, but rather weighted least squares"* (Stanley and Doucouliagos, 2012: 110).

A decisive factor in WLS meta-regression concerns the weights assigned to the effect size estimates.[49] From a statistical point of view, the inverse variance of the errors, which is equal to the inverse variance of the effect size estimates, is the optimal weight in WLS regression with heteroscedasticity in the error term (Wooldridge, 2012: 283). Although inverse variance is optimal from a statistical point of view, applied meta-regressions also use alternative weighting schemes to address practical challenges of inverse variance weighting. As shown in the meta-regression on the hedging-firm value nexus in the Sections 3.7.2 and 3.7.3, different weighting schemes in a meta-regression may influence the estimated mean effects and the coefficients quantifying the impact of the moderator variables, as well as the statistical inferences from hypothesis testing using the meta-regression coefficients.

To illustrate the variety of weighting schemes applied in meta-regression research, Table 21 shows the distribution of the meta-regression weights used in the sample of 76 prior meta-analyses in finance that were reviewed in Chapter 2.

Table 21. Weighting Schemes Applied in Meta-Regression Analyses in Finance

Weighting Scheme	Number of Studies*
Inverse of the effect size variance	39
No weights / Equal weighting	20
Inverse number of effect sizes per study	9
Primary study sample size	7
Study quality	4
Inverse of effect size variance × inverse number of effect sizes per study	2
Citations	1

Notes: This table presents the distribution of the weighting schemes applied in the meta-regressions of 76 previous meta-analysis studies in finance. See Section 2.3.1 for the sample selection. * Multiple counts per study possible.

Three main groups of weighting schemes can be observed in the previous meta-analysis literature in finance: (1) the inverse variance of the effect size estimate or the primary study sample size, (2) the number of effect sizes reported in the primary study or the interaction of the inverse number of estimates and the inverse variance of the effect size estimate, (3) no weighting. Other weights, such as quality rankings[50] and citations, are rarely used. The first group of weights relates to the precision of the effect size estimates. Accordingly, more precise estimates, those with smaller variance and a larger sample size, receive greater weight in the meta-regression. The second group of weights considers the number of estimates reported in each primary study. If multiple effect size estimates are observed from the same study, smaller

[49] See also the previous Section 3.5.2, where the WLS meta-regression was introduced with alternative weights.
[50] The quality ranking reported in Table 21 is used in four studies, all authored by the same researchers.

weights are assigned to each estimate when more estimates are reported in the same study. The third group does not consider weights and therefore estimates meta-regression by OLS. Similar groups of weighting schemes can be observed in other research fields. For example, Nelson and Kennedy (2009: 361) review 140 meta-analyses in environmental and resource economics. They find that among the meta-regressions reported in the 140 meta-analyses, 5 studies use standard errors of the effect size estimates, 11 studies use the variances of the effect size estimates, 13 studies use the primary study sample size, 16 studies use the number of observations collected from the primary study, and 7 studies use other weights.

Given that different weighting schemes are used in meta-analysis, there are several previous (simulation) studies in the literature that examine the optimal weights for meta-analysis in general (among others, Brannick et al., 2011; Kepes et al., 2013; Marin-Martinez and Sanchez-Meca, 2010; Stanley and Doucouliagos, 2015) and meta-regression in particular (among others, Koetse et al., 2010; Reed et al., 2015; Stanley and Doucouliagos, 2017; Zigraiova and Havranek, 2016). Summarizing the key results of those studies, it appears that precision weighting using the inverse of the effect size variance is generally considered the optimal weighting scheme due to its favorable statistical properties (Hedges and Olkin, 1985; Stanley and Doucouliagos, 2017).

One aspect that has not been comprehensively addressed in previous simulation studies concerns cases in which primary studies report not just one but multiple effect size estimates, and the number of estimates varies across studies. However, in economics and finance, we typically find primary studies reporting a varying number of estimates for the effect in question, resulting in an unbalanced meta-data set. However, with multiple estimates per study, inverse variance weighting assigns more weight to studies simply because they report more estimates (among others, Fidrmuc and Lind, 2020: 7; Havranek and Irsova, 2017: 372; Reed et al., 2015: 25; Stanley, 2001: 138). This may cause meta-regression results to be overly influenced by studies that provide many effect size estimates (Stanley and Doucouliagos, 2012: 72).

While WLS with inverse variance weighting is the optimal approach from a statistical point of view when studies report only a single effect size estimate and the effect size variance is heteroscedastic, there may be other weighting schemes with better relative performance in the case of multiple effect size estimates per study. To accommodate the practical challenges of inverse variance weighting, meta-researchers have been applying alternative weighting schemes that consider the number of effect sizes reported in each study (among others, Anderson et al., 2018: 67; Bandaranayake et al., 2020: 7-10; de Batz and Kočenda, 2021: 23; Havranek and Irsova, 2017: 372-392; Iwasaki and Kočenda, 2017: 545). When the inverse number of reported estimates per study is considered in the weighting scheme, each study is assigned the same weight regardless of the number of reported estimates. This reduces the impact of studies reporting many estimates and, hence, the impact of within-study correlation that comes from estimates reported in the same study. However, this approach implicitly assumes that the effect size variances are equal (homoscedasticity), an assumption that is routinely violated in applied meta-regressions. Alternatively, the combined weight of the inverse variances times the inverse number of estimates per study is used to account for the fact that studies may report many effect size estimates and that the variances of the estimates are usually unequal.

Although there are prior simulation studies that compare different weighting schemes and meta-regression estimators (among others, Koetse et al., 2010; Reed et al., 2015; Stanley and Doucouliagos, 2017), there is not yet a comprehensive approach that considers the case of an unbalanced panel data structure whereby each study reports multiple estimates of the examined effect. This is where this chapter aims to contribute by comparing the statistical properties (i.e., bias, mean squared error, and coverage) of inverse variance weighting to the inverse number of reported effects per study when primary studies report varying numbers of non-independent effect size estimates (Zigraiova and Havranek, 2016: 973). To analyze the statistical performance of different weighting schemes, I apply Monte Carlo simulation building on the design of recent and established simulation studies of meta-regression estimators, in particular that of Stanley and Doucouliagos (2017).

4.2. Literature Review of Simulation Studies on Meta-Analysis Methods

There are two related streams of previous simulation studies in the literature. The first strand analyzes weighted mean effect sizes calculated in a simple meta-analysis.[51] For example, Sanchez-Meca and Marin-Martinez (1998) apply Monte Carlo simulation to generate weighted average effects of standardized mean differences. The authors compare the statistical properties of the Hedges and Olkin (1985) estimator with those of the Hunter and Schmidt (1990) estimator. As a key result, they find that the Hedges and Olkin estimator, which uses the inverse variance of the effect sizes as the weight to calculate the weighted mean effect size, is more efficient but also more biased than the Hunter and Schmidt estimator, which uses the sample size of the primary studies as weights. The results are confirmed in a later study by Marin-Martinez and Sanchez-Meca (2010).

In another study, Brannick et al. (2011) use Monte Carlo simulation to compare unit, sample size, and inverse variance weighting when estimating aggregated effect sizes of standardized mean differences and Pearson correlations. The authors conclude that the Hunter and Schmidt estimator is preferable for meta-analyses using Pearson correlations as effect size measure and

[51] See also Section 2.2.1 for an introduction to simple meta-analysis.

that the Hedges and Olkin estimator performs well for meta-analyses using mean differences as effect sizes.

Reed et al. (2015) design a Monte Carlo simulation to evaluate the performance of several meta-analysis estimators in the presence of two types of publication selection bias: (1) selection of statistically significant estimates and (2) rejection of estimates with incorrect signs. Their simulation considers the case in which each study reports one effect size estimate, and all studies have the same true effect (fixed effects). The authors also consider a scenario in which the true effect differs across studies (random effects). In addition, random effects in a panel environment are simulated in which each study reports multiple estimates, and the true effect varies within and between studies (panel random effects). The authors conclude that estimators that do not explicitly correct for publication selection bias (fixed effects, WLS, and random effects) perform well or even better than the estimators that control for publication selection bias (precision effect test and precision effect estimate with standard errors).[52] Moreover, the random effects estimator is often more biased but also more efficient than other estimators. However, in a reader's comment, Stanley (2015) points out several weaknesses in the simulation design of Reed et al. (2015), particularly their sampling approach, which biases the reported results. The results of the simulation study should therefore be interpreted with caution.

Other studies in this first literature stream analyze further aspects of weighting in simple meta-analysis while considering different effect size measures, alternative meta-analysis estimators, and characteristics of specific research disciplines (among others, Park and Beretvas, 2018; Stanley and Doucouliagos, 2015).

The second stream in the literature analyzes weighting schemes in meta-regression analyses where the effect sizes are regression estimates and the model accounts for the sources of heterogeneity within and across studies as well as the effects of publication selection. For example, Koetse et al. (2010) apply Monte Carlo simulation to investigate the behavior of unweighted OLS meta-regression, WLS meta-regression with inverse variance weighting, and mixed effects meta-regression in the presence of heteroscedasticity as well as heterogeneity in terms of random effect size variation. The mixed-effects model splits the variance component into the effect size variance and the between-study variance, the latter of which is unknown and estimated by the model. In their simulation, the authors examine the statistical properties of the meta-estimators for increasing levels of heteroscedasticity in the primary study error variance and increasing levels of the between-study variance. They find that as heteroscedasticity rises, the OLS variance decreases substantially as compared to WLS and the

[52] For an overview of the common meta-analysis estimators, see Sections 2.2.1 and 2.2.2.

mixed effects model. This result changes when random variation of the true underlying effect is introduced. Random variation increases the variance of all three estimators and in particular the variance of the WLS estimator. In this case, the mixed effects estimator is preferable. However, when random variation is introduced only in a subset of the effect sizes and not in the overall population, the results are different because the bias of the OLS and mixed effects estimators increases when the proportion of estimates from misspecified models is large. In contrast, the WLS estimator remains unaffected. In summary, the authors conclude that *"using the mixed effects estimator in empirical applications of meta-analysis is suboptimal under these circumstances, and that applying WLS is clearly preferable"* (Koetse et al., 2010: 235).

Zigraiova and Havranek (2016: 971-974) discuss the challenges of inverse variance weighting in practical applications and propose the inverse of the number of observed effect size estimates per study as alternative weight. The authors see five challenges that arise from inverse variance weighting. First, previous Monte Carlo simulations of meta-analysis methods that show that inverse variance weighting leads to optimal results do not account for cases in which the number of effect size estimates observed from each study is different. Second, inverse variance weighting introduces artificial variation in moderator variables that are defined at the study level, such as the publication year of the study. Although such factors are constant within a study, inverse variance weighting adds variation to these fixed effects variables and it is not clear which weighting scheme is preferable in this scenario. Third, effect size estimates and their standard errors could be endogenous if methodological decisions in the primary study affect both the effect size estimate and its standard error. If endogeneity exists and a particular methodological choice affects both effect size estimates and their standard errors in the same direction, any meta-regression specification that includes the standard error as an explanatory variable will be biased. Fourth, inverse variance weights are sensitive to outliers in the effect size variance.[53] When primary studies report small standard errors of their effects, this is usually taken as evidence of precision that improves the estimate of the overall mean effect. However, when highly precise estimates are biased by incorrect model specification or the application of some idiosyncratic methods, they can have a significant impact on the meta-regression results because inverse variance weighting assigns larger weights to these observations. Fifth, the alternative weighting scheme based on the number of observed effect size estimates per study allocates equal importance to each study, rather than assigning greater weight to a study that reports more estimates, as with inverse variance weighting.

[53] It should be noted that there are several methods for detecting and correcting outliers, such as winsorizing or DFBETAS (Bollen and Jackman, 1985: 518-519).

Stanley and Doucouliagos (2017) extend the simulation study by Koetse et al. (2010). They apply Monte Carlo simulation to compare the statistical performance of the (unrestricted) WLS estimator to the mixed effects/random effects estimator while considering heterogeneity and publication selection bias. The authors consider two sources of heterogeneity in their simulation design: systematic heterogeneity caused by omitted-variable bias and random unexplained heterogeneity. Random heterogeneity is either modeled as omitted-variable bias, as direct additive heterogeneity, or random deviations of a moderator variable. As a result, they find that the unrestricted WLS estimator yields statistical properties (bias, mean squared error, and coverage) similar to those of the mixed effects model when there is no publication selection bias. However, when publication selection bias is present, the WLS estimator shows less bias than the mixed effects model. Since in practical applications it is difficult to rule out publication selection bias, due to the low power of publication selection bias tests, Stanley and Doucouliagos (2017) recommend WLS as the preferred meta-regression estimator.

Bom and Rachinger (2020) extend previous meta-regression simulations, especially that of Stanley and Doucouliagos (2017), by explicitly modeling overlapping samples. Sample overlaps are often observed in research fields that use aggregated observational data or in which the same data is used by several primary studies. Sample overlap also occurs when multiple effect size estimates are sampled from the same primary study. The authors show that overlapping samples causes high rates of false positives when using the fixed effects, WLS, or random effects estimators. Therefore, they introduce a new Generalized Weighting (GW) estimator that explicitly considers the dependency structure in the variance-covariance matrix that arises from overlapping samples. The degree of sample overlap is defined by the country and year of observation in the primary data. The GW estimator assigns weights to each primary study estimate based on its *"share of independent sampling information"* (Bom and Rachinger, 2020: 830), which ensures consistency and maximizes efficiency in the case of overlapping sample data.

Summarizing the previous literature, inverse variance weighting and sample size weighting are the most intensively analyzed weighting schemes for simple meta-analysis (Brannick et al., 2011; Marin-Martinez and Sanchez-Meca, 2010; Sanchez-Meca and Marin-Martinez, 1998). For meta-regression, WLS using inverse variance weighting performs better under real conditions, including publication selection bias and heterogeneity, than fixed or random/mixed effects meta-regression models (Koetse et al., 2010; Stanley and Doucouliagos, 2017). Several studies consider additional practical conditions in their simulation approaches, including various sources of publication selection bias, systematic or random heterogeneity, or the fact that studies often report non-independent effect size estimates (Bom and Rachinger, 2020; Reed et al., 2015; Stanley and Doucouliagos, 2017).

Despite the variety of previous simulation studies on meta-analysis weights, there are aspects that are not or not comprehensively considered in the previous literature. While Reed et al. (2015) and Bom and Rachinger (2020) account for cases in which studies report multiple estimates for the effect in question, the results do not consider instances in which studies report a varying number of effect sizes. Moreover, following the discussions by Zigraiova and Havranek (2016: 971-974), who recommend the inverse number of estimates per study as the preferred weighting scheme in meta-regression, there is not yet a simulation study comparing the statistical properties of inverse variance weighting with the inverse number of reported estimates per study used as alternative weights.

This is where this chapter contributes to the literature. It extends the previous simulation studies by comparing the statistical properties of three alternative weighting schemes: inverse variance, the inverse number of estimates per study, as well as the interaction of the inverse variance with the inverse number of effect size estimates. The three weights are compared in a scenario in which the number of effect sizes varies across studies and the observed estimates from the same study are not independent.

4.3. Weighting Schemes in Meta-Regression

This section describes the WLS meta-regression and the three weighting schemes for which the statistical performance is compared in the subsequent simulation analysis.

4.3.1. Weighted Least Squares Meta-Regression

In meta-regression analysis, excess systematic variation in a sample of effect size estimates is explained by a set of moderator variables:

$$\boldsymbol{y} = \boldsymbol{M\beta} + \mathbf{e}, \text{ with } \mathbf{e} \sim N(0, \boldsymbol{V}), \tag{17}$$

where $\boldsymbol{y}$ is an $L \times 1$ vector of all reported effect size estimates collected from a set of primary studies in which each primary study reports one effect size estimate, $\boldsymbol{M}$ is an $L \times P$ matrix including the constant and the meta-explanatory variables capturing heterogeneity as well as (a)symmetry of the effect size estimates[54], $\boldsymbol{\beta}$ is a $P \times 1$ vector of meta-regression coefficients, $\boldsymbol{e}$ represents an $L \times 1$ vector capturing the estimation errors of the reported effect size estimates, $\boldsymbol{V}$ is the variance-covariance matrix of $\boldsymbol{y}$, and $E(\boldsymbol{\varepsilon\varepsilon}^t) = \boldsymbol{V}$.

[54] The common meta-explanatory variables are $SE(y_{ij})$ to control for publication selection, as well as the vectors $\boldsymbol{M}$ and $\boldsymbol{M}'$ of the moderator variables to control for model misspecification and various sources of systematic heterogeneity (see Eq. 5 in Section 2.2.2).

When the Gauss-Markov assumptions are valid, the meta-regression coefficients, $\boldsymbol{\beta}$, can be estimated with OLS (Greene, 2011: 16):

1. Linearity: The model defines a linear relationship between the dependent variable and the independent variable(s).

2. Full rank: There is no exact linear relation between the independent variables.

3. Exogeneity: The expected value of the random error term ($\boldsymbol{e}$) is not a function of the independent variables.

4. Homoscedasticity: The errors have the same finite variance.

5. Non-autocorrelation: The errors have no serial dependence.

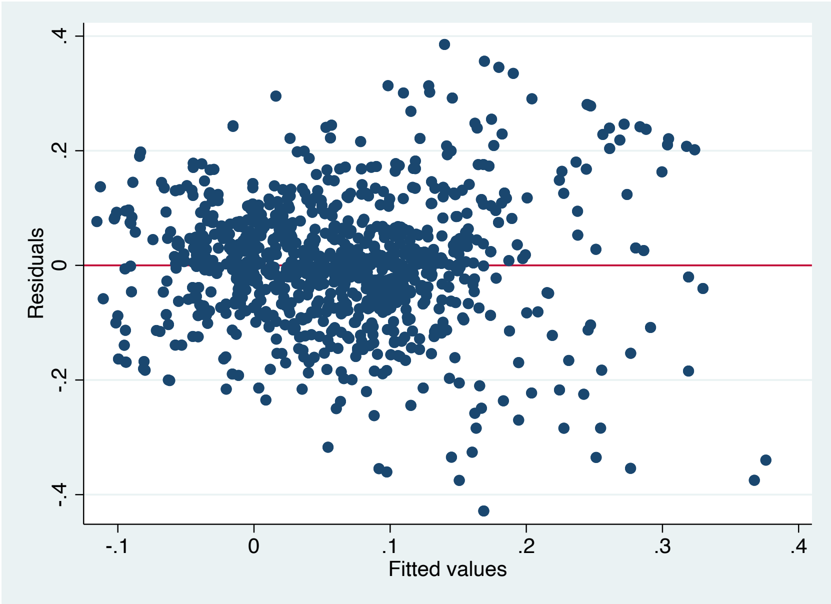
The errors that meet assumptions (4) and (5) are also referred to as 'spherical' errors (Greene, 2011: 21-22). According to the Gauss-Markov theorem, OLS is the best linear unbiased estimator (BLUE) under these assumptions (Wooldridge, 2012: 102). When OLS is BLUE, the expected values of the estimated regression coefficients equal the values describing the relation between $\boldsymbol{y}$ and $\boldsymbol{M}$ (unbiased estimator), and the variance of the OLS estimator is smaller than the variance of any other linear and unbiased estimator (best estimator) (Greene, 2011: 60). The validity of the Gauss-Markov assumptions should always be tested before applying OLS.

In almost all meta-regression applications, the primary studies are based on different data sets with varying sample sizes and alternating model specifications, causing heteroscedasticity of the errors. Accordingly, assumption (4), which is the homoscedasticity of the random error term is routinely violated in meta-regressions (Stanley and Doucouliagos, 2012: 110). In the case of homoscedasticity, the diagonal elements σ^2 in the variance-covariance matrix are identical variances and the off-diagonal elements represent the covariances, which are all zero as, according to assumption (5), the errors must be independent (Greene, 2011: 22). Hence, the variance-covariance matrix $\boldsymbol{V}$ of the effect size estimates $\boldsymbol{y}$ is given by:

$$\boldsymbol{V} = \begin{bmatrix} \sigma^2 & 0 & . & . & 0 \\ 0 & \sigma^2 & & & 0 \\ . & & . & & . \\ . & & & . & . \\ 0 & 0 & . & . & \sigma^2 \end{bmatrix}. \tag{18}$$

To illustrate the heteroscedasticity that is common in meta-regression analysis, Figure 30 plots the residuals against the fitted values from a simple OLS meta-regression of the estimates for the hedging premium and the moderator variables as defined in Eq. (15) in Section 3.5.1.

Figure 30. Graphical Illustration of Heteroscedasticity in Meta-Regression



Notes: This graph shows a plot of residuals and fitted values from a simple OLS meta-regression of the primary study estimates for the hedging premium and a set of moderator variables as defined in Eq. (15) in Section 3.5.1.

The scatter plot shows a pattern with larger residuals towards the right end of the graph, which indicates heteroscedasticity. As a result of heteroscedasticity, the variance-covariance matrix of $\boldsymbol{y}$ will not have the constant structure as in Eq. (18). Since $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^t) = \boldsymbol{V}$, the variances of the errors depend on the primary study from which the effect size estimates are obtained. Accordingly, the diagonal elements in the variance-covariance matrix are not identical because each of the $i = 1, \ldots, L$ primary studies has its own study-specific variance:

$$\boldsymbol{V} = \begin{bmatrix} \sigma_1^2 & 0 & . & . & 0 \\ 0 & \sigma_2^2 & & & 0 \\ . & & . & & . \\ . & & & . & . \\ 0 & 0 & . & . & \sigma_L^2 \end{bmatrix}. \tag{19}$$

Since there is still the assumption that the errors are independent, the covariances are again zero. When the random error term in the meta-regression is heteroscedastic, the OLS estimator is no longer the BLUE. OLS is still unbiased, but no longer efficient, and inefficiency increases when the dispersion of the individual variances grows (Greene, 2011: 270). As a result of inefficiency, erroneous inferences may be derived from hypothesis testing using the incorrect standard errors of the OLS estimator (White, 1980: 817).

As OLS is not efficient under heteroscedasticity, there is another estimator with a smaller variance. Accordingly, the routine procedure to account for heteroscedasticity is to estimate the meta-regression model using WLS, which compensates for non-constant error variances by weighting each observation according to its precision.[55] WLS is a special case of generalized least squares (GLS) estimators, where the variance-covariance matrix $\boldsymbol{V}$ has the diagonal structure as in Eq. (19). Aitken (1935) generalized the Gauss-Markov theorem and shows that GLS, and hence WLS, are minimum variance within the class of BLUEs when the parameters of the variance-covariance matrix $\boldsymbol{V}$ are known.[56]

While OLS minimizes the sum of the squared residuals, WLS minimizes the 'weighted' sum of squared residuals. When the correct weights are known a priori, each residual is weighted by $1/w_i$ (Wooldridge, 2012: 283):

$$(\boldsymbol{y}-\boldsymbol{M\beta})^{\boldsymbol{T}}\boldsymbol{W^{-1}}(\boldsymbol{y}-\boldsymbol{M\beta}), \tag{20}$$

where $\boldsymbol{W^{-1}}$ is a diagonal weight matrix:

$$\boldsymbol{W^{-1}} = \begin{bmatrix} 1/w_1 & 0 & . & . & 0 \\ 0 & 1/w_2 & & & 0 \\ . & & . & & . \\ . & & & . & . \\ 0 & 0 & . & . & 1/w_L \end{bmatrix}. \tag{21}$$

Assuming that one effect size estimate is observed in each study, the number of primary studies, L, is equal to the number of effect size estimates, K. The WLS estimators are the values of $\hat{\boldsymbol{\beta}}$ that minimize Eq. (20).[57] The general WLS estimator is given by (Greene, 2011: 279):

$$\hat{\boldsymbol{\beta}}_{\text{WLS}} = (\boldsymbol{M}^T\boldsymbol{W^{-1}M})^{\boldsymbol{-1}}\ \boldsymbol{M}^T\boldsymbol{W^{-1}y}. \tag{22}$$

When applying WLS, each effect size estimate in the regression has a specific weight, w_i. While OLS is the special case in which each effect size estimate is weighted equally, efficient WLS weighting assigns more weight, and thus more importance, to estimates that provide more information about the relation between the dependent and the independent variables (Mosteller and Tukey, 1977: 346). Through the weighting, WLS produces residuals with a constant variance (homoscedasticity). Applying incorrect weights in a WLS regression causes the estimator to be inefficient, i.e., the estimator does not achieve the smallest variance, and the standard errors will be incorrect (Greene, 2011: 278).

[55] There are also other estimators used in meta-analysis, such as the mixed effects or random effects meta-regression. However, previous research shows that the (unrestricted) WLS is equivalent or even superior to these estimators (Stanley and Doucouliagos, 2017). Thus, the other estimators will not be considered here.
[56] If the weights are unknown, the model can also be estimated by OLS using the White (1980) estimator or another estimator for an asymptotic covariance matrix.
[57] It can be shown that WLS is equivalent to dividing the regression by $\sqrt{w_i}$ and running OLS on the adjusted regression model (Chatterjee and Hadi, 2006: 166).

4.3.2. Inverse Variance Weighting

According to Wooldridge (2012: 283), *"a weighted least squares estimator can be defined for any set of positive weights"*. Among the multitude of possible weights, the most efficient procedure to overcome heteroscedasticity is to weight the squared residuals in Eq. (20) by the inverse of the error variance (Wooldridge, 2012: 283). This is the BLUE in the presence of heteroscedasticity. However, in most empirical studies, WLS comes with the challenge that the error variances are unknown. But as long as the error variance can be predicted from another variable and the estimates of σ_i^2 are consistent, the WLS estimates are also consistent and asymptotically efficient (Wooldridge, 2002: 160-162).

Meta-analysts have the advantage that they can observe the effect size estimates, y_i, along with an estimate of their variance, $\hat{\sigma}_i^2 = SE(y_i)^2$ (Stanley and Doucouliagos, 2012: 111). Accordingly, the primary studies usually provide a direct estimate of the variance. If these observed estimates of the effect size variances are consistent, the weights matrix $\boldsymbol{W^{-1}}$ for the WLS estimator is given by:

$$\boldsymbol{W^{-1}} = \begin{bmatrix} 1/\hat{\sigma}_1^2 & 0 & . & . & 0 \\ 0 & 1/\hat{\sigma}_2^2 & & & 0 \\ . & & . & & . \\ . & & & . & . \\ 0 & 0 & . & . & 1/\hat{\sigma}_L^2 \end{bmatrix}. \tag{23}$$

Under this weighting scheme, each effect size estimate is given a weight based on its inverse variance. By applying inverse variance weighting, the WLS estimator considers the precision of the information that is incorporated into each effect size estimate. The more precise effect size estimates (that is, those with small effect size variance) are given greater weight in the estimation of the meta-regression coefficients, while less precise effect size estimates (that is, those with larger effect size variance) receive a smaller weight.

4.3.3. Unbalanced Panel Data

Although from a statistical point of view, WLS with inverse variance weighting is the BLUE in the case of heteroscedasticity, reality is often more complex. One of these factors is that primary studies typically report more than one effect size estimate per study to consider different primary model specifications and subsamples of the data set. It is likely that such multiple estimates within the same study are interdependent because they are obtained from the same or a similar data set, or that multiple estimates could be influenced by an unreported or unobservable common factor such as the quality of the study or the ideology of the authors (Stanley and Doucouliagos, 2012: 113). When effect size estimates are not independent, the

Gauss-Markov assumption (5) is violated, which requires that the errors in the variance-covariance matrix $\boldsymbol{V}$ be uncorrelated.

If we assume a meta-data set where the first study ($i = 1$) and the last study ($i = L$) report three estimates each, with a non-zero within-study correlation and a zero between-study correlation for all estimates, the variance-covariance matrix has non-zero values on the off-diagonal elements:

$$\boldsymbol{V} = \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,12}^2 & \sigma_{1,13}^2 & . & . & 0 & 0 & 0 \\ \sigma_{1,21}^2 & \sigma_{1,2}^2 & \sigma_{1,23}^2 & . & . & 0 & 0 & 0 \\ \sigma_{1,31}^2 & \sigma_{1,32}^2 & \sigma_{1,3}^2 & . & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & \sigma_{L,1}^2 & \sigma_{L,12}^2 & \sigma_{L,13}^2 \\ 0 & 0 & 0 & . & . & \sigma_{L,21}^2 & \sigma_{L,2}^2 & \sigma_{L,23}^2 \\ 0 & 0 & 0 & . & . & \sigma_{L,31}^2 & \sigma_{L,32}^2 & \sigma_{L,3}^2 \end{bmatrix}, \tag{24}$$

where $\sigma_{1,1}^2$ is the variance of the first effect size estimate reported in study $i = 1$ and $\sigma_{1,12}^2$ is the covariance between the first and the second estimates observed in the first study of the sample.

When the errors in the meta-regression are correlated, OLS is still unbiased, but no longer provides the minimum variance. As a consequence, the OLS standard errors are incorrect because they assume uncorrelated errors (Wooldridge, 2012: 414). Similar to the case of heteroscedasticity, erroneous inferences might be derived from hypothesis tests using incorrect standard errors. To ensure the validity of the meta-regression estimates, the dependency structure of the effect size estimates must be considered in the estimation of the meta-regression model. In the case of correlated errors, GLS is the efficient estimator, but requires that the covariances of the errors be known (Greene, 2011: 924). As an alternative to GLS, consistent estimates of the standard errors can be obtained using, for example, the Newey-West approach (Newey and West, 1987). The standard routine in meta-regression research to account for heteroscedasticity and correlation of the errors is to use WLS with cluster-robust standard errors or WLS panel regression methods (Havranek et al., 2020: 472).[58]

Weighting by the Inverse Number of Estimates per Study. In applied meta-regression, the number of reported effect size estimates is often not the same for all studies but rather varies from study to study. For example, in Chapter 3, the number of effect size estimates for the hedging and firm value relation ranges from 1 to 82, while the average is 14 estimates reported per study. If all effect size estimates were to have the same estimated standard errors, the

[58] Meta-analysts also apply bootstrapped standard errors and multilevel regression models.

conventional WLS meta-regression approach with inverse variance weighting would assign the same weight to all estimates. Accordingly, a study that reports 82 estimates would implicitly receive 82 times the weight of the study that reports only one estimate. This causes an 'over-representativeness' of studies that report many estimates. Although the number of estimates is not indicative of a 'better' study, the results of a meta-regression using inverse variance weighting are strongly influenced by studies with many estimates, while the impact of studies with a small number of estimates may be marginal (Zigraiova and Havranek, 2016: 971-973). The effect of unintentional weighting is greater the more unbalanced the meta-data set is.

Meta-analysts use different approaches to mitigate the challenge of unintentional weighting, including the selection of a representative effect size estimate or taking the average effect size estimate per study as the dependent variable in the meta-regression (Stanley and Doucouliagos, 2012: 32). However, every selection approach poses the challenge of finding objective criteria for deciding on what the 'best' estimate is, and it often reduces the variability of the effect size estimates in the meta-regression (Stanley and Doucouliagos, 2012: 32). An alternative approach is to retain all observed effect size estimates in the data set, but to consider the number of estimates observed per study, m_i, in the weights of the meta-regression.

An exemplary weighting matrix for a sample of three studies reporting $m_1 = 1$, $m_2 = 2$, and $m_3 = 3$ effect size estimates is given by:

$$\boldsymbol{W^{-1}} = \begin{bmatrix} 1/m_1 & 0 & . & . & . & 0 \\ 0 & 1/m_2 & & & & . \\ . & & 1/m_2 & & & . \\ . & & & 1/m_3 & & . \\ . & & & & 1/m_3 & . \\ 0 & . & . & . & . & 1/m_3 \end{bmatrix}, \tag{25}$$

where m_i is the number of estimates observed per study. Accordingly, the weights are $1/m_1 = 1$, $1/m_2 = 0.5$, and $1/m_3 = 0.\overline{33}$.

Weighting the meta-regression model by the inverse of the number of estimates per study instead of the inverse of the effect size variance assigns equal importance to all studies, regardless of the number of effect size estimates observed in each study. Recent examples applying this weighting scheme in meta-regression analysis are Anderson et al. (2018: 67), de Batz and Kočenda (2021: 23), Iwasaki and Kočenda (2017: 545), Lichter et al. (2015: 103-104), Reed and Sidek (2015: 5-6), and Zigraiova and Havranek (2016: 950-964).

Although the WLS estimator is consistent independently of the weights used, efficiency requires that the weights be proportional to the inverse of the variance (Greene, 2011: 278). Because the number of effect size estimates per study is not proportional to the inverse of the effect size variance, WLS regression using the inverse number of the estimates per study as weights is inefficient. As an alternative to WLS, another common approach to addressing non-constant variance in regression models is the White heteroscedasticity-consistent estimator (White, 1980). White (1980) proposes using the inefficient (but consistent) OLS estimator with a heteroscedasticity-consistent covariance matrix. The resulting standard errors are heteroscedasticity robust. This approach is often used in empirical economics, especially when the exact functional form of heteroscedasticity is unknown (Greene, 2011: 273).[59]

Weighting by the Inverse Number of Estimates and the Inverse Variance. To account for both heteroscedasticity and the unintentional weighting of studies with many estimates, the two meta-regression weights outlined above can be combined using the inverse variance times the inverse number of estimates per study as weights. An exemplary weights matrix for a sample of two studies with $m_1 = 2$ and $m_2 = 3$ is given by:

$$\boldsymbol{W^{-1}} = \begin{bmatrix} 1/(\sigma_{1,1}^2\ m_1) & 0 & . & . & 0 \\ 0 & 1/(\sigma_{1,2}^2\ m_1) & & & . \\ . & & 1/(\sigma_{2,1}^2\ m_2) & & . \\ . & & & 1/(\sigma_{2,2}^2\ m_2) & . \\ 0 & . & . & . & 1/(\sigma_{2,3}^2\ m_2) \end{bmatrix}. \tag{26}$$

For example, if $\sigma_{1,1}^2 = 0.1$ and $\sigma_{2,1}^2 = 0.05$, the corresponding weights are $w_{1,1} = \frac{1}{0.1} * \frac{1}{2} = 5$ and $w_{2,1} = \frac{1}{0.05} * \frac{1}{3} = 6.\overline{66}$.

When applying this weighting scheme, studies with a different number of estimates are weighted equally, but within each equally weighted study, the estimates with lower variance receive a higher weight. Recent examples applying these weights in meta-regression are Bandaranayake et al. (2020: 7-10), Geyer-Klingeberg et al. (2018b: 2181-2185), and Kočenda and Iwasaki (2022: 23-26).

In summary, inverse variance weighting is the superior method for WLS meta-regression in the case of heteroscedasticity. Although inverse variance weighting is dominant in theory, its statistical advantages may not always manifest themselves in practical applications. If studies report an unbalanced number of effect size estimates, inverse variance weighting could impose excessive weights on certain studies. Which weights are optimal for practical applications of meta-regression when studies report a varying number of correlated effect size estimates? This is the research question to be addressed in this chapter. The next step is to design a Monte

[59] In small samples, the White (1980) estimator is usually downward biased, which may impact the inferences of the regression model (Chesher and Jewitt, 1987; Furno, 1996). But it is a consistent estimator when the errors are heteroscedastic and the bias decreases with increasing sample size (Hayes and Cai, 2007: 712).

Carlo simulation for meta-regression that allows for the evaluation of the statistical properties of the three outlined weighting schemes under practical conditions of finance and economics.

4.4. Simulation Design

This section describes the design of the Monte Carlo simulation and the statistical properties used to compare the three weighting schemes discussed in the previous section.

4.4.1. Objectives of the Simulation

This simulation builds on prior simulation studies of meta-regression methods (especially Bom and Rachinger, 2019; Bom and Rachinger, 2020; Koetse et al., 2010; Reed et al., 2015; Stanley, 2008, 2017; Stanley and Doucouliagos, 2015, 2017). I extend these studies by introducing an unbalanced panel data structure in which the simulated primary studies report multiple effect size estimates, and the number of reported estimates varies between studies. The statistical properties of the three weighting schemes are evaluated at varying levels of unbalancedness of the panel data and varying degrees of sample overlap of the estimates reported in the same study. To the best of my knowledge, this is the first simulation study examining the three meta-regression weights (inverse of the effect size variance, inverse number of estimates per study, and inverse of the effect size variance times the inverse number of estimates per study) with unbalanced panel data at the meta-level.

The general idea of the Monte Carlo simulation is to generate a representative random sample of meta-estimates with the desired statistical properties and to evaluate the performance of the weighting schemes for a large number of replications of the simulation (Kočenda and Cerný, 2015: 193). The design of this simulation is closely following Stanley and Doucouliagos (2017).[60] Their simulation process generates a sample of primary studies that test a specific regression coefficient, α_1.[61] α_1 represents the empirical effect to be analyzed in meta-regression. The simulation in this chapter extends Stanley and Doucouliagos (2017) in the sense that in the primary data generation, each study can produce one or multiple non-independent effect size estimates, creating a panel data structure. Another aspect to be considered in the simulation is that meta-analyses in economics and finance typically find large heterogeneity (Stanley and Doucouliagos, 2017: 28), i.e., variation beyond sampling error. While some heterogeneity can usually be explained by the moderators included in the meta-regression

[60] The approach by Stanley and Doucouliagos (2017) differs from other Monte Carlo simulations of meta-analysis methods, such as Bijmolt and Pieters (2001) or Sanchez-Meca and Marin-Martinez (1998), by explicitly modeling and controlling the process of data generation in the primary studies before estimating the meta-analysis results.
[61] As regression analyses are the *"workhorse statistical models for empirical analysis in finance"* (Sollis, 2012: 49), no other statistical tests are simulated in this chapter.

model, a substantial amount of heterogeneity often remains unexplained. Previous simulations have shown that the magnitude of unexplained heterogeneity is a decisive factor of the statistical properties of meta-regression estimators (Bom and Rachinger, 2019; Moreno et al., 2009; Stanley, 2008, 2017; Stanley and Doucouliagos, 2015, 2017). Following Stanley and Doucouliagos (2017), two sources of heterogeneity are considered in this simulation: systematic heterogeneity and random unexplained heterogeneity.

The baseline simulation consists of five steps that are designed to reflect real meta-regression applications:

1. Generating random primary study data
2. Estimating the primary regression model
3. Constructing a meta-sample of primary regression estimates and their standard errors
4. Computing meta-regression estimates with different weighting schemes
5. Replicating the previous steps 10,000 times to evaluate the statistical properties of the three weighting schemes

Steps 1–3 and 4–5 of the simulation process are illustrated in Figure 31 and Figure 34, respectively.

4.4.2. Primary Study Simulation

Data Generating Process. Random variables are generated to simulate the data input for the primary study regression. There are three independent variables ($X_{1,h}$, $X_{2,h}$, and $X_{3,h}$) and one dependent variable (z_h). The variable of interest is $X_{1,h}$, which is a random uniform variable, $X_{1,h}$~$U(100, 200)$, where $h = 1, \dots, N_{ij}$ and N_{ij} is the sample size of primary study i that is used for the estimation of the jth effect. The other independent variables, $X_{2,h}$ and $X_{3,h}$, are defined in such a way that they are correlated with $X_{1,h}$:

$$X_{2,h} = X_{1,h} + N(0, 50^2),\ X_{3,h} = X_{1,h} + N(0, 50^2), \tag{27}$$

where $N(0, 50^2)$ is a random disturbance term. The dependent variable, z_h, is defined as follows:

$$z_h = 100 + \alpha_1 X_{1,h} + \alpha_2 X_{2,h} + \alpha_3 X_{3,h} + 100u_h, \tag{28}$$

where u_h~$N(0, 1^2)$, $\alpha_1 = \{0, 1\}$, $\alpha_2 = 0.5$, and α_3~$N(0, \sigma_g^2)$. The regression coefficient of interest is α_1, which is a deterministic variable that can take two possible values {0,1}.

Figure 31. Process for the Simulation of Primary Studies

1. Data Generating Process

$$X_{1,h} \sim U(100,200) \qquad X_{2,h} = X_{1,h} + N(0,50^2) \qquad X_{3,h} = X_{1,h} + N(0,50^2)$$

$$z_h = 100 + \alpha_1 X_{1,h} + \alpha_2 X_{2,h} + \alpha_3 X_{3,h} + 100u_h$$

$$\alpha_1 = \{0,1\} \qquad \alpha_2 = 0.5 \qquad \alpha_3 \sim N(0,\sigma_g^2) \qquad u_h \sim N(0,1^2)$$

2. Sample Data Set

For all random variables, draw N_{ij} observations, where N_{ij} is the sample size of study i and estimate j, which is uniformly distributed $= \{62, 125, 250, 500, 1000\}$.

3. Primary Study Regression

$$[1]\ z_h = \alpha_0 + \alpha_1 X_{1,h} + \alpha_2 X_{2,h} + u'_h, \text{ or}$$

$$[2]\ z_h = \alpha_0 + \alpha_1 X_{1,h} + u''_h$$

Half of the effect size estimates come from regression [1], while the other half of the estimates come from a regression [2] that omits $X_{2,h}$.

4. Multiple Estimates per Study

Each study reports m_{i} estimates of $\hat{\alpha}_1$, where m_{i} varies between studies depending on the level of the Gini coefficient, $G = \{0, 0.33, 0.55, 0.67\}$.

The samples N_{ij} in the same study i are overlapping by taking $C_i = \min_{\forall j=1,\dots,K} \{N_{ij}\}$ observations that are the same for all estimates in a study i, and $N_{ij} - C_i$ observations that are randomly generated for each estimate.

Random heterogeneity, α_3 , is either effect size-specific and varies within a study or, alternatively, random heterogeneity is constant within a study.

5. Publication Selection Bias

If $\hat{\alpha}_1$ is not statistically significant at the 5% level, $(f * 100)\%$ of the researchers draw a new sample of the random variables with $N_{ij} - C_i$ observations and re-estimate step 3 until reaching statistical significance.

6. Primary Study Output

Steps 1 to 5 are repeated for $i = 1, \dots, 20$ studies with $j = 1, \dots, m_i$ effect size estimates per study and $\sum_{i=1}^{20} m_i =$ 80. This produces a sample of 80 effect size estimates, $y_{ij} = \hat{\alpha}_{1,ij}$, and corresponding standard errors, $SE(y_{ij})$.

Notes: This figure describes the steps for the simulation of the primary studies.

Sample Data Set. As routinely observed in empirical primary research, a wide range of sample sizes is used to estimate the empirical effect, α_1. Stanley and Doucouliagos (2017) simulate either 20 or 80 primary studies with one effect size estimate from each study. Accordingly, their meta-regression sample consists of either 20 or 80 effect size estimates. I use the larger figure of 80 as the number of effect size estimates and distribute the 80 estimates across 20 studies to create an unbalanced panel data structure.[62] Each primary study i has an estimate-specific sample size, N_{ij}, that defines the number of draws for the random set of primary data used for the estimation of the effect size j. The sample size is uniformly distributed among possible sample sizes $= \{62, 125, 250, 500, 1000\}$. Accordingly, a meta-regression with 80 effect size estimates contains 16 estimates for each primary sample size. The sample size for the estimation of an effect size is randomly selected from this distribution. Accordingly, a study that reports multiple effect size estimates may have estimates based on different sample sizes that are drawn randomly from this distribution.[63]

Primary Study Regression. The researcher of study i uses the primary data with the respective sample size N_{ij} and estimates one of two OLS regressions:

$$z_h = \alpha_0 + \alpha_1 X_{1,h} + \alpha_2 X_{2,h} + u'_h, \tag{29a}$$

$$z_h = \alpha_0 + \alpha_1 X_{1,h} + u''_h. \tag{29b}$$

The result of the regression is the estimated effect size, $\hat{\alpha}_1$, and its standard error, $SE(\hat{\alpha}_1)$.

Systematic heterogeneity is introduced by omitted-variable bias. Omitting a relevant variable like X_2, that is correlated with X_1, causes the effect size estimate, $\hat{\alpha}_1$, to be biased. X_2 is randomly omitted in half of the regressions (Eq. 29b). When a study generates multiple effect size estimates, the researcher randomly decides for each regression whether to omit X_2. Across all estimates in the meta-sample, half of the regressions are estimated without X_2.[64]

In addition to systematic heterogeneity, random heterogeneity is imposed by a second omitted variable, X_3. In economics and finance research, as in any other social science, a certain phenomenon might be impacted by a large number of variables and factors that cannot all be

[62] According to Stanley and Doucouliagos (2017: 24), the sample size of 80 effect size estimates in the meta-regression *"is chosen because it provides sufficient power in most meta-regression applications"*.
[63] In an extension of the simulation design, a scenario could be simulated whereby the sample sizes used for multiple effect size estimates may have certain clusters of similar sample sizes used in the same study or in which the sample size depends on the number of effect size estimates reported per study.
[64] In an extension of the simulation design, a scenario could be simulated whereby omitted-variable bias is not randomly distributed among the 80 effect size estimates, but clustered in specific studies, e.g., the studies with lower sample sizes.

controlled for in a primary study (Stanley and Doucouliagos, 2017: 27). X_3 accounts for this by adding random omitted-variable bias. Following Stanley and Doucouliagos (2017: 27), such random omitted-variable bias is the main source of unexplained heterogeneity and selection bias in applied economics and finance research. Unlike X_2, X_3 is omitted in all primary study regressions. Two scenarios are considered in which α_3 either varies between studies but is fixed within a study (random heterogeneity at the study level, see Tables 26–30) or, alternatively, α_3 varies also between the estimates within the same study (random heterogeneity at the effect size level, see Tables 23 and 25). Due to random heterogeneity, the mean of the sampling distribution, $E_{\alpha_3}(\hat{\alpha}_1)$, is not α_1, but $\alpha_1 + \alpha_3$, where α_1 measures the direct effect of X_1 on z, and α_3 captures the indirect effect of X_1 on z through X_3. Thus, there is no single true effect, but rather a distribution of true effects that are normally distributed around their mean, which is $\alpha_1 + \alpha_3$. If random heterogeneity is effect size-specific, each estimate has its own mean of the sampling distribution, $E_{\alpha_{3,ij}}(\hat{\alpha}_1) = \alpha_1 + \alpha_{3,ij}$. This might be the case when estimates in the same study are estimated for different sample sizes that do not share the same sample mean – for example, when multiple estimates are reported for different countries, where each country has its own mean of the sampling distribution.[65] In contrast, if random heterogeneity is constant within a study, each study shares the same mean of the sampling distribution, $E_{\alpha_{3,i}}(\hat{\alpha}_1) = \alpha_1 + \alpha_{3,i}$. Accordingly, all estimates come from a study level-specific distribution with the same mean – for example, if the samples and methods applied to produce several estimates in a study are similar. Modeling heterogeneity in such a way implies a panel structure in the meta-sample with study-level fixed effects.

The level of unexplained heterogeneity is calibrated by $\sigma_g = \{0, 0.125, 0.25, 0.5, 1, 2, 4\}$, which is the standard deviation of the random variable X_3. The magnitude of the implied heterogeneity can be quantified by I^2, which measures the proportion of unexplained heterogeneity in relation to the total variation (Higgins and Thompson, 2002: 1545):

$$I^2 = \frac{\hat{\tau}^2}{\hat{\tau}^2 + \sigma^2}, \tag{30}$$

where $\hat{\tau}^2$ is the estimate of the between-study variance and σ^2 is the within-study sampling variance. I^2 is defined in the range of 0–100%. When $I^2 = 0\%$, all variation in the effect size estimates is due to sampling error within the studies. Higgins et al. (2003) specify benchmark categories with values of 25%, 50%, and 75% to be considered as low, moderate, and high.

[65] As an extension of the simulation design, a scenario could be simulated whereby the effect size-specific random heterogeneity is different within a study as compared to the heterogeneity between studies.

According to Stanley and Doucouliagos (2017: 28), $I^2 = 80\%$ and more *"are the norm"* in economics.[66] In the hedging and firm value meta-analysis presented in Chapter 3, the I^2 is 83%, indicating a high level of heterogeneity.

Multiple Effect Size Estimates per Study. So far, the simulation design is similar to that of Stanley and Doucouliagos (2017). However, the defining feature of this chapter is the varying number of estimates generated in the primary studies. Stanley and Doucouliagos (2017), as well as many of the other prior simulation studies, assume that each primary study reports one estimate $\hat{\alpha}_1$. Accordingly, the number of studies equals the number of estimates. However, in economics and finance research, primary studies regularly report multiple estimates in the same study. Therefore, I assume that the primary researchers run the regression model (Eq. 29a or 29b) m_i times, where m_i varies across studies depending on the level of unbalancedness, measured by the Gini coefficient G.[67] The higher the Gini coefficient, the greater the unbalancedness among the number of estimates reported per study. When $G = 0$, all primary studies report the same number of estimates. When $G = 1$, one study reports all estimates.

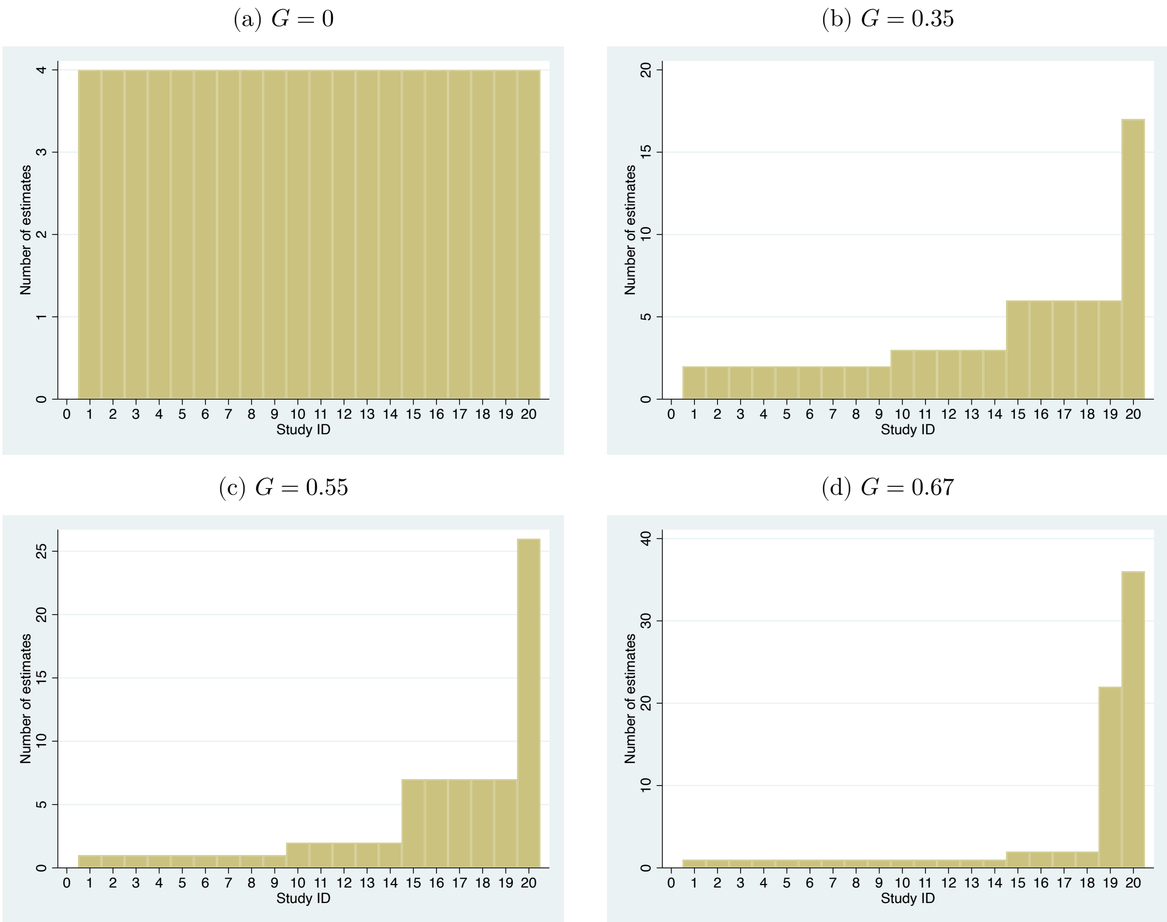
The Gini coefficient of the effect size estimates in the meta-analysis on hedging and firm value reported in Chapter 3 is 0.55. I take the value of 0.55 as a benchmark and produce a reasonable distribution of estimates per study that leads to this value of the Gini coefficient.[68] To examine the statistical properties of the meta-regression weights for different levels of unbalancedness of the meta-data set, four different values of the Gini coefficient $G = \{0, 0.35, 0.55, 0.67\}$ are considered. As Figure 32 shows, higher Gini coefficients indicate greater unbalancedness of the panel data, whereas $G > 0.67$ is an extreme scenario where the estimates from one primary study would dominate the entire sample. For each study, the number of estimates, m_i, is determined by G.

To generate multiple and non-independent effect size estimates in the same study, the researcher of study i draws C_i observations of the random variables, which are constant across the multiple estimates in the same study, and $N_{ij} - C_i$ observations that are randomly drawn for each estimate. The sample overlap, C_i, reflects that primary studies often report multiple effect size estimates for different subsamples with similar data sets. The number of overlapping observations, C_i, is defined by the minimum of the randomly selected sample sizes for study i

[66] Among nine of their meta-analysis studies in economics, the average I^2 is 93%.
[67] The Gini methodology was introduced by Corrado Gini in 1912 and is widely used in the areas of income distribution, risk analysis, and other research fields in economics (Yitzhaki and Schechtman, 2013: 1).
[68] A better calibration would be the distribution of Gini coefficients from the meta-data samples of the 76 finance meta-analyses reviewed in Chapter 2. However, most of these meta-analyses do not report the number of estimates observed from each study.

times the factor $d = [0,1]$, which calibrates the level of sample overlap. If $d = 0$, there is no sample overlap, while $d = 1$ indicates identical samples. For example, if $m_i = 3$, $N_{i1} = 62$, $N_{i2} = 250$, $N_{i3} = 1000$, and $d = 0.5$, a study runs the primary regression with three different sample sizes to produce three effect size estimates. Due to sample overlap, $C_i = 0.5 * \min\{62, 250, 1000\} = 31$ observations are the same for all effect size estimates in this study. The remaining observations, $N_{i1} - C_i = 31$, $N_{i2} - C_i = 219$, and $N_{i3} - C_i = 969$, are drawn randomly for each new estimate in the study.

Figure 32. Distribution of the Number of Effect Size Estimates for 20 Primary Studies



Notes: The graphs show the distribution of the number effect size estimates in a set of 20 primary studies. The number of reported estimates is determined by the Gini coefficient G. Higher values of G indicate greater unbalancedness of the meta-sample.

Publication Selection Bias. Without publication selection, effect size estimates are published regardless of their sign and statistical significance. The simulation model allows for the selection of statistical significance by a group of researchers who actively search for statistically significant estimates, $\hat{\alpha}_1$, and a second group that does not actively select specific results. If a researcher is in the group of the ‘file drawers’, she re-estimates the primary regression (Eq. 29a or 29b) when the obtained estimate, $\hat{\alpha}_1$, is not statistically significant at the 5% level. In the case of re-estimation, the researcher draws a new number for the overlapping observations, C_i, and takes a new set of the $N_{ij} - C_i$ random variables at the estimate level to run the primary regression again with the new data. The procedure continues until she finds, by chance, a value of the *t*-statistic that exceeds the desired threshold of statistical significance at the 5% level.

The number of file drawers is calibrated by f. For example, if $f = 0.5$, 50% of the estimates are biased by publication selection. Publication selection is randomly distributed across the 80 effect size estimates.[69] Accordingly, a study with multiple effect size estimates might report some effects without active searching for statistical significance, whereas other estimates are biased. If $f = 0.5$, half of the 80 estimates are based on re-estimation of the regression until a random but statistically significant estimate is found. For the other half of the estimates, the researcher reports the first observed estimate, independent of its statistical significance. Initially, I assume that there is no publication selection. Later in the chapter, active selection of statistically significant and positive estimates is considered (Tables 28 and 29).

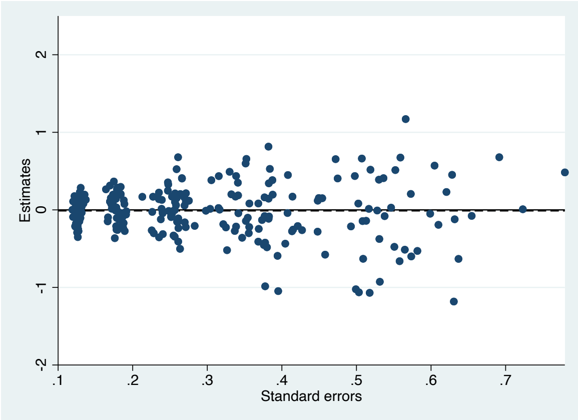
To illustrate the effect of publication selection and heterogeneity, Figure 33 shows scatter plots of 250 randomly generated effect size estimates, $\hat{\alpha}_1$, and their standard errors, $SE(\hat{\alpha}_1)$, following the steps 1–3 and 5 of the simulation process and assuming only one reported estimate per study. In the plots, it is supposed that $\alpha_1 = 0$, i.e., there is no true effect. Publication selection is either absent (left column) or calibrated at $f = 0.5$ (right column). Figures 33(a) and 33(b) assume that there is neither systematic heterogeneity (all regressions include X_2), nor random heterogeneity ($\sigma_g = 0$). Figures 33(c) and 33(d) assume that there is systematic heterogeneity, with 50% of the regressions omitting X_2. Figures 33(e) and 33(f) assume that, in addition to systematic heterogeneity, there is random heterogeneity calibrated at $\sigma_g = 0.5$.

Figure 33(a) forms a triangular shape with estimates distributed almost symmetrically around the true zero effect. Sampling error induces larger variation of estimates with larger standard errors (smaller sample sizes). When systematic heterogeneity is added (Figure 33c and 33d), omitted-variable bias increases the variation of the effect sizes and the mean effect size rises. Adding random heterogeneity (Figure 33e and 33f) increases the variation even more, also for precise estimates. When allowing for publication selection (Figures 33b, 33d, 33e), the plots become truncated with more estimates falling in the range of positive values, since negative and statistically insignificant estimates remain unpublished.

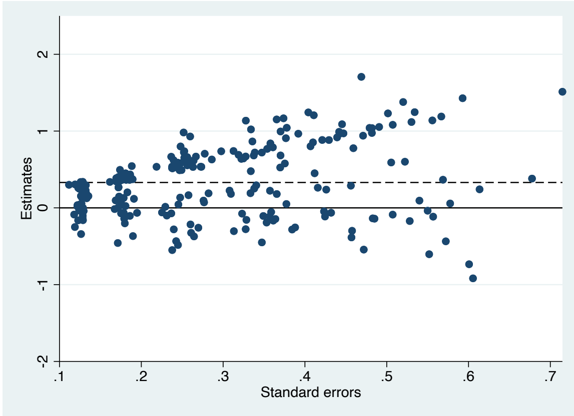
[69] In an extension of the simulation design, a scenario could be simulated whereby publication selection is not randomly distributed among the 80 effect size estimates, but clustered in specific studies, e.g., the studies that report a high number of effect size estimates.

Figure 33. Simulated Scatter Plot of 250 Effect Size Estimates and Standard Errors

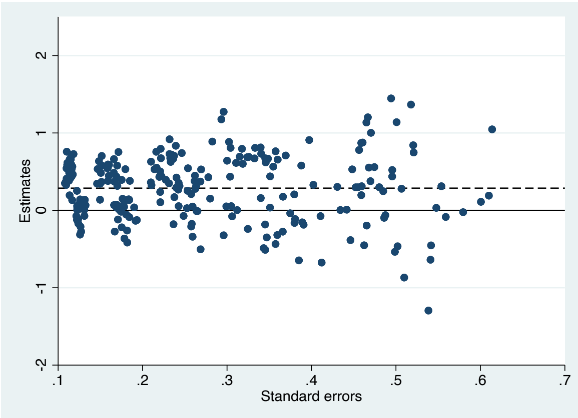
(a) No publication selection, No systematic heterogeneity, No random heterogeneity



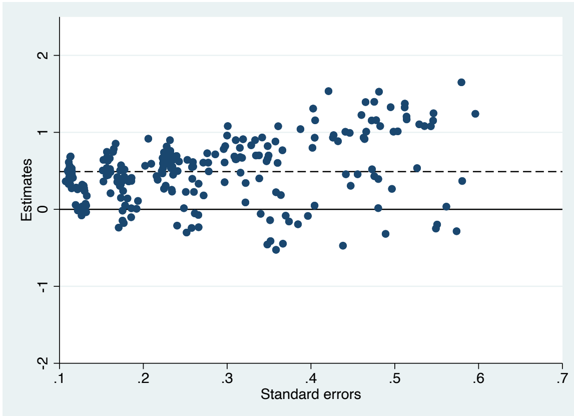
(b) Publication selection (50%), No systematic heterogeneity, No random heterogeneity



(c) No publication selection, Systematic heterogeneity, No random heterogeneity

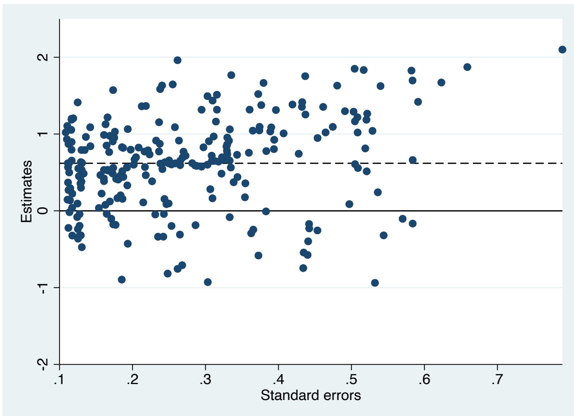


(d) Publication selection (50%), Systematic heterogeneity, No random heterogeneity



(e) No publication selection, Systematic heterogeneity, Random heterogeneity ($\sigma_g = 0.5$)

(f) Publication selection (50%), Systematic heterogeneity, Random heterogeneity ($\sigma_g = 0.5$)



Notes: These graphs show the simulated scatter plots of 250 effect size estimates generated by steps 1–3 and 5 of the simulation process (Figure 31). The true effect, α_1, is 0. For the cases without systematic heterogeneity, all estimates come from the regression model including X_2 (Eq. 29a). Random heterogeneity is parametrized by $\sigma_g = 0.5$. The solid line shows the true effect, while the dashed lines represent the average of the 250 estimated effects.

Primary Study Output. The steps above are repeated to produce a random sample of $i = 1, \dots, 20$ primary studies, where each study reports $j = 1, \dots, m_i$ effect size estimates and $\sum_{i=1}^{20} m_i = 80$. This leads to a final sample of 80 effect size estimates, $y_{ij} = \hat{\alpha}_{1,ij}$, and the corresponding standard errors, $SE(y_{ij})$.

4.4.3. Meta-Regression Simulation

After the simulation of the primary regressions, the meta-regression consumes the primary study results and estimates the mean effect size (Figure 34).

Figure 34. Process for the Simulation of Meta-Regression Results

1. Meta-Regression Data

Use the output of the simulation of primary study regressions (Figure 31).

2. Meta-Regression Model

Estimate $\hat{\beta}_0$ in the meta-regression model, including a moderator, M_{ij}, that is equal to 1 if the primary regression omits $X_{2,h}$, and 0 otherwise:

$$[1]\ y_{ij} = \beta_0 + \beta_1 M_{ij} + e'_{ij}$$

In the case of publication selection, the meta-regression considers the standard error of the effect size estimates as an additional variable to correct for publication selection bias:

$$[2]\ y_{ij} = \beta_0 + \beta_1 M_{ij} + \beta_2 SE(y_{ij}) + e''_{ij}$$

Estimate the meta-regression models [1] and [2] using WLS with three alternative weighting schemes:

1. $w_{ij} = 1/SE(y_{ij})^2$
2. $w_{ij} = 1/m_i$
3. $w_{ij} = 1/(SE(y_{ij})^2\ m_i)$

3. Meta-Regression Output

Repeat the previous steps (Figures 31 and 34) 10,000 times. For each iteration, record the estimated values, $\hat{\beta}_0$, and the corresponding standard error. Evaluate the performance of the meta-regression estimates by computing coverage proportions, mean squared errors, and bias of the estimator and average them across the 10,000 replications.

Notes: This figure describes the steps for the estimation of the meta-regression models using alternatives weights.

Meta-Regression Data. The reviewer collects the effect size estimates, y_{ij}, and the corresponding standard errors, $SE(y_{ij})$, from the 20 primary studies to run the meta-regression analysis. The reviewer can observe whether a primary study estimates a regression with or without X_2. Studies omitting X_2 are coded as $M_{ij} = 1$. In contrast to X_2, the third variable, X_3, which determines the true effect through its correlations with X_1, is omitted in all primary studies. Accordingly, meta-regression cannot correct for its omission.

Meta-Regression Model. After the data collection, the reviewer estimates the weighted meta-regression model including the moderator variable M_{ij}:

$$y_{ij} = \beta_0 + \beta_1 M_{ij} + e'_{ij}, \tag{31}$$

where M_{ij} corrects for the omitted-variable bias of X_2, and $\hat{\beta}_0$ is the estimate of the genuine effect. When the simulation allows for publication selection, the meta-regression model is extended by including the estimate's standard error as a second moderator variable:

$$y_{ij} = \beta_0 + \beta_1 M_{ij} + \beta_2 SE(y_{ij}) + e''_{ij}, \tag{32}$$

where $\hat{\beta}_2$ measures the impact of publication selection. Following Egger et al. (1997), the presence of publication selection bias can be detected by rejecting H_0: $\beta_2 = 0$.

The meta-regression model (Eq. 31 or 32) is estimated with three alternative weighting schemes that are explained in more detail in the Section 4.3:

1. Inverse variance of the effect size estimate: $w_{ij} = \frac{1}{SE(y_{ij})^2} = V_{ij}^{-1}$
2. Inverse number of reported effect size estimates per study: $w_{ij} = \frac{1}{m_i} = m_i^{-1}$
3. Combined weight of (1) and (2): $w_{ij} = \frac{1}{SE(y_{ij})^2\ m_i} = \left(V_{ij}\ m_i\right)^{-1}$

Meta-Regression Output. All steps described above are repeated 10,000 times, simulating 10,000 meta-analyses and their respective primary study samples. Previous Monte Carlo simulations with a similar simulation design show that the results are rather stable for 10,000 runs (Bom and Rachinger, 2019; Reed et al., 2015; Stanley, 2017; Stanley and Doucouliagos, 2017). For each run, the estimated mean effect size, $\hat{\beta}_0$, and the corresponding standard error, $SE(\hat{\beta}_0)$, are recorded. Averages of the statistical properties (Section 4.4.4) are calculated to evaluate the performance of the three weighting schemes.

Alternative Calibrations. The simulations are repeated for different calibrations of the input parameters in order to test the sensitivity of the meta-regression estimates. Each 'group' of simulations consists of 56 runs of the simulation procedure with 10,000 replications each. The 56 runs iterate through three dimensions:

- the level of unbalancedness of the reported estimates: $G = \{0, 0.33, 0.55, 0.67\}$,
- the level of unexplained heterogeneity: $\sigma_g = \{0, 0.125, 0.25, 0.5, 1, 2, 4\}$,
- the true effect: $\alpha_1 = \{0{,}1\}$.

Several alternative groups of simulations are evaluated by changing the level of sample overlap $d = \{0, 0.25, 0.5, 0.95\}$, adding study-level random heterogeneity, allowing for publication selection, and estimating the meta-regression model with cluster-robust standard errors, which is a common routine in applied meta-regression analyses.

4.4.4. Statistical Properties of Meta-Regression Weights

To evaluate and compare the performance of the meta-regression weights, coverage proportions, mean squared error (MSE), and the bias of the meta-regression estimator are calculated and averaged across the 10,000 replications of each simulation. The three statistical properties are often used to evaluate performance in simulation studies (among others, Gómez-Déniz and Calderín-Ojeda, 2011; Stanley and Doucouliagos, 2017; Trikalinos et al., 2013).

Coverage rates are computed by constructing for each run, r, of the simulation, the 95% CIs around the meta-regression estimate, $\hat{\beta}_{0,r}$. The coverage is the proportion of the 10,000 CIs of the estimated effect size that actually contain the true mean effect $\alpha_1 = \{0{,}1\}$ (Stanley and Doucouliagos, 2017: 30):

$$\text{Coverage} = \frac{1}{R} \sum_{r=1}^{R} I(\alpha_1 \in [95\% \text{ CI of } \hat{\beta}_{0,r}]), \tag{33}$$

where R is the total number of runs in the simulation and $I(\bullet)$ is an indicator that takes the value of 1 if the argument is true and 0 otherwise. The desirable value of the coverage rate is the 95% nominal level.

The MSE is the average squared difference between the simulated estimate and the true mean effect (Trikalinos et al., 2013: 14):

$$\text{MSE} = \frac{1}{R} \sum_{r=1}^{R} \left(\hat{\beta}_{0,r} - \alpha_1\right)^2. \tag{34}$$

Lower values of the MSE indicate higher efficiency of the estimator (Bom and Rachinger, 2020: 824). As the difference is squared, positive and negative deviations of the estimate from the true mean do not cancel out. Accordingly, the MSE can be high, even when the bias is low.

The bias is calculated as the absolute value of the difference between the average of the meta-regression estimates across the 10,000 replications and the true mean effect (Stanley and Doucouliagos, 2017: 30):

$$\text{Bias} = \left| \left(\tfrac{1}{R} \textstyle\sum_{r=1}^{R} \hat{\beta}_{0,r} \right) - \alpha_1 \right|. \tag{35}$$

Moreover, I^2 (Eq. 30) is used to indicate the level of unexplained heterogeneity. The values for I^2 are estimated for a meta-regression when there is no systematic heterogeneity and no publication selection. The values of I^2 are reported as averages across the 10,000 replications.

4.5. Simulation Results

The presentation of the results is structured as follows. Section 4.5.1 presents the simulation outcomes with varying levels of sample overlap when random heterogeneity is defined at the effect size level. Section 4.5.2 defines random heterogeneity at the study level. Section 4.5.3 introduces publication selection, and Section 4.5.4 applies cluster-robust standard errors in the meta-regression to account for non-independent errors.

Table 22 shows an overview of the groups of simulations that are reported and described in the subsequent sections. d denotes the level of sample overlap, α_3 is the random heterogeneity, f calibrates the level of publication selection, and clustered errors refer to the application of cluster-robust standard errors at the study level.

Table 22. Overview of Calibrations of the Simulation Models

No.	d	α_3	f	Clustered errors	Table
1	0	Effect size level	0	No	23
2	0.25	Effect size level	0	No	Appendix C.1
3	0.5	Effect size level	0	No	Appendix C.3
4	0.95	Effect size level	0	No	25
5	0	Study level	0	No	26
6	0.5	Study level	0	No	Appendix C.5
7	0.5	Study level	0	No	Appendix C.7
8	0.95	Study level	0	No	27
9	0.5	Study level	0.25	No	Appendix C.9
10	0.5	Study level	0.50	No	28
11	0.5	Study level	0.75	No	29
12	0.5	Study level	0	Yes	30

Notes: This table shows the groups of simulations that are constructed by changing key parameters in the simulation design.

The subsequent results tables report the averages of 10,000 simulated meta-analyses. Each row presents the findings of 10,000 runs for a specific set of a parameters that calibrate the simulation. In all cases, the meta-regression estimate of the true effect, $\hat{\beta}_0$, is the quantity of investigation. G is the Gini coefficient that defines the unbalancedness of the meta-panel data arising from the varying number of estimates reported in the same study. σ_g is the standard deviation of α_3 determining random heterogeneity. I^2 measures random heterogeneity relative to sampling error. α_1 is the true effect that is either 0 or 1. The columns report the coverage rates (Eq. 33 times 100), mean squared errors (Eq. 34 times 1,000), and the bias of the estimates (Eq. 35 times 1,000). $\boldsymbol{V}^{-1}$ refers to the results with inverse variance weighting, $\boldsymbol{m}^{-1}$ is the weighting by the inverse number of estimates per study, and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the combined weight. For now, it is assumed that there is no publication selection ($f = 0$).

4.5.1. Effect Size-Level Heterogeneity

In this section, random heterogeneity, α_3, is defined at the level of the individual effect size estimates. Accordingly, α_3 is randomly generated for each effect size estimate and varies from estimate to estimate. The mean of the distribution of a given estimated effect is $\alpha_1 + \alpha_{3,ij}$.

Sample Overlap $\boldsymbol{d} = \mathbf{0}$. Since the sample overlap parameter d is calibrated to 0, multiple estimates from the same study are based on independent random data samples. This scenario with no sample overlap and random heterogeneity at the effect size level is the simulation that is most similar to the case with 80 studies and random omitted-variable bias reported by Stanley and Doucouliagos (2017: 27-28). Because Stanley and Doucouliagos (2017) calibrate the simulation design such that each primary study reports one effect size estimate, random heterogeneity is defined analogously to the effect size-specific heterogeneity in this section.

Unsurprisingly, inverse variance, which is the BLUE in the case of heteroscedasticity, performs best in this scenario. Looking at the average coverage rates, MSE, and biases across all parametrizations, inverse variance weighting outperforms the alternative weights, having the lowest average MSE and bias, and an average coverage rate of 92.97%, which exhibits the lowest deviation from the nominal of 95%. A coverage rate of 92.97% indicates that in 9,297 of the 10,000 replications, the CIs around the MRA estimates actually contain the true mean effect. When the 10,000 replications of the simulation are repeated five times, the mean absolute deviation from one average bias reported in Table 23 to another is approximately 0.134 for $\boldsymbol{V}^{-1}$, 0.317 for $\boldsymbol{m}^{-1}$, and 0.144 for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. For the MSE, the mean absolute deviation is 0.072 for $\boldsymbol{V}^{-1}$, 0.384 for $\boldsymbol{m}^{-1}$, and 0.170 for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. Coverage rates vary by 0.036 for $\boldsymbol{V}^{-1}$, 0.041 for $\boldsymbol{m}^{-1}$, and 0.043 for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ from one of the 10,000 replications to another.

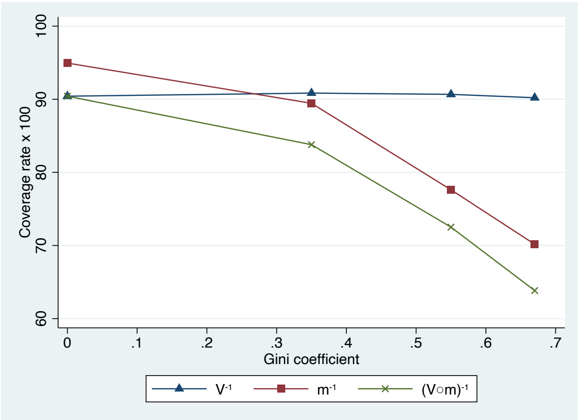
Table 23. Simulation Results ($d = 0$, $f = 0$, Random Heterogeneity at ES Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.37	0	94.99	93.61	94.99	1.06	2.62	1.06	0.07	0.62	0.07
0	0.125	28.24	0	92.72	93.33	92.72	1.77	3.08	1.77	0.12	0.25	0.12
0	0.25	60.85	0	91.00	94.46	91.00	3.67	4.13	3.67	0.42	0.36	0.42
0	0.5	85.54	0	90.42	94.95	90.42	10.38	8.82	10.38	0.22	1.18	0.22
0	1	95.40	0	91.78	94.96	91.78	28.59	27.83	28.59	0.41	1.80	0.41
0	2	98.45	0	93.76	95.03	93.76	70.10	104.56	70.10	1.32	1.95	1.32
0	4	99.40	0	95.82	95.23	95.82	149.61	414.16	149.61	6.96	1.22	6.96
0.35	0	4.40	0	95.07	88.17	89.93	1.07	3.66	1.50	0.71	0.66	0.66
0.35	0.125	28.13	0	93.03	88.80	87.72	1.74	4.12	2.40	0.35	0.06	0.27
0.35	0.25	60.97	0	91.14	88.96	84.50	3.65	5.91	5.17	0.51	0.54	0.79
0.35	0.5	85.58	0	90.84	89.45	83.79	10.20	12.63	14.49	1.02	0.16	0.96
0.35	1	95.43	0	91.96	89.21	85.66	28.42	41.02	40.91	1.61	1.36	1.50
0.35	2	98.45	0	94.25	90.29	88.81	68.45	147.98	96.74	1.56	4.61	2.60
0.35	4	99.40	0	95.55	90.19	90.62	152.20	581.98	216.45	3.34	5.12	0.43
0.55	0	4.50	0	95.28	76.85	79.23	1.04	6.27	2.55	0.28	0.86	0.38
0.55	0.125	28.01	0	92.92	76.20	75.30	1.77	7.43	4.19	0.55	0.59	1.31
0.55	0.25	60.93	0	91.04	76.68	71.70	3.68	10.49	8.72	0.50	2.93	2.12
0.55	0.5	85.60	0	90.67	77.62	72.51	10.28	21.98	23.58	1.75	0.45	2.36
0.55	1	95.43	0	91.82	77.78	73.28	28.56	69.38	67.50	3.16	1.13	3.48
0.55	2	98.45	0	94.09	78.34	76.18	69.51	256.85	165.84	1.24	3.60	5.57
0.55	4	99.40	0	95.38	79.32	79.28	151.50	972.82	371.71	1.54	0.33	0.26
0.67	0	4.54	0	94.98	68.77	71.55	1.05	8.34	3.53	0.34	1.14	0.97
0.67	0.125	28.04	0	92.76	68.38	68.09	1.75	9.84	5.62	0.81	0.66	1.48
0.67	0.25	60.91	0	90.96	69.29	64.00	3.76	13.70	11.88	0.28	0.86	0.66
0.67	0.5	85.56	0	90.21	70.17	63.84	10.62	29.01	33.08	0.29	0.31	0.24
0.67	1	95.41	0	91.54	70.71	64.97	28.77	92.73	92.84	0.85	2.09	0.84
0.67	2	98.45	0	94.07	70.92	68.48	68.01	344.96	237.60	0.48	2.03	1.08
0.67	4	99.40	0	96.11	71.19	72.02	147.40	1357.61	542.15	0.44	2.95	6.07
0	0	4.52	1	95.17	94.07	95.17	1.04	2.57	1.04	0.08	0.07	0.08
0	0.125	28.12	1	92.94	94.06	92.94	1.75	2.97	1.75	0.35	1.15	0.35
0	0.25	60.89	1	91.24	94.19	91.24	3.68	4.20	3.68	0.02	0.66	0.02
0	0.5	85.55	1	90.71	94.79	90.71	10.41	8.97	10.41	1.90	1.36	1.90
0	1	95.42	1	91.80	94.82	91.80	28.90	28.28	28.90	0.20	0.17	0.20
0	2	98.45	1	94.24	94.91	94.24	68.29	105.95	68.29	0.35	2.31	0.35
0	4	99.40	1	95.65	95.21	95.65	150.19	408.61	150.19	8.86	13.31	8.86
0.35	0	4.53	1	94.91	88.19	89.75	1.07	3.64	1.51	0.34	0.08	0.13
0.35	0.125	28.19	1	92.93	88.34	87.10	1.77	4.21	2.45	0.30	0.61	0.19
0.35	0.25	60.94	1	90.96	88.69	84.54	3.66	5.95	5.14	0.41	0.78	1.01
0.35	0.5	85.55	1	90.08	89.81	84.04	10.37	12.41	14.12	0.70	1.70	1.23
0.35	1	95.41	1	91.88	89.72	85.92	28.40	40.03	39.79	2.03	2.38	2.47
0.35	2	98.45	1	93.83	90.00	88.36	69.17	145.20	97.56	2.63	0.70	2.78
0.35	4	99.40	1	95.59	90.48	90.66	151.04	566.81	214.54	6.55	14.02	7.75
0.55	0	4.43	1	95.28	76.89	79.91	1.03	6.15	2.48	0.11	0.10	0.38
0.55	0.125	28.31	1	93.03	76.66	75.07	1.75	7.21	4.15	0.34	0.62	0.29
0.55	0.25	60.97	1	90.91	76.99	72.53	3.71	10.12	8.57	0.94	0.05	1.37
0.55	0.5	85.56	1	90.13	77.67	70.67	10.28	21.92	24.39	0.46	4.00	2.95
0.55	1	95.42	1	91.04	78.56	73.28	29.15	68.87	67.98	2.30	0.23	1.06
0.55	2	98.45	1	93.81	78.76	75.82	69.39	256.65	168.87	2.07	9.47	3.73
0.55	4	99.40	1	95.77	78.67	79.18	149.05	997.45	369.56	6.38	14.86	13.62
0.67	0	4.52	1	94.95	68.85	71.47	1.04	8.50	3.50	0.20	0.17	0.06
0.67	0.125	28.09	1	93.10	68.14	67.64	1.74	9.97	5.72	0.25	0.21	0.01
0.67	0.25	60.90	1	90.74	68.80	64.01	3.79	14.04	12.00	0.36	0.90	2.00
0.67	0.5	85.56	1	89.83	70.24	62.79	10.56	29.11	33.44	0.74	0.84	2.07
0.67	1	95.41	1	92.07	71.15	64.85	27.98	91.51	93.27	1.07	4.34	3.01
0.67	2	98.45	1	94.51	71.43	68.74	67.19	334.88	231.91	1.49	0.63	1.28
0.67	4	99.40	1	95.29	70.45	70.84	155.32	1381.19	535.58	1.65	7.09	0.64
Average				**92.97**	**82.85**	**80.73**	**37.86**	**163.24**	**78.83**	**1.33**	**2.19**	**1.85**

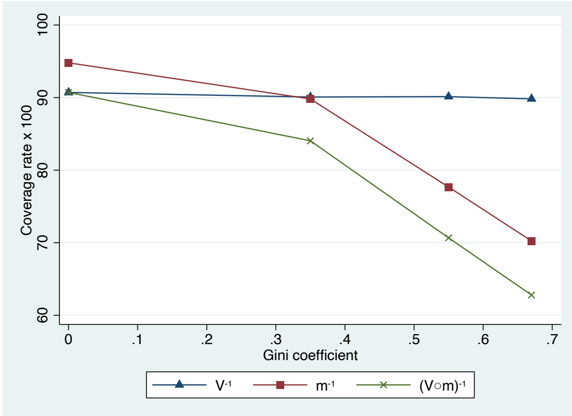
Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. V^{-1} is the meta-regression with inverse variance weighting, m^{-1} is the weighting by the inverse number of estimates per study, and $(V \circ m)^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Figure 35. Coverage, MSE, Bias ($d = 0$, $f = 0$, Random Heterogeneity at ES Level)

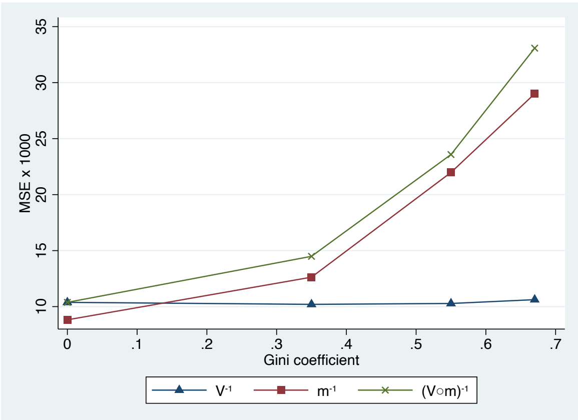
(a) Coverage; $\alpha_1 = 0$



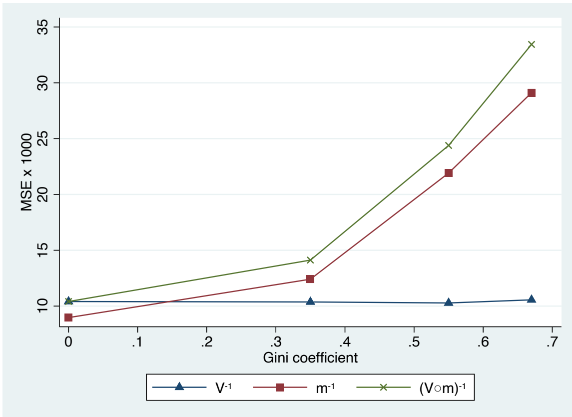
(b) Coverage; $\alpha_1 = 1$



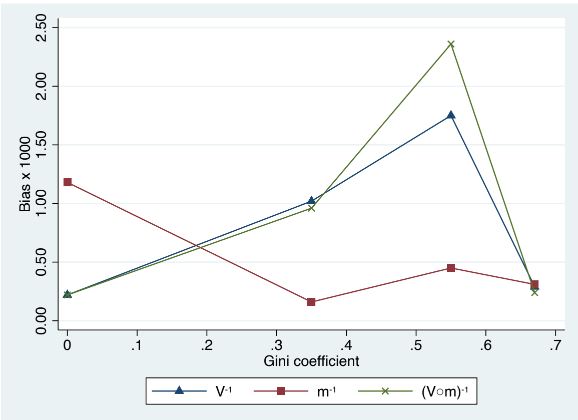
(c) MSE; $\alpha_1 = 0$



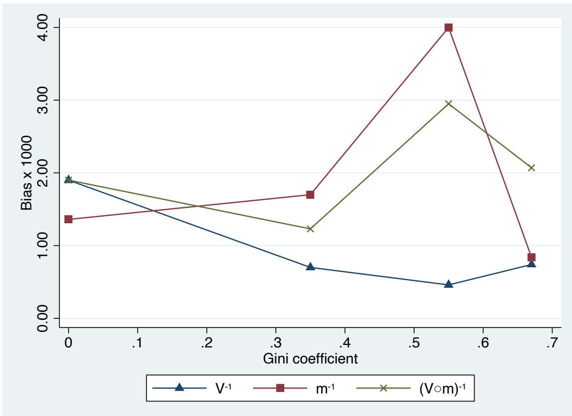
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: These graphs show the coverage, MSE, and bias reported in Table 23 with alternating levels of unbalancedness, measured by the Gini coefficient, plotted on the horizonal axis and the true effect being either $\alpha_1 = 0$ or $\alpha_1 = 1$. Graphs (a) and (b) show the coverage rates. Graphs (b) and (d) show the mean squared errors. Graphs (e) and (f) plot the bias. Random heterogeneity is parametrized at $\sigma_g = 0.5$.

Similar to Stanley and Doucouliagos (2017), the coverage rates of inverse variance weighting are around 95% for low and high values of σ_g, especially $\sigma_g = \{0{,}4\}$. The largest deviation from the nominal can be observed for medium random heterogeneity, $\sigma_g = 0.5$, where coverage rates are around 90% for $\boldsymbol{V}^{-1}$. Interestingly, when $G = 0$ and $\alpha_1 = 0$, the coverage rates for $\boldsymbol{m}^{-1}$ decreases from an average of 95% across all calibrations of σ_g to an average of 69% when $G = 0.67$ and $\alpha_1 = 0$. Hence, larger unbalancedness drives the coverage rates of the alternative weights, $\boldsymbol{m}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$, away from the nominal. At the same time, the MSE of the two alternative weights is much higher than the MSE of $\boldsymbol{V}^{-1}$, making the alternative weights non-efficient estimators. The bias of all estimators is rather small; practically speaking, they are almost nil. For example, the average bias of $\boldsymbol{V}^{-1}$ is a bit larger than 0.1%, a result that is also reported by Stanley and Doucouliagos (2017: 30). When the Gini coefficient is 0, meaning that each study reports the same number of estimates, and σ_g is at a medium to high level, I can find a few cases in which coverage rates of $\boldsymbol{m}^{-1}$ outperform the other weighting schemes. For example, if $G = 0$ and $\alpha_1 = 0$, the coverage rate of $\boldsymbol{m}^{-1}$ at $\sigma_g = 0.5$ is 94.95%, in contrast to 90.42% for $\boldsymbol{V}^{-1}$. Another finding is that, regardless of the magnitude of the Gini coefficient, MSE and bias increase as random heterogeneity rises, which is plausible because heterogeneity increases the variation of the effect size estimates, causing the estimates to deviate more from the true mean effect. For example, at $G = 0.55$ and $\alpha_1 = 0$, the MSE of $\boldsymbol{V}^{-1}$ for $\sigma_g = 4$ is 146 times higher than for $\sigma_g = 0$. For the other weights, the spread is even higher, 155 times for $\boldsymbol{m}^{-1}$ and 144 times for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$.

To illustrate the performance of the meta-regression weights as a function of the Gini coefficient, Figure 35 depicts the results of the statistical properties at alternating levels of G. Random heterogeneity is calibrated at a medium level, $\sigma_g = 0.5$, which is the case with the lowest coverage rates for inverse variance weighting. Figure 35(a) shows the coverage when $\alpha_1 = 0$. Figure 35(b) depicts the coverage of the meta-regression results when $\alpha_1 = 1$, which is the proportion of meta-regression tests that correctly identify the true effect. The coverage rate is the highest for $\boldsymbol{m}^{-1}$ when $G = 0$. However, when the unbalancedness of the meta-panel data increases, i.e., when studies report a wider range of estimates per study, the coverage proportions of $\boldsymbol{V}^{-1}$ are superior. The relative performance of the other two alternative weights becomes worse when the Gini coefficient increases. A similar pattern can be detected for the MSE in the Figures 35(c) and 35(d), where inverse variance weighting dominates as it is the most efficient estimator. According to the Figures 36(e) and 36(f), the biases of all three estimators are rather small and do not exceed rounding errors. Hence, the biases are not interpreted any further.

Sample Overlap $\boldsymbol{d} = \{\mathbf{0.25}, \mathbf{0.5}, \mathbf{0.95}\}$. In the previous scenario, where $d = 0$, each of the 80 effect size estimates in the 20 primary studies is estimated by a new random sample. However, when a study reports multiple effect size estimates, the effect sizes are often estimated with overlapping samples. In this case, the effect size estimates are not independent. For example, considering a study that examines an empirical effect in a specific country and estimates that effect for overlapping time periods, for example, 1990–2000 and 1995–2005, it is reasonable to assume that the effect sizes estimated for the respective time periods depend on each other, because they are based on a sample from the same country and overlapping time periods. In applied meta-regressions in economics and finance, such sample overlap is a common feature. Therefore, it is reasonable to assume that $d > 0$.

The factor d determines the level of within-study sample overlap. Larger values of d imply a higher number of observations that are constant across multiple estimates in the same study. The larger the sample overlap, the lower the share of independent sampling information that is stored in each of the individual primary study estimates (Bom and Rachinger, 2020: 829).

Following Bom and Rachinger (2020: 820), the within-study correlation implied by sample overlap can be quantified by the average correlation between the pairs of overlapping samples. The pairwise correlation between any two overlapping estimates, y_{ij} and y_{ij+1}, is given by:

$$\rho_{y_{ij},y_{ij+1}} = \frac{C_i}{\sqrt{N_{ij}N_{ij+1}}}, \tag{36}$$

where C_i is the number of overlapping observations within a study.

Give an example with $N_{ij} = 62$, $N_{ij+1} = 250$, and $C_i = 31$, the pairwise correlation is 0.25. For the average within-study correlation, all pairwise correlations are transformed by Fisher's z prior to averaging. As Silver and Dunlap (1987) show, the average of z_{trans} is less biased than the average of ρ. The average values are transformed back into the correlation metric for interpretation. Hence, $\bar{z}_{trans,i}$ is the average of all pairwise and transformed correlations within a study, i.e., all off-diagonal elements in the variance-covariance matrix (Eq. 24).

The average across all studies and replications of the simulation is given by:

$$\bar{z}_{trans,R} = \frac{1}{R}\sum\nolimits_{r=1}^{R}\left(\frac{1}{L}\sum\nolimits_{i=1}^{L}\bar{z}_{trans,i,r}\right), \tag{37}$$

where $\bar{z}_{trans,R}$ is the average transformed correlation of a simulation with 10,000 replications. Transforming the result back into the correlation metric leads to the average correlations, $\bar{\rho}_R$, which are reported in Table 24 as a function of the sample overlap that is determined by the factor d.

Table 24. Average Within-Study Correlation for Different Levels of d

d	$\bar{\rho}_R$
0	0.00
0.25	~ 0.13
0.5	~ 0.25
0.95	~ 0.46

Notes: This table shows the average within-study correlations for the same simulation as in Table 23, with 10,000 runs and varying levels of the sample overlap factor d.

Appendices C.1 and C.2 show the results for the same simulation configuration as in Table 23 when the sample overlap increases. Appendix C.1 presents the results for $d = 0.25$ and Appendix C.2 for $d = 0.5$. The findings for the 'extreme' scenario, when $d = 0.95$, are shown in the subsequent Table 25. Comparing the results for $d = \{0.25, 0.5, 0.95\}$ with the case in which there is no sample overlap (Table 23), we can see that the average values for the coverage, MSE, and bias are rather similar. For example, when $d = 0.95$, the coverage for the inverse variance weighting $\boldsymbol{V}^{-1}$ is 92.97% for $d = 0$ (Table 23) and 92.92% for $d = 0.95$ (Table 25). Similarly, coverage ratios are 82.85% (Table 23) and 82.79% (Table 25) for $\boldsymbol{m}^{-1}$ and 80.73% (Table 23) and 80.80% (Table 25) for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. The MSE and biases are also similar. Therefore, sample overlap as modeled in this simulation does not appear to have a significant impact on the performance of the WLS weights when random heterogeneity is effect size-specific.

Figure 36 shows the change in the statistical properties of the three weighting schemes as a function of the magnitude of the Gini coefficient when $\sigma_g = 0.5$. The coverage proportions and MSE are almost identical to the case without sample overlap (Figure 35).

In summary, a simple overlap structure, where each study that reports two or more estimates has a fixed number of overlapping observations, does not affect the relative performance of the three examined weighting schemes. Accordingly, in the case of effect size-level random heterogeneity, inverse variance meta-regression is the superior approach.

4.5.2. Study-Level Heterogeneity

Thus far, random heterogeneity has been defined at the effect size level. Accordingly, each effect size estimate, within the same study or across studies, has its own mean of the sampling distribution, $\alpha_1 + \alpha_{3,ij}$. In this section, random heterogeneity is constant within a single study, but varies between studies. If a study reports multiple effect size estimates, all estimates within this study have the same mean of the sampling distribution, $\alpha_1 + \alpha_{3,i}$.

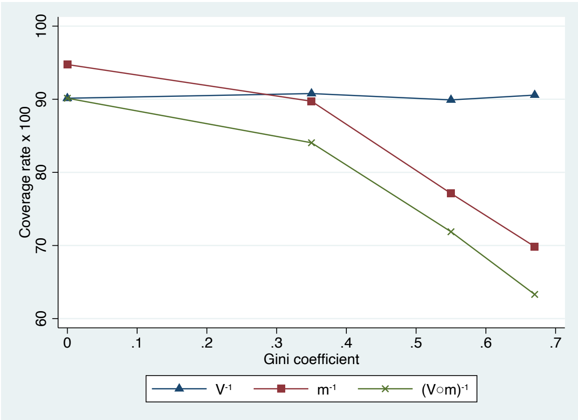
Table 25. Simulation Results ($d = 0.95$, $f = 0$, Random Heterogeneity at ES Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.46	0	95.00	93.57	95.00	1.06	2.58	1.06	0.15	0.42	0.15
0	0.125	28.03	0	93.59	93.59	93.59	1.67	3.01	1.67	0.40	0.06	0.40
0	0.25	60.89	0	90.78	94.55	90.78	3.76	4.17	3.76	0.16	0.49	0.16
0	0.5	85.54	0	90.14	94.75	90.14	10.40	8.98	10.40	1.77	1.42	1.77
0	1	95.41	0	91.68	94.61	91.68	28.69	28.80	28.69	0.14	2.16	0.14
0	2	98.45	0	93.90	94.68	93.90	69.17	105.92	69.17	3.72	0.59	3.72
0	4	99.40	0	95.52	95.25	95.52	149.74	407.51	149.74	2.22	2.33	2.22
0.35	0	4.38	0	95.59	88.43	90.58	1.02	3.67	1.45	0.46	0.63	0.59
0.35	0.125	28.27	0	93.16	88.70	87.56	1.75	4.16	2.44	0.02	0.76	0.15
0.35	0.25	60.83	0	90.85	88.86	84.95	3.74	5.91	5.09	0.30	0.94	0.21
0.35	0.5	85.56	0	90.78	89.73	84.05	10.11	12.47	14.23	0.23	0.24	0.45
0.35	1	95.41	0	91.44	90.00	85.28	28.87	39.85	40.97	1.36	1.26	1.43
0.35	2	98.44	0	94.23	89.56	88.87	67.58	150.11	95.02	2.24	2.64	1.74
0.35	4	99.40	0	95.59	89.99	90.76	148.77	580.45	214.64	1.71	0.12	2.62
0.55	0	4.51	0	95.16	76.94	79.54	1.04	6.11	2.49	0.05	1.32	0.53
0.55	0.125	28.17	0	92.45	76.18	75.01	1.78	7.32	4.18	0.64	1.01	0.67
0.55	0.25	60.89	0	91.14	77.40	72.42	3.64	10.10	8.59	0.14	0.55	0.24
0.55	0.5	85.56	0	89.91	77.12	71.88	10.53	22.15	23.83	0.63	1.31	0.40
0.55	1	95.41	0	91.91	78.61	72.96	28.31	68.56	68.14	0.75	0.17	0.57
0.55	2	98.44	0	93.75	78.17	76.72	69.35	261.62	169.03	1.25	1.68	1.20
0.55	4	99.40	0	95.58	78.97	79.64	148.48	990.50	366.56	4.82	7.02	4.80
0.67	0	4.50	0	95.05	68.14	71.51	1.07	8.43	3.46	0.02	2.10	0.75
0.67	0.125	28.06	0	92.73	68.89	66.98	1.75	9.64	5.81	0.34	0.88	0.24
0.67	0.25	60.96	0	90.78	68.89	64.62	3.70	14.05	11.68	0.24	0.29	0.20
0.67	0.5	85.55	0	90.57	69.85	63.31	10.42	29.63	33.53	0.55	1.72	1.38
0.67	1	95.38	0	91.63	70.44	64.75	28.60	92.13	93.04	1.00	3.67	1.61
0.67	2	98.44	0	94.36	71.18	68.86	66.51	339.08	230.91	3.19	2.40	0.63
0.67	4	99.39	0	95.59	71.07	71.46	147.95	1360.63	541.91	0.54	28.03	13.62
0	0	4.44	1	95.13	93.79	95.13	1.05	2.57	1.05	0.38	0.00	0.38
0	0.125	28.17	1	93.21	93.84	93.21	1.73	2.99	1.73	0.46	0.06	0.46
0	0.25	60.80	1	90.62	93.95	90.62	3.75	4.28	3.75	1.23	0.51	1.23
0	0.5	85.59	1	90.93	94.54	90.93	10.14	8.86	10.14	0.21	0.02	0.21
0	1	95.41	1	91.39	94.63	91.39	29.50	29.27	29.50	0.57	1.11	0.57
0	2	98.45	1	93.70	95.01	93.70	69.21	104.61	69.21	1.36	2.45	1.36
0	4	99.40	1	95.42	95.10	95.42	155.91	411.50	155.91	4.39	6.54	4.39
0.35	0	4.50	1	94.81	88.55	89.99	1.07	3.60	1.51	0.10	0.46	0.37
0.35	0.125	28.06	1	92.68	88.20	87.45	1.75	4.30	2.44	0.06	0.18	0.05
0.35	0.25	60.83	1	90.90	88.50	84.73	3.70	5.97	5.14	0.27	0.85	0.63
0.35	0.5	85.58	1	90.37	89.65	84.08	10.38	12.50	14.44	0.01	0.40	0.90
0.35	1	95.40	1	91.55	90.27	86.50	28.21	38.76	38.88	2.41	2.37	1.79
0.35	2	98.44	1	94.00	89.43	88.49	69.09	150.50	95.84	4.30	0.02	4.35
0.35	4	99.40	1	95.61	90.01	90.15	150.80	581.29	219.83	1.72	5.52	1.98
0.55	0	4.39	1	94.87	76.65	79.41	1.07	6.26	2.53	0.15	0.50	0.30
0.55	0.125	28.08	1	92.82	77.18	75.20	1.79	7.12	4.23	0.28	0.46	0.92
0.55	0.25	60.91	1	90.64	76.31	72.30	3.82	10.42	8.77	0.00	1.88	0.49
0.55	0.5	85.56	1	90.31	77.64	71.96	10.45	21.57	23.75	0.21	0.70	0.57
0.55	1	95.41	1	92.16	78.74	73.31	28.31	67.91	66.83	1.79	3.27	2.43
0.55	2	98.45	1	93.50	78.38	76.09	70.96	255.70	167.18	1.47	1.98	1.43
0.55	4	99.40	1	95.31	77.30	78.78	154.99	1034.42	387.75	4.35	1.07	5.99
0.67	0	4.44	1	94.92	68.55	71.42	1.07	8.39	3.53	0.11	0.47	0.07
0.67	0.125	28.08	1	93.15	69.49	68.17	1.74	9.68	5.60	0.83	2.00	1.08
0.67	0.25	60.86	1	90.95	68.97	65.01	3.68	13.72	11.74	1.20	0.64	0.44
0.67	0.5	85.50	1	90.05	70.57	63.14	10.56	29.13	33.28	1.39	0.18	0.75
0.67	1	95.39	1	92.00	70.60	65.30	28.35	91.75	91.20	2.33	3.49	1.90
0.67	2	98.44	1	94.39	70.94	69.22	68.14	343.66	231.18	0.91	0.34	4.06
0.67	4	99.39	1	95.09	71.25	71.64	153.32	1330.28	539.10	2.30	20.07	2.51
Average				**92.92**	**82.79**	**80.80**	**37.93**	**163.72**	**79.06**	**1.14**	**2.21**	**1.50**

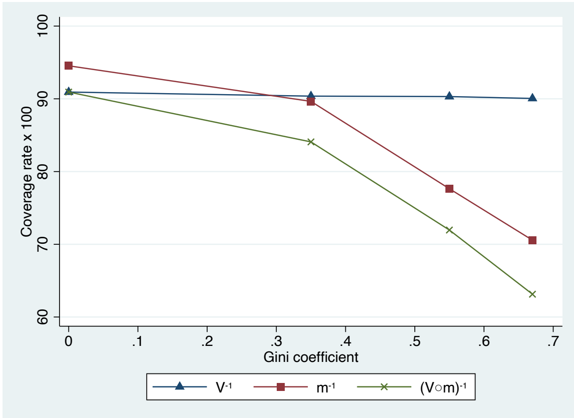
Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. V^{-1} is the meta-regression with inverse variance weighting, m^{-1} is the weighting by the inverse number of estimates per study, and $(V \circ m)^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Figure 36. Coverage, MSE, Bias ($d = 0.95$, $f = 0$, Random Heterogeneity at ES Level)

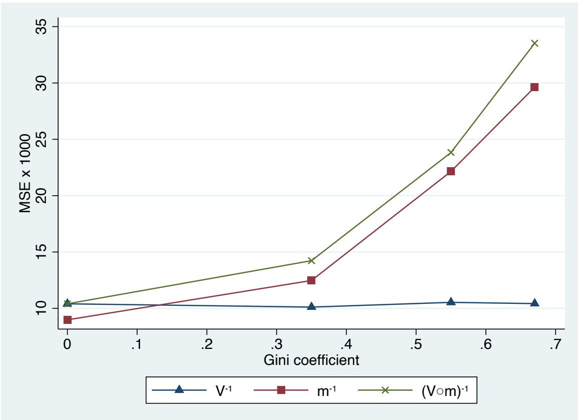
(a) Coverage; $\alpha_1 = 0$



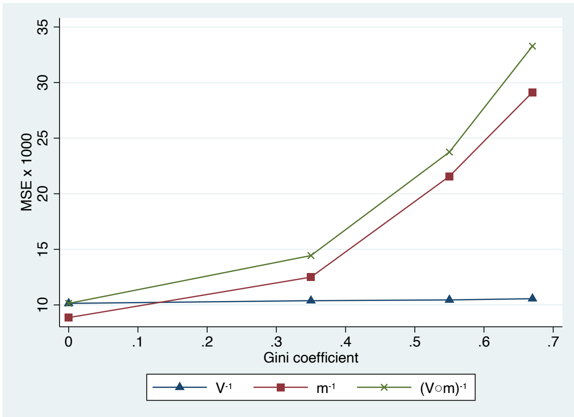
(b) Coverage; $\alpha_1 = 1$



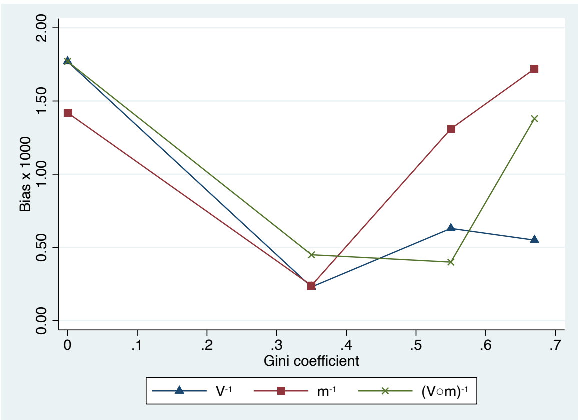
(c) MSE; $\alpha_1 = 0$



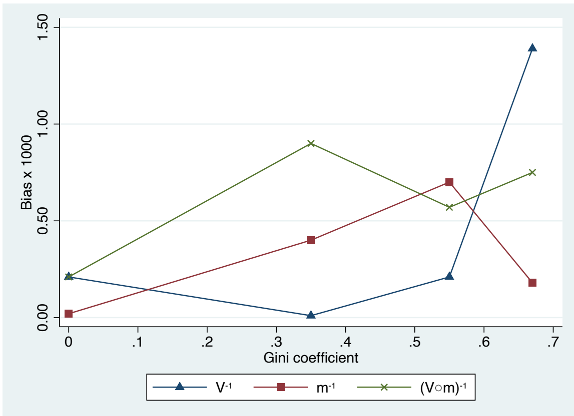
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: These graphs show the coverage, MSE, and bias reported in Table 23 with alternating levels of unbalancedness, measured by the Gini coefficient, plotted on the horizonal axis and the true effect being either $\alpha_1 = 0$ or $\alpha_1 = 1$. Graphs (a) and (b) show the coverage rates. Graphs (b) and (d) show the mean squared errors. Graphs (e) and (f) plot the bias. Random heterogeneity is parametrized at $\sigma_g = 0.5$.

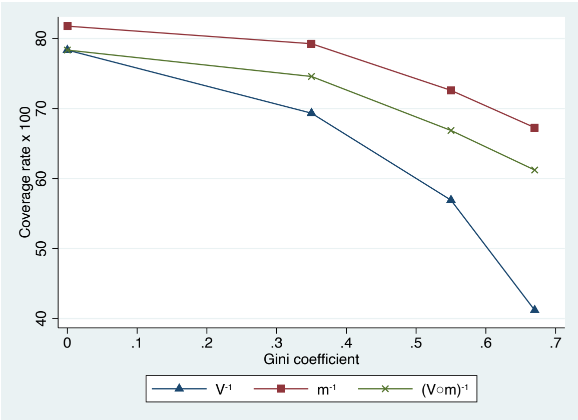
Table 26. Simulation Results ($d = 0$, $f = 0$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.50	0	94.97	93.99	94.97	1.03	2.57	1.03	0.37	0.36	0.37
0	0.125	26.99	0	88.72	91.19	88.72	2.30	3.52	2.30	0.33	0.09	0.33
0	0.25	58.83	0	81.42	86.73	81.42	5.85	6.52	5.85	0.01	0.05	0.01
0	0.5	84.23	0	78.35	81.77	78.35	17.98	18.09	17.98	0.41	0.69	0.41
0	1	94.91	0	79.64	79.16	79.64	51.78	64.92	51.78	3.25	2.06	3.25
0	2	98.25	0	82.31	77.78	82.31	128.24	249.79	128.24	0.18	3.60	0.18
0	4	99.30	0	85.88	77.38	85.88	284.62	1028.95	284.62	4.75	1.90	4.75
0.35	0	4.47	0	94.93	88.73	90.03	1.05	3.64	1.50	0.03	0.24	0.43
0.35	0.125	26.17	0	84.13	85.86	83.30	2.91	4.81	2.98	0.37	0.44	0.45
0.35	0.25	57.36	0	74.93	83.22	77.84	7.83	7.81	6.91	2.33	1.13	2.68
0.35	0.5	83.26	0	69.34	79.26	74.58	25.48	20.60	20.97	0.24	0.17	0.65
0.35	1	94.49	0	70.79	76.94	76.32	72.42	70.53	58.91	4.10	5.16	2.43
0.35	2	98.10	0	72.41	75.56	79.34	192.79	273.01	150.10	1.63	3.19	1.75
0.35	4	99.23	0	75.58	75.92	81.84	441.21	1095.13	341.57	5.86	1.22	3.07
0.55	0	4.45	0	94.87	76.07	79.30	1.07	6.34	2.55	0.44	0.08	0.20
0.55	0.125	24.31	0	77.93	76.13	73.34	3.72	7.38	4.48	0.26	1.61	0.46
0.55	0.25	54.72	0	64.78	74.22	69.95	11.48	11.31	9.76	1.95	1.50	0.59
0.55	0.5	81.29	0	56.92	72.60	66.87	38.48	27.35	28.11	1.46	0.21	0.68
0.55	1	93.79	0	56.09	71.80	68.95	121.32	87.90	81.67	3.76	2.73	2.92
0.55	2	97.82	0	58.72	70.43	71.33	310.32	340.76	207.29	2.69	6.28	6.74
0.55	4	99.09	0	60.63	70.97	75.04	731.37	1341.56	473.38	3.86	2.42	7.42
0.67	0	4.52	0	95.08	68.42	71.61	1.05	8.48	3.57	0.74	2.28	0.82
0.67	0.125	21.44	0	66.90	68.62	66.37	5.78	9.84	5.97	0.35	1.43	0.42
0.67	0.25	47.82	0	50.73	68.24	62.96	18.57	14.07	12.37	0.26	0.35	0.71
0.67	0.5	76.37	0	41.20	67.28	61.22	68.69	32.25	35.54	0.56	0.16	1.47
0.67	1	91.57	0	38.91	68.50	63.52	219.38	100.48	98.26	3.37	4.59	5.06
0.67	2	96.99	0	42.03	67.86	66.36	606.36	388.61	262.57	1.62	5.85	5.24
0.67	4	98.72	0	44.81	67.12	69.19	1560.62	1524.42	616.90	13.16	1.97	3.39
0	0	4.56	1	94.75	93.30	94.75	1.08	2.59	1.08	0.49	0.15	0.49
0	0.125	27.10	1	88.37	91.14	88.37	2.36	3.59	2.36	0.02	0.44	0.02
0	0.25	58.91	1	81.55	86.72	81.55	5.87	6.45	5.87	0.86	1.38	0.86
0	0.5	84.30	1	78.08	81.45	78.08	18.31	18.75	18.31	0.56	0.62	0.56
0	1	94.93	1	79.07	78.32	79.07	52.64	65.85	52.64	0.86	2.35	0.86
0	2	98.25	1	82.72	77.77	82.72	126.13	253.76	126.13	2.72	1.07	2.72
0	4	99.30	1	85.48	77.84	85.48	282.82	1004.06	282.82	0.70	2.91	0.70
0.35	0	4.47	1	94.86	87.75	89.82	1.07	3.80	1.51	0.21	0.62	0.16
0.35	0.125	26.29	1	85.24	86.20	84.15	2.79	4.70	2.86	0.02	0.64	0.67
0.35	0.25	57.37	1	73.55	82.42	76.76	8.21	8.00	7.18	0.32	0.45	0.56
0.35	0.5	83.20	1	69.08	78.80	74.24	25.46	20.47	21.24	2.22	0.84	1.28
0.35	1	94.50	1	69.89	76.90	76.52	74.97	70.43	59.38	3.31	2.57	3.71
0.35	2	98.10	1	72.76	75.66	78.78	189.74	280.98	151.30	6.10	7.91	11.29
0.35	4	99.23	1	75.29	75.13	81.90	437.87	1101.58	339.84	1.77	10.60	5.89
0.55	0	4.38	1	95.16	76.51	80.43	1.02	6.35	2.46	0.11	0.33	0.04
0.55	0.125	24.48	1	77.40	74.32	73.79	3.83	7.83	4.46	0.41	1.25	0.58
0.55	0.25	54.67	1	63.96	73.48	67.74	11.72	11.54	10.20	1.09	0.48	0.33
0.55	0.5	81.44	1	56.62	72.52	66.36	39.71	26.72	28.43	0.97	1.80	1.13
0.55	1	93.74	1	56.47	71.20	68.94	117.77	88.20	81.98	1.83	0.51	0.63
0.55	2	97.82	1	57.78	70.55	71.64	313.12	341.51	205.45	5.70	8.77	0.86
0.55	4	99.09	1	61.07	69.86	74.42	728.21	1350.06	477.25	0.69	4.03	7.90
0.67	0	4.51	1	95.09	68.87	71.37	1.04	8.43	3.54	0.07	0.91	0.38
0.67	0.125	21.10	1	66.27	68.58	67.66	5.83	9.83	5.70	0.13	0.18	0.19
0.67	0.25	47.94	1	48.44	68.57	63.19	19.62	14.14	12.51	0.85	0.32	0.52
0.67	0.5	76.21	1	40.30	68.07	60.96	68.73	31.82	35.45	0.78	3.20	0.80
0.67	1	91.54	1	40.02	67.39	62.63	214.42	105.19	102.28	2.05	2.21	3.35
0.67	2	96.99	1	42.09	68.64	66.87	598.93	378.13	251.76	6.29	0.68	3.78
0.67	4	98.70	1	43.91	66.73	69.44	1562.72	1542.71	594.71	2.04	5.04	2.05
Average				**70.68**	**76.37**	**75.68**	**175.89**	**241.39**	**103.68**	**1.81**	**2.04**	**1.95**

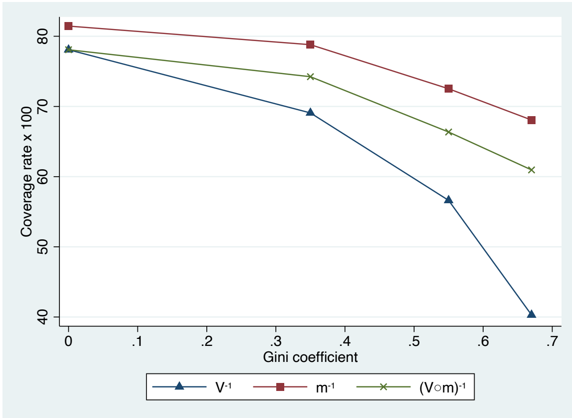
Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. $\boldsymbol{V}^{-1}$ is the meta-regression with inverse variance weighting, $\boldsymbol{m}^{-1}$ is the weighting by the inverse number of estimates per study, and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Figure 37. Coverage, MSE, Bias ($d = 0$, $f = 0$, Random Heterogeneity at Study Level)

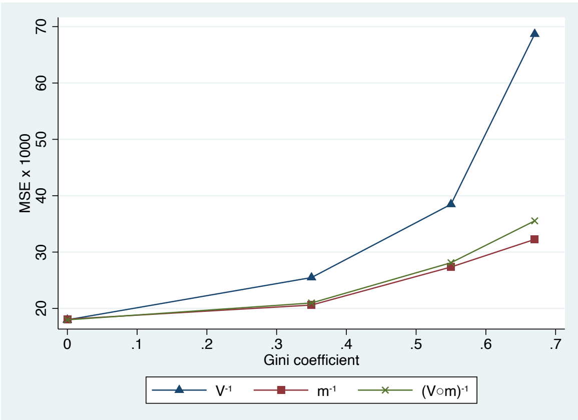
(a) Coverage; $\alpha_1 = 0$



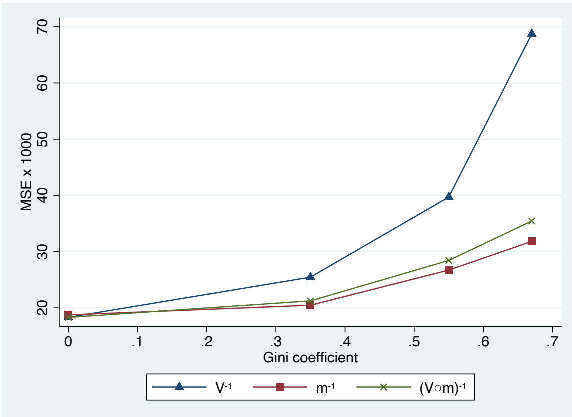
(b) Coverage; $\alpha_1 = 1$



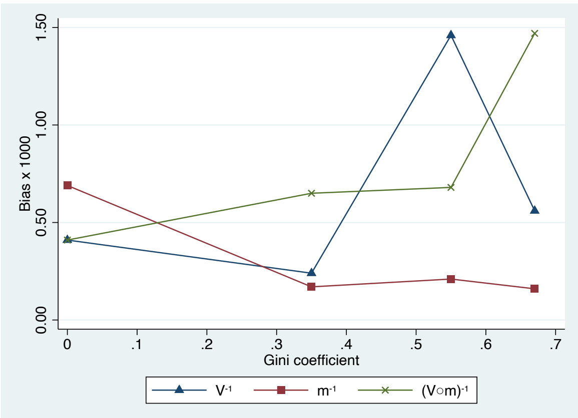
(c) MSE; $\alpha_1 = 0$



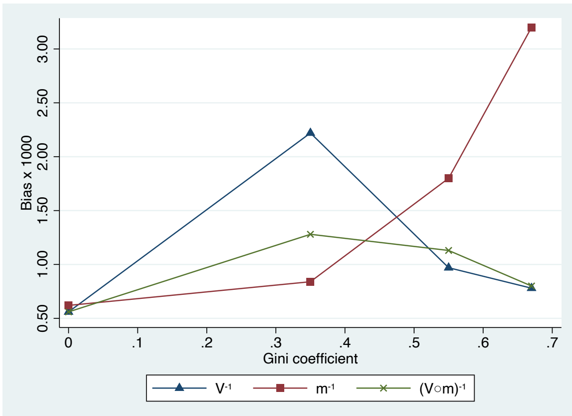
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: These graphs show the coverage, MSE, and bias reported in Table 23 with alternating levels of unbalancedness, measured by the Gini coefficient, plotted on the horizonal axis and the true effect being either $\alpha_1 = 0$ or $\alpha_1 = 1$. Graphs (a) and (b) show the coverage rates. Graphs (b) and (d) show the mean squared errors. Graphs (e) and (f) plot the bias. Random heterogeneity is parametrized at $\sigma_g = 0.5$.

***Sample Overlap* $\boldsymbol{d = 0}$.** According to Table 26, the average coverage ratios are lower for all three weighting schemes, while the MSE is much higher as compared to the previous case when random heterogeneity was effect size-specific (Table 23). The drop in the performance is most pronounced for $\boldsymbol{V}^{-1}$, with an average coverage of 70.68% compared to 92.97% in the case of effect size-level heterogeneity. As compared to $\boldsymbol{V}^{-1}$, the CIs for $\boldsymbol{m}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ are on average 5.69 and 5.00 percentage points closer to the nominal level of 95%. The average MSE is the lowest for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ and highest for $\boldsymbol{m}^{-1}$, making $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ the most efficient estimator when random heterogeneity is study level-specific.

Similar to the previous case, where random heterogeneity is defined at the effect size level, the largest deviation of $\boldsymbol{V}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ from the nominal can be observed for medium random heterogeneity, $\sigma_g = 0.5$. In contrast, the coverage of $\boldsymbol{m}^{-1}$ is lowest when random heterogeneity is at the highest level, $\sigma_g = 4$. The MSE is also increasing with higher levels of random heterogeneity. While the performance of $\boldsymbol{V}^{-1}$ in terms of coverage rates and MSE is independent of the level of G when random heterogeneity was defined at the effect size level (Table 23), the results for study-level heterogeneity are dependent on the distribution of effect sizes across studies. The higher the level of G, the lower the coverage rates of $\boldsymbol{V}^{-1}$. For example, when $G = 0$, the coverage is 94.97% for $\alpha_1 = 0$ and $\sigma_g = 0$. In contrast, the coverage drops to only 38.91% when $G = 0.67$, $\alpha_1 = 0$, and $\sigma_g = 1$.

Figure 37 presents the statistical properties of the three weighting schemes with changing levels of G, while keeping random heterogeneity at a medium level, $\sigma_g = 0.5$. In this particular case with medium random heterogeneity, $\boldsymbol{m}^{-1}$ outperforms the other weights as it has the highest coverage and lowest MSE for all levels of the Gini coefficient.

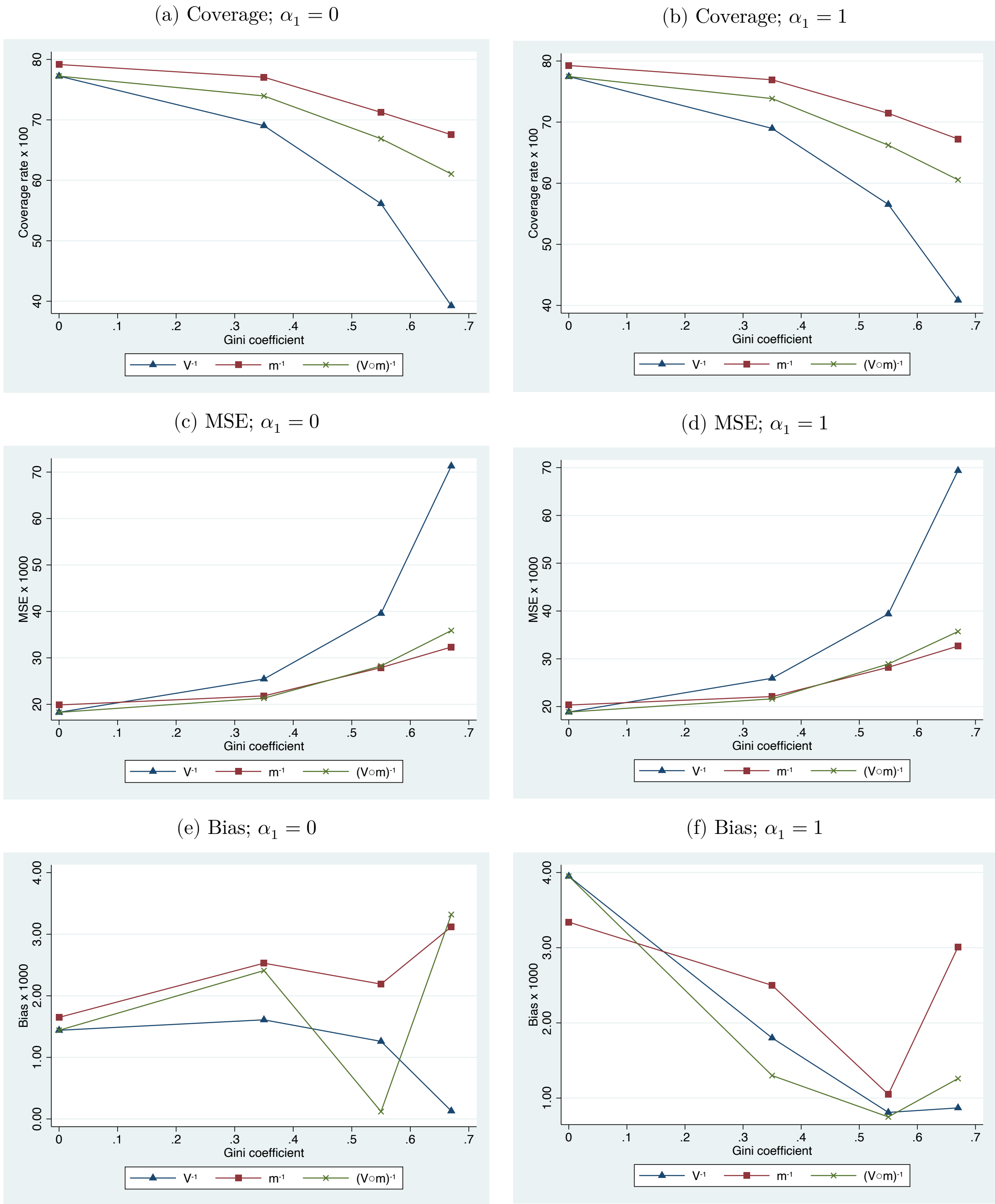
Sample Overlap $\boldsymbol{d = \{0.25, 0.5, 0.95\}}$. The previous simulation in Table 26 assumed that there is no overlap in the randomly generated data samples, neither within nor between studies. Appendices C.3 and C.4 show the results for sample overlap with $d = 0.25$ and $d = 0.5$. The findings for the highest level of overlap, $d = 0.95$, are reported in Table 27. Comparing the outcomes with the previous case without sample overlap shows that the average coverage ratios are slightly lower for $\boldsymbol{V}^{-1}$ and $\boldsymbol{m}^{-1}$. The average coverage is 67.98% for $\boldsymbol{V}^{-1}$, 74.24% for $\boldsymbol{m}^{-1}$, and 74.62% for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$, as compared to 70.68%, 76.37% and 75.68%, respectively, when $d =$ 0 (Table 26). The MSE is almost the same and the bias increases slightly for all three weighting schemes. The effect of overlapping samples on $\boldsymbol{V}^{-1}$ is the highest when $\sigma_g = 0$. Previously, the coverages of $\boldsymbol{V}^{-1}$ almost coincided with the nominal (95%) for all levels of G, when $d = 0$ and $\sigma_g = 0$ (Table 26). But when $d = 0.95$ and $\sigma_g = 0$ (Table 27), the coverage ratios decrease with larger values of G. For example, the coverage of $\boldsymbol{V}^{-1}$ is 90.54% when $G = 0$ and $\alpha_1 = 0$ as compared to 94.97% without sample overlap. For $G = 0.67$, $\sigma_g = 0$, and $\alpha_1 = 0$, the difference

Table 27. Simulation Results ($d = 0.95$, $f = 0$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.64	0	90.54	85.05	90.54	1.45	4.11	1.45	0.14	0.85	0.14
0	0.125	26.57	0	84.50	83.20	84.50	2.78	5.30	2.78	0.31	0.24	0.31
0	0.25	58.45	0	80.58	82.65	80.58	6.23	7.98	6.23	0.50	1.03	0.50
0	0.5	83.94	0	77.24	79.16	77.24	18.27	19.86	18.27	1.44	1.65	1.44
0	1	94.86	0	79.36	77.89	79.36	52.40	67.73	52.40	1.98	2.07	1.98
0	2	98.24	0	82.61	76.88	82.61	128.65	255.40	128.65	0.19	5.81	0.19
0	4	99.30	0	85.67	76.87	85.67	285.61	1026.04	285.61	4.46	2.70	4.46
0.35	0	4.59	0	88.14	81.08	85.14	1.67	5.14	1.96	0.31	0.04	0.66
0.35	0.125	25.15	0	79.28	80.20	79.73	3.54	6.15	3.43	0.30	0.32	0.19
0.35	0.25	56.33	0	72.13	78.23	75.97	8.46	9.39	7.41	0.92	0.29	0.39
0.35	0.5	82.78	0	69.04	77.06	73.96	25.45	21.81	21.32	1.61	2.53	2.41
0.35	1	94.42	0	70.19	76.22	75.72	74.32	72.21	60.36	3.94	1.41	1.59
0.35	2	98.09	0	72.90	76.06	79.01	191.26	276.72	152.40	1.00	1.05	6.16
0.35	4	99.23	0	76.01	75.61	82.89	433.11	1094.66	336.33	3.31	7.46	2.31
0.55	0	4.28	0	84.55	72.43	76.69	2.00	7.20	2.79	0.42	0.26	0.59
0.55	0.125	23.32	0	72.03	72.20	72.46	4.67	8.58	4.65	0.39	1.30	0.50
0.55	0.25	53.05	0	61.20	72.26	67.83	12.50	11.97	10.02	0.86	0.40	1.01
0.55	0.5	80.70	0	56.14	71.25	66.89	39.57	27.90	28.26	1.26	2.19	0.12
0.55	1	93.62	0	56.49	71.53	68.27	119.30	89.04	81.56	3.54	6.89	4.67
0.55	2	97.77	0	57.36	70.79	71.40	311.17	336.21	204.23	0.02	2.70	2.35
0.55	4	99.08	0	61.15	70.52	74.86	724.97	1356.69	471.27	9.67	6.05	8.95
0.67	0	3.90	0	77.11	67.04	70.98	2.81	8.83	3.62	0.49	0.01	0.07
0.67	0.125	19.05	0	59.34	66.88	66.08	7.54	10.48	6.04	0.43	1.03	0.90
0.67	0.25	45.92	0	46.17	68.14	62.33	21.15	14.46	12.60	2.33	1.82	0.89
0.67	0.5	75.06	0	39.26	67.55	61.06	71.30	32.33	35.88	0.13	3.12	3.32
0.67	1	91.25	0	38.74	67.49	62.14	223.14	103.44	101.35	1.26	1.69	3.72
0.67	2	96.91	0	41.89	67.24	66.01	598.98	385.94	253.68	9.63	9.96	2.59
0.67	4	98.69	0	43.89	67.18	69.20	1571.85	1543.32	617.31	1.49	1.01	6.56
0	0	4.56	1	89.64	84.21	89.64	1.49	4.27	1.49	0.31	0.36	0.31
0	0.125	26.62	1	84.93	83.55	84.93	2.73	5.24	2.73	0.13	0.16	0.13
0	0.25	58.21	1	79.63	81.64	79.63	6.39	8.15	6.39	0.09	0.75	0.09
0	0.5	84.00	1	77.46	79.27	77.46	18.86	20.33	18.86	3.95	3.34	3.95
0	1	94.84	1	78.57	77.22	78.57	53.68	69.10	53.68	0.20	1.08	0.20
0	2	98.24	1	82.94	77.91	82.94	128.07	252.57	128.07	2.44	4.57	2.44
0	4	99.30	1	86.00	77.50	86.00	284.17	1013.93	284.17	1.86	1.10	1.86
0.35	0	4.46	1	87.97	80.44	84.85	1.67	5.11	1.95	0.48	2.40	0.95
0.35	0.125	25.15	1	78.91	80.00	79.72	3.53	6.30	3.45	0.27	0.74	0.53
0.35	0.25	56.59	1	72.44	79.01	76.20	8.46	9.35	7.41	0.94	0.45	0.59
0.35	0.5	83.07	1	68.97	76.93	73.85	25.93	22.10	21.62	1.80	2.50	1.30
0.35	1	94.51	1	69.74	75.15	75.12	76.31	74.22	61.91	0.91	0.94	2.19
0.35	2	98.08	1	73.28	75.69	79.23	186.62	278.52	147.97	2.83	2.02	4.50
0.35	4	99.23	1	75.60	75.91	82.41	445.14	1104.32	343.49	1.09	1.30	0.97
0.55	0	4.38	1	84.66	72.42	76.45	2.00	7.11	2.82	0.08	0.59	0.17
0.55	0.125	23.22	1	72.06	72.94	72.80	4.67	8.32	4.55	0.76	0.22	0.43
0.55	0.25	53.19	1	61.80	72.22	68.96	12.40	12.36	9.96	0.72	0.85	0.64
0.55	0.5	80.69	1	56.53	71.44	66.23	39.39	28.22	28.91	0.81	1.05	0.75
0.55	1	93.62	1	55.79	71.66	67.80	119.96	88.93	82.48	2.54	1.44	1.36
0.55	2	97.79	1	58.53	71.21	72.25	308.84	330.72	198.58	4.03	6.86	0.49
0.55	4	99.09	1	61.54	71.00	75.75	715.96	1329.77	459.60	1.61	14.74	3.05
0.67	0	3.91	1	77.06	67.19	70.25	2.79	8.96	3.66	0.23	0.51	0.26
0.67	0.125	19.42	1	57.91	67.20	66.54	7.74	10.24	5.90	0.73	0.13	0.63
0.67	0.25	45.67	1	45.91	67.75	62.68	21.23	14.81	12.70	1.73	1.81	0.52
0.67	0.5	75.10	1	40.86	67.24	60.56	69.39	32.66	35.71	0.87	3.01	1.26
0.67	1	91.17	1	38.34	67.45	63.57	223.61	102.09	99.03	6.56	0.50	2.34
0.67	2	96.91	1	42.07	67.59	66.09	611.28	379.91	253.93	13.20	9.03	8.79
0.67	4	98.69	1	44.40	67.22	69.10	1540.42	1521.03	608.03	14.69	18.82	16.76
Average				**67.98**	**74.24**	**74.62**	**176.19**	**242.10**	**103.59**	**2.11**	**2.63**	**2.08**

Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. $\boldsymbol{V}^{-1}$ is the meta-regression with inverse variance weighting, $\boldsymbol{m}^{-1}$ is the weighting by the inverse number of estimates per study, and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Figure 38. Coverage, MSE, Bias ($d = 0.95$, $f = 0$, Random Heterogeneity at Study Level)



Notes: These graphs show the coverage, MSE, and bias reported in Table 23 with alternating levels of unbalancedness, measured by the Gini coefficient, plotted on the horizonal axis and the true effect being either $\alpha_1 = 0$ or $\alpha_1 = 1$. Graphs (a) and (b) show the coverage rates. Graphs (b) and (d) show the mean squared errors. Graphs (e) and (f) plot the bias. Random heterogeneity is parametrized at $\sigma_g = 0.5$.

in the coverage ratios between the results for $d = 0.95$ and $d = 0$ is even greater: 77.11% when $d = 0.95$ and 95.08% when $d = 0$. In this particular case, the CIs of $\boldsymbol{V}^{-1}$, are on average 17.97

percentage points closer to the nominal level of 95% when there is no sample overlap. Interestingly, when $G = 0$, there is a spread in the coverage ratios of $\boldsymbol{m}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ when σ_g increases from lower to higher levels. For example, when $G = 0$, $\sigma_g = 0$, and $\alpha_1 = 0$, the coverage of $\boldsymbol{m}^{-1}$ is 85.05% and goes down to 76.87% when $\sigma_g = 4$. However, when $G = 0.67$, this spread in the coverage almost disappears. For example, when $\alpha_1 = 0$, the coverage ratio of $\boldsymbol{m}^{-1}$ is 67.04% when $\sigma_g = 0$ and $G = 0.67$, whereas the coverage is 67.18% when $\sigma_g = 4$.

Figure 38 illustrates the statistical properties with changing levels of G, while holding the random heterogeneity at a medium level, $\sigma_g = 0.5$. The coverage rates are the lowest for $\boldsymbol{V}^{-1}$ and highest for $\boldsymbol{m}^{-1}$. The MSE is lowest for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ and $\boldsymbol{m}^{-1}$, making the two alternatives the more efficient weights in this specific scenario. In summary, holding random heterogeneity constant at the study level, which implies a fixed effects panel structure in the meta-regression sample, has a strong impact on the performance of inverse variance weighting. While the three weighting schemes all show a lower relative performance as compared to the previous case of effect size-level random heterogeneity (Section 4.5.1), $\boldsymbol{V}^{-1}$ is affected the most. In the extreme scenario of maximum random heterogeneity and maximum unbalancedness of the number of effect size estimates reported per study, the difference in coverage ratios between $\boldsymbol{V}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is more than 20 percentage points. Given the lowest values for the MSE and bias, $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the preferred weight when random heterogeneity is constant at the study level.

4.5.3. Publication Selection

So far, the simulation has not considered selective reporting of statistically significant results. However, in practical applications, such as in the meta-analysis of the impact of hedging on firm value in Chapter 3 (Table 11), meta-researchers routinely find evidence for publication selection (among many others, Anderson et al., 2018: 66; Askarov and Doucouliagos, 2013: 613; Doucouliagos and Stanley, 2013: 320-322; Doucouliagos et al., 2012: 198-201; Doucouliagos and Laroche, 2009: 161; Doucouliagos and Stanley, 2009: 411-415; Efendic et al., 2011: 592; Feld et al., 2013: 2855; Fidrmuc and Lind, 2020: 5; Hang et al., 2018: 220; Havranek et al., 2016: 140; Havranek et al., 2012: 205; Havranek et al., 2015b: 401; Lichter et al., 2015: 105; Stanley and Doucouliagos, 2012: 51-53; Zigraiova et al., 2021: 10). If unaccounted for, publication selection can bias the meta-analysis results (Ioannidis et al., 2017: F241-F242; Stanley and Doucouliagos, 2010: 182-183).

In this section, the simulation design incorporates publication selection, which is simulated by selecting those effect size estimates in the primary regression that have a significantly positive result at the 5% level. If an effect size estimate is not significant, the original primary regression (Eq. 29a or 29b) is re-run by drawing a new set of random variables. This procedure continues until a significant result is obtained by chance. The decision of whether an effect size estimate is prone to publication selection is randomly assigned with a probability f. Following Stanley (2008: 127), the probability of publication selection is calibrated at three levels, $f = \{0.25, 0.5, 0.75\}$. For example, if $f = 0.5$, a new sample is drawn for 50% of the effect size estimates until a random but statistically significant result is found. In the other 50% of the cases, the first random estimate is reported, significant or not. 50% is selected as a 'moderate' value for publication selection, 25% is classified as 'light', and 75% is an 'extreme' level of publication selection (Stanley, 2008: 127). Meta-regression analysis provides the ability to detect and control publication selection. Therefore, the meta-regressions models with the three alternative weights are estimated by including the standard error of the effect size estimate as a second moderator (Eq. 32). In this model, the constant of the meta-regression, $\hat{\beta}_0$, is the estimate for the genuine empirical effect beyond publication selection.

Publication Selection $f = 0.5$ and Random Heterogeneity at Study Level. As in the previous section, random heterogeneity is study level-specific. Table 28 reports the results when 50% of the estimates are reported only when they are positive and statistically significant. The findings for $f = 0.25$ are reported in Appendix C.9. The average coverage ratios are rather similar for all three weighting schemes, with values ranging from 73.86% for $\boldsymbol{V}^{-1}$ to 76.86% for $\boldsymbol{m}^{-1}$. Although the meta-regression model accounts for the effects of publication selection by including the standard error of the effect size estimates as a moderator, all three weights suffer from publication selection, resulting in greater bias and lower efficiency. As compared to the case without publication selection, the overall average MSE across the three weights is about 1.65 times higher, while the bias is 47.74 times higher. $\boldsymbol{m}^{-1}$ has the highest average bias and the highest MSE among the three weighting schemes. For the other two weights, the results indicate that $\boldsymbol{V}^{-1}$ is more biased and less efficient (higher MSE) than $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. On average, $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ dominates the other two weights when there is publication selection. In only a few cases when there is a true effect with no or medium random heterogeneity, $\boldsymbol{V}^{-1}$ is slightly less biased than $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. The improvement of $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ over the other two weights is greatest when both σ_g and G are high. For example, when $G = 0.67$, $\sigma_g = 4$, and $\alpha_1 = 0$, the MSE is about 2.5 times larger for $\boldsymbol{V}^{-1}$ as compared to $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$, while the bias is more than twice as large.[70]

[70] A larger value of G implies a higher likelihood that a study with many estimates reports both estimates that are not affected by publication selection and estimates that are actively selected based on their statistical significance.

Table 28. Simulation Results ($d = 0.5$, $f = 0.5$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.46	0	87.44	91.42	87.44	7.60	14.69	7.60	45.71	66.28	45.71
0	0.125	28.03	0	67.73	79.69	67.73	16.06	26.81	16.06	101.44	124.38	101.44
0	0.25	60.89	0	70.57	76.48	70.57	24.55	38.21	24.55	118.71	150.78	118.71
0	0.5	85.54	0	72.91	73.82	72.91	58.92	76.95	58.92	176.01	211.61	176.01
0	1	95.41	0	78.79	77.06	78.79	137.50	191.87	137.50	249.08	306.59	249.08
0	2	98.45	0	86.67	82.32	86.67	254.63	501.65	254.63	244.25	367.32	244.25
0	4	99.40	0	88.29	81.97	88.29	496.80	1579.08	496.80	173.37	569.50	173.37
0.35	0	4.38	0	84.77	87.12	81.39	9.48	19.04	9.89	54.04	70.49	49.69
0.35	0.125	28.27	0	65.86	76.25	65.53	18.35	30.07	17.56	102.60	121.37	97.66
0.35	0.25	60.83	0	68.79	75.07	69.50	28.34	41.30	26.87	117.51	141.41	111.02
0.35	0.5	85.56	0	69.18	71.49	69.82	68.96	84.02	65.67	173.35	203.16	163.60
0.35	1	95.41	0	73.80	74.94	75.82	166.43	211.67	156.29	241.44	292.61	232.52
0.35	2	98.44	0	79.67	78.27	81.84	349.15	602.67	311.13	237.82	373.43	231.58
0.35	4	99.40	0	80.07	78.59	84.33	714.40	1806.33	584.54	188.68	600.87	165.06
0.55	0	4.51	0	77.80	78.19	72.27	14.88	26.73	13.82	65.23	65.93	48.72
0.55	0.125	28.17	0	61.85	71.22	63.29	23.62	35.62	20.40	105.98	108.59	86.12
0.55	0.25	60.89	0	61.42	69.68	63.38	39.11	49.06	34.17	123.13	124.92	99.91
0.55	0.5	85.56	0	59.90	66.22	62.36	97.59	100.24	87.21	186.75	184.30	154.47
0.55	1	95.41	0	64.19	66.94	67.04	243.57	280.36	217.06	259.57	284.73	226.76
0.55	2	98.44	0	67.89	70.69	74.24	546.14	779.60	430.48	256.95	347.06	226.73
0.55	4	99.40	0	69.27	71.80	77.99	1079.58	2240.81	796.19	223.61	541.02	176.35
0.67	0	4.50	0	70.67	71.96	64.10	27.07	32.17	17.58	82.25	59.35	46.57
0.67	0.125	28.06	0	55.35	67.26	59.21	37.88	41.01	23.76	112.40	95.10	73.60
0.67	0.25	60.96	0	51.93	65.34	58.59	67.25	55.24	40.61	141.22	112.66	89.86
0.67	0.5	85.55	0	48.03	62.06	57.80	164.53	116.53	105.49	217.99	169.03	143.03
0.67	1	95.38	0	49.86	62.18	61.05	422.38	326.28	268.83	302.89	261.26	215.05
0.67	2	98.44	0	49.97	65.07	67.75	1008.36	931.41	557.70	355.32	338.95	225.39
0.67	4	99.39	0	50.17	66.76	71.93	2191.03	2696.84	1071.65	371.61	512.80	168.58
0	0	4.44	1	91.62	92.92	91.62	4.03	7.76	4.03	23.66	37.52	23.66
0	0.125	28.17	1	90.00	92.27	90.00	6.46	9.44	6.46	22.49	35.62	22.49
0	0.25	60.80	1	87.36	90.68	87.36	13.42	14.64	13.42	21.61	34.40	21.61
0	0.5	85.59	1	88.23	90.58	88.23	32.08	31.03	32.08	1.68	1.39	1.68
0	1	95.41	1	90.19	89.76	90.19	76.88	98.05	76.88	48.59	72.30	48.59
0	2	98.45	1	90.94	84.82	90.94	183.81	407.07	183.81	45.48	157.54	45.48
0	4	99.40	1	89.15	83.08	89.15	396.53	1334.74	396.53	27.38	350.98	27.38
0.35	0	4.50	1	90.95	88.75	87.38	4.39	10.13	5.26	21.14	35.10	22.49
0.35	0.125	28.06	1	86.88	87.95	84.73	7.39	12.27	8.46	19.35	33.52	20.62
0.35	0.25	60.83	1	83.88	86.50	81.75	15.71	19.14	17.55	13.40	31.28	19.02
0.35	0.5	85.58	1	83.56	86.41	81.61	38.77	40.67	43.30	14.35	1.70	2.07
0.35	1	95.40	1	85.61	85.11	85.22	100.33	125.15	101.35	62.61	70.30	43.40
0.35	2	98.44	1	85.72	81.61	87.08	248.02	463.23	226.06	46.57	117.22	20.95
0.35	4	99.40	1	83.12	81.00	86.09	553.97	1449.34	471.22	5.53	317.76	40.68
0.55	0	4.39	1	86.61	78.98	76.88	5.88	15.90	8.43	17.00	34.76	22.18
0.55	0.125	28.08	1	80.22	77.59	72.96	10.14	19.57	13.80	15.71	32.62	21.01
0.55	0.25	60.91	1	77.19	77.44	71.87	20.49	28.86	27.10	7.20	32.32	21.54
0.55	0.5	85.56	1	76.15	77.10	71.80	55.22	63.30	68.77	27.42	13.04	8.19
0.55	1	95.41	1	77.83	75.64	75.09	143.82	191.99	162.88	82.14	44.18	30.51
0.55	2	98.45	1	77.04	73.83	78.75	361.79	638.22	337.15	89.43	114.38	22.51
0.55	4	99.40	1	72.45	73.58	78.51	851.72	1961.50	691.09	61.12	300.85	48.65
0.67	0	4.44	1	78.75	71.43	68.83	9.16	20.82	11.71	13.64	33.39	20.18
0.67	0.125	28.08	1	70.02	70.91	66.07	14.98	25.98	18.86	10.01	34.01	23.14
0.67	0.25	60.86	1	62.21	69.74	63.54	32.88	39.65	38.39	4.10	33.79	22.87
0.67	0.5	85.50	1	61.10	70.12	63.67	87.89	82.66	94.49	61.45	9.98	3.33
0.67	1	95.39	1	60.39	69.81	67.13	246.61	246.11	225.89	137.28	32.74	26.56
0.67	2	98.44	1	60.31	67.98	71.43	679.45	797.36	465.76	197.09	75.71	2.98
0.67	4	99.39	1	56.02	68.58	71.51	1650.23	2364.44	948.87	195.29	214.67	85.52
Average				**73.86**	**76.86**	**75.38**	**252.97**	**418.84**	**188.44**	**112.89**	**162.65**	**86.79**

Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. $\boldsymbol{V}^{-1}$ is the meta-regression with inverse variance weighting, $\boldsymbol{m}^{-1}$ is the weighting by the inverse number of estimates per study, and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Table 29. Simulation Results ($d = 0.5$, $f = 0.75$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.46	0	80.23	88.35	80.23	7.73	14.12	7.73	61.49	77.92	61.49
0	0.125	28.03	0	57.70	70.71	57.70	14.42	26.57	14.42	104.49	132.65	104.49
0	0.25	60.89	0	55.89	62.37	55.89	26.38	42.53	26.38	139.72	174.46	139.72
0	0.5	85.54	0	50.00	49.03	50.00	78.64	105.61	78.64	242.51	286.31	242.51
0	1	95.41	0	57.45	50.84	57.45	194.11	277.68	194.11	368.89	455.07	368.89
0	2	98.45	0	74.15	65.80	74.15	302.95	610.58	302.95	392.74	599.47	392.74
0	4	99.40	0	84.65	70.87	84.65	437.46	1680.75	437.46	308.86	918.08	308.86
0.35	0	4.38	0	75.41	82.07	72.74	10.83	18.64	10.01	71.94	83.30	66.18
0.35	0.125	28.27	0	55.90	69.01	56.64	17.34	29.67	16.16	106.82	130.80	103.54
0.35	0.25	60.83	0	53.09	62.21	55.78	32.00	44.78	28.64	142.48	167.33	134.68
0.35	0.5	85.56	0	47.39	51.90	53.01	93.78	106.53	81.13	244.53	270.81	225.67
0.35	1	95.41	0	52.33	53.20	59.58	232.57	279.69	197.31	370.77	427.21	338.66
0.35	2	98.44	0	66.89	65.65	73.51	395.07	639.45	324.07	391.99	563.86	357.04
0.35	4	99.40	0	76.53	69.47	81.54	633.49	1771.45	517.46	339.18	869.90	281.03
0.55	0	4.51	0	70.46	75.01	64.79	16.12	25.50	13.54	80.61	78.71	63.30
0.55	0.125	28.17	0	53.66	67.53	57.30	22.10	33.95	18.54	107.28	116.23	94.99
0.55	0.25	60.89	0	47.01	62.80	56.27	44.57	49.75	34.73	152.17	149.40	125.55
0.55	0.5	85.56	0	41.60	54.50	53.27	128.32	112.84	94.87	263.28	242.63	212.50
0.55	1	95.41	0	43.35	55.11	56.79	321.81	311.61	238.57	401.86	396.26	324.89
0.55	2	98.44	0	56.04	63.03	66.88	580.98	792.31	446.75	449.31	539.85	352.41
0.55	4	99.40	0	65.95	67.44	74.94	939.68	2053.18	711.53	416.64	780.04	274.19
0.67	0	4.50	0	70.46	75.01	64.79	16.12	25.50	13.54	80.61	78.71	63.30
0.67	0.125	28.06	0	41.60	64.61	56.51	38.22	36.84	21.49	117.20	102.81	85.48
0.67	0.25	60.96	0	35.09	59.54	54.78	75.94	53.02	38.67	171.11	130.54	108.81
0.67	0.5	85.55	0	27.37	53.52	51.80	226.26	120.04	105.86	304.99	218.49	188.74
0.67	1	95.38	0	28.26	53.29	54.04	556.66	340.53	280.38	455.53	352.25	289.87
0.67	2	98.44	0	36.31	59.70	62.74	1070.10	889.07	540.70	544.71	476.58	312.10
0.67	4	99.39	0	44.88	62.75	70.66	1897.33	2433.00	917.48	589.33	707.31	231.85
0	0	4.44	1	89.14	89.23	89.14	4.58	9.20	4.58	35.78	58.76	35.78
0	0.125	28.17	1	87.51	89.09	87.51	7.15	10.84	7.15	34.61	57.20	34.61
0	0.25	60.80	1	86.15	89.27	86.15	13.48	15.18	13.48	28.19	50.02	28.19
0	0.5	85.59	1	87.17	90.44	87.17	31.07	28.30	31.07	1.92	6.38	1.92
0	1	95.41	1	89.12	88.72	89.12	70.37	87.33	70.37	58.53	90.47	58.53
0	2	98.45	1	90.51	86.12	90.51	147.18	306.87	147.18	33.55	176.92	33.55
0	4	99.40	1	87.36	83.46	87.36	314.16	1054.24	314.16	105.73	455.23	105.73
0.35	0	4.50	1	87.77	84.82	84.81	4.96	11.24	5.74	34.25	55.45	33.41
0.35	0.125	28.06	1	85.37	84.69	82.63	7.87	13.42	9.00	30.81	52.98	32.77
0.35	0.25	60.83	1	82.54	84.69	80.63	15.91	19.06	17.79	23.96	49.52	31.30
0.35	0.5	85.58	1	83.58	86.03	81.60	37.94	37.27	41.31	13.28	11.47	4.44
0.35	1	95.40	1	83.85	83.99	84.08	97.05	111.76	95.64	76.53	79.04	46.61
0.35	2	98.44	1	84.48	82.67	85.88	210.85	359.21	195.27	50.77	143.57	14.97
0.35	4	99.40	1	81.59	81.49	83.46	437.39	1128.23	395.10	57.94	403.03	117.75
0.55	0	4.39	1	84.46	76.59	74.78	6.20	16.74	8.94	31.13	50.93	32.19
0.55	0.125	28.08	1	79.37	76.49	72.72	10.39	19.50	14.01	25.20	50.18	29.49
0.55	0.25	60.91	1	76.12	76.16	70.74	20.39	28.92	27.67	13.48	47.74	29.28
0.55	0.5	85.56	1	75.30	77.20	71.97	52.94	58.38	64.71	38.38	16.63	3.80
0.55	1	95.41	1	75.44	76.49	74.56	138.38	164.25	147.15	121.78	67.49	47.45
0.55	2	98.45	1	74.98	75.43	77.59	324.56	515.16	300.01	119.10	126.23	18.05
0.55	4	99.40	1	74.68	75.73	76.95	623.60	1499.89	565.12	37.82	339.02	117.14
0.67	0	4.44	1	76.37	69.84	68.09	8.91	21.62	12.24	22.63	49.53	30.54
0.67	0.125	28.08	1	68.12	69.98	65.38	16.26	25.91	18.84	12.34	46.63	28.20
0.67	0.25	60.86	1	60.85	69.43	63.00	34.41	38.41	37.59	10.43	42.75	28.00
0.67	0.5	85.50	1	57.95	70.30	63.14	94.12	76.88	89.60	82.80	18.79	8.10
0.67	1	95.39	1	55.93	70.17	66.95	267.64	217.05	204.57	198.18	39.40	33.34
0.67	2	98.44	1	55.67	68.58	70.76	616.03	657.62	413.95	238.92	91.00	1.12
0.67	4	99.39	1	56.84	70.34	71.21	1244.29	1835.77	789.00	232.54	268.90	126.84
Average				**66.24**	**71.54**	**70.37**	**237.23**	**380.33**	**174.25**	**164.75**	**222.69**	**124.41**

Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. V^{-1} is the meta-regression with inverse variance weighting, m^{-1} is the weighting by the inverse number of estimates per study, and $(V \circ m)^{-1}$ is the combined weight. The columns report the coverage proportions, mean squared errors, and bias of the estimates. The three measures are averaged across the 10,000 replications. d is the level of sample overlap. Random heterogeneity, α_3, is defined at the effect size level, and f is the level of publication selection.

Publication Selection $f = 0.75$, Random Heterogeneity at Study Level. Table 29 reports the simulation results when 75% of the estimates in the primary study regression are re-estimated until they are statistically significant and positive. Comparing the results of Table 28 (50% publication selection) with Table 29 (75% publication selection) reveals that the relative performance of the weights is not dependent on the degree of publication selection. Although bias increases when there is 75% selection for statistical significance, $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is still, on average, the superior weight, with the strongest relative performance gains against the other weights, especially when both G and σ_g are high.

4.5.4. Cluster-Robust Standard Errors

Defining random heterogeneity at the study level and introducing sample overlap within studies leads to correlation among the estimates reported in the same study. This correlation violates the Gauss-Markov assumption (5), which requires independently and identically distributed errors (Section 4.3). The dependence of the errors affects the efficiency of the meta-regression estimates. Thus, if the meta-researcher does not account for the correct error structure, the standard errors and *t*-values of the meta-regression results may be miscalculated, leading to the false appearance of statistical significance (Stanley and Doucouliagos, 2012: 71).

The common approaches to accounting for dependent errors in meta-regression are panel or multilevel methods or, alternatively, cluster-robust standard errors (Stanley and Doucouliagos, 2012: 68-72). When cluster-robust errors are used, each study is treated as a cluster with a specific within-cluster dependence. In this section, cluster-robust standard errors defined at the level of the individual studies are used to calculate the *t*-statistics of the meta-regression coefficients. As cluster-robust standard errors do not affect the effect size estimates, MSE and bias are not influenced by changes in the standard errors; only the coverage ratios might change due to the adjusted confidence bands using the cluster-robust standard errors.

Table 29 shows the coverage ratios when heterogeneity is defined at the effect size level (Columns 5–7) or at the study level (Columns 8–10). The overlap factor d is set to a medium level of 0.5. Comparing the findings with those for $d = 0.5$ when standard errors are not corrected (Appendices C.3 and C.7), reveals a slight drop in the coverage ratio of $\boldsymbol{V}^{-1}$ and an increase for $\boldsymbol{m}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. When α_3 is defined at the effect size level, the average coverage of $\boldsymbol{V}^{-1}$ is 4.87 precentage points closer to the nominal when the meta-regression is estimated without cluster-robust errors. However, the coverage ratios of $\boldsymbol{m}^{-1}$ and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ are much closer to the nominal when estimated with clustered errors: 11.20 precentage points for $\boldsymbol{m}^{-1}$ and 10.34 precentage points for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$.

Table 30. Simulation Results ($d = 0.5$, $f = 0$, Clustered Standard Errors)

Simulation				α_3 at Effect Size Level			α_3 at Study Level		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.50	0	93.75	94.51	93.75	94.40	94.33	94.40
0	0.125	24.19	0	93.73	94.94	93.73	93.84	94.99	93.84
0	0.25	54.16	0	93.27	94.41	93.27	93.70	94.83	93.70
0	0.5	80.25	0	93.24	94.64	93.24	92.97	94.62	92.97
0	1	92.47	0	93.10	94.66	93.10	93.52	94.66	93.52
0	2	96.86	0	92.20	94.29	92.20	93.19	94.62	93.19
0	4	98.38	0	92.82	94.87	93.60	92.07	94.89	92.07
0.35	0	4.47	0	92.53	94.41	93.33	92.17	94.58	92.89
0.35	0.125	23.77	0	92.14	94.56	92.65	92.49	94.18	93.10
0.35	0.25	53.85	0	91.77	94.47	92.40	92.26	94.39	93.16
0.35	0.5	80.13	0	91.73	93.99	92.61	91.74	94.38	92.54
0.35	1	92.44	0	91.27	94.75	91.96	91.51	94.39	92.09
0.35	2	96.83	0	90.09	94.61	91.04	91.54	94.53	91.57
0.35	4	98.37	0	89.56	94.25	91.95	90.05	94.54	90.57
0.55	0	4.34	0	88.75	93.79	91.40	89.57	94.05	91.75
0.55	0.125	23.42	0	88.80	93.94	91.12	88.64	94.11	91.25
0.55	0.25	53.69	0	88.56	93.79	90.47	88.40	93.64	91.08
0.55	0.5	80.19	0	87.88	93.48	90.06	88.73	93.73	90.45
0.55	1	92.46	0	87.70	93.80	89.65	88.34	93.62	90.37
0.55	2	96.86	0	86.31	93.59	88.87	87.74	93.74	89.29
0.55	4	98.38	0	80.97	93.27	90.15	86.12	93.24	88.05
0.67	0	3.89	0	81.38	93.91	90.30	81.34	93.48	90.38
0.67	0.125	22.74	0	79.81	93.22	89.57	80.89	93.52	90.48
0.67	0.25	53.18	0	80.30	93.19	88.92	80.47	93.18	89.06
0.67	0.5	79.75	0	78.65	93.49	87.70	79.84	93.00	88.54
0.67	1	92.32	0	77.37	92.73	86.58	79.70	93.07	87.78
0.67	2	96.78	0	74.41	93.07	84.82	77.18	92.94	86.48
0.67	4	98.33	0	94.07	94.54	94.07	74.20	93.34	84.91
0	0	4.44	1	93.88	94.79	93.88	94.22	94.90	94.22
0	0.125	23.98	1	93.37	94.91	93.37	93.57	94.53	93.57
0	0.25	54.13	1	93.56	94.90	93.56	93.61	94.48	93.61
0	0.5	80.25	1	93.63	94.96	93.63	92.97	94.41	92.97
0	1	92.51	1	92.93	94.82	92.93	93.75	95.08	93.75
0	2	96.85	1	92.63	94.99	92.63	93.09	95.05	93.09
0	4	98.38	1	92.40	94.09	92.73	92.30	94.70	92.30
0.35	0	4.21	1	91.72	94.41	92.71	92.37	94.32	93.14
0.35	0.125	23.74	1	91.72	94.04	92.77	91.87	94.45	92.63
0.35	0.25	53.97	1	91.90	93.58	92.52	91.85	94.57	92.28
0.35	0.5	80.18	1	91.55	94.39	92.07	92.22	94.15	92.57
0.35	1	92.44	1	91.17	94.43	91.46	91.49	94.26	92.26
0.35	2	96.84	1	89.84	94.77	90.48	91.14	94.52	91.99
0.35	4	98.37	1	89.59	94.36	91.86	90.24	94.62	90.89
0.55	0	4.17	1	89.10	93.76	90.86	89.12	93.97	91.44
0.55	0.125	23.50	1	89.10	94.13	91.37	89.78	93.84	91.26
0.55	0.25	53.87	1	88.33	93.68	90.99	88.96	93.72	90.83
0.55	0.5	80.11	1	88.61	93.82	90.46	88.44	93.39	90.50
0.55	1	92.49	1	87.46	93.09	89.11	88.63	93.54	90.45
0.55	2	96.86	1	85.43	93.32	88.20	88.06	93.44	89.47
0.55	4	98.38	1	80.85	93.75	91.07	86.65	93.38	88.39
0.67	0	4.01	1	80.84	93.26	90.05	81.30	93.98	90.37
0.67	0.125	22.64	1	80.16	93.60	88.95	80.98	93.51	89.82
0.67	0.25	53.03	1	81.20	92.86	88.36	81.05	93.28	89.10
0.67	0.5	79.86	1	78.35	93.26	87.77	79.76	92.76	88.57
0.67	1	92.28	1	77.27	93.11	86.77	78.67	93.38	88.29
0.67	2	96.79	1	75.01	93.30	84.31	76.97	92.83	86.19
0.67	4	98.31	1	88.07	94.05	91.10	74.86	92.66	84.58
Average				**88.07**	**94.05**	**91.10**	**88.12**	**94.01**	**91.04**

Notes: This table presents the results of the Monte Carlo simulation (Figures 31 and 34) with 10,000 runs. G is the Gini coefficient measuring the unbalancedness of the panel data, σ_g is the standard deviation of the random excess heterogeneity, I^2 measures the random excess heterogeneity relative to sampling error, and α_1 is the true effect. $\boldsymbol{V}^{-1}$ is the meta-regression with inverse variance weighting, $\boldsymbol{m}^{-1}$ is the weighting by the inverse number of estimates per study, and $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$ is the combined weight. The columns report the coverage proportions, which are averaged across the 10,000 replications. d denotes the level of sample overlap.

When random heterogeneity is defined at the study level, clustered errors have a positive impact on the coverage proportions of all three weights. The coverage of $\boldsymbol{V}^{-1}$ increases to 88.12%, 94.01% for $\boldsymbol{m}^{-1}$, and 91.04% for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. This corresponds to an increase of 19.05 percentage points for $\boldsymbol{V}^{-1}$, 18.83 percentage points for $\boldsymbol{m}^{-1}$, and 15.95 percentage points for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$. Considering the results for the MSE and bias (Appendices C.3 and C.7), which are lowest for $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$, it seems evident that the combined weight, $(\boldsymbol{V} \circ \boldsymbol{m})^{-1}$, coupled with cluster-robust standard errors, achieves the best relative performance compared to the other two weights (given the specific assumptions of the simulation design).

4.6. Discussion

4.6.1. Major Findings and Implications

What are the optimal weights when meta-regression analysis is estimated by WLS? According to statistical theory, inverse variance weighting is superior in the presence of heteroscedasticity. But what are the implications for the dominance of inverse variance weighting when multiple effect sizes are estimated in the same primary study and the number of estimates varies across studies? This is the research questions to be addressed in this chapter.

The Monte Carlo simulation in this chapter extends the simulation design by Stanley and Doucouliagos (2017) in the sense that it allows each primary study to report one or multiple effect size estimates. The varying number of estimates per study is determined by the Gini coefficient. To consider that multiple effect sizes within the same study may not be independent, the simulation adds two dimensions of within-study dependence. First, the random samples from which the estimates are computed are overlapping in the same study. Second, random heterogeneity is constant at the study level, resulting in a study-specific mean of the sampling distribution that is the same for all effect sizes estimated in the same study.

The main findings of the simulation can be summarized as follows. When random heterogeneity is effect size-specific and studies report a varying number of multiple estimates, the inverse of the effect sizes' variance is the dominant weight. Because random heterogeneity is defined for each individual effect size, the mean of the sampling distribution varies across effect sizes (within and across studies), implying that effect sizes are independent of each other even when assigned to the same study. The average coverage proportions of the meta-regression estimates using inverse variance weighting is about 93% for all levels of sample overlap, which is close to the nominal of 95%. Sample overlap in the same study implies within-study correlation. The finding that a higher sample overlap does not affect the dominant performance of inverse variance weighting is differs from Bom and Rachinger (2020), who find lower coverage ratios and a higher MSE for the WLS estimator with inverse variance weighting when the sample overlap increases. The difference in the results might be explained by the differences in the simulation design. Bom and Rachinger (2020) set the number of effect sizes to be either 32, 128, or 512, and show that the negative impact of sample overlap increases in larger samples, especially when the number of effect size estimates is 512. However, when the number of effects is only 32, the changes in their results for the WLS estimator are marginal when the correlation of the overlapping estimates increases. This is a similar result to that seen when comparing the results in Table 23 and Table 25. The number of effects in this chapter is 80, which is rather at the lower end of the range of estimates in Bom and Rachinger (2020). Using a larger number of estimates in the simulation (e.g., 128 or 512) may increase the effect of overlapping samples. In addition, the overlap structure in the simulation approach used in this chapter is similar, but not identical, to Bom and Rachinger (2020), which could also affect the results.

As in previous simulations of meta-regressions (among others, Stanley, 2008; Stanley and Doucouliagos, 2014, 2017), random heterogeneity is the most influential parameter for the statistical properties of the meta-estimators. When random heterogeneity is constant for all estimates in the same study, there is an equal mean of the sampling distribution for all estimated effects of this study. This implies a fixed effects structure at the study level. In this case, the effect sizes within the same study have the same mean, but the means vary randomly between studies, imposing a dependence structure for the estimates reported in the same study. In this scenario of study-level random heterogeneity, the relative performance of the three examined weighting schemes is negatively affected because the dependence among the estimates within the same study *"drives a wedge between the nominal and the informative sample, causing the standard errors of conventional meta-analysis methods to be underestimated"* (Bom and Rachinger, 2020: 829). In contrast to the case where heterogeneity is defined at the effect size level, a larger unbalancedness of the effect size estimates per study and a larger sample overlap have an additional negative impact. With increasing sample overlap and higher Gini coefficients, the share of independent sampling information included in each effect size estimate decreases as more estimates are sampled from a distribution with the same mean. Accordingly, a higher Gini coefficient increases the dependence that is induced by study-level heterogeneity. The relative performance of inverse variance weighting is lowest when a high level of unbalancedness ($G = 0.67$) is combined with high levels of random heterogeneity ($\sigma_g = \{1,2,4\}$). In the extreme, i.e., with maximum random heterogeneity and the largest unbalancedness, the coverage ratio of inverse variance weighting falls below 50% with high values for the MSE and bias. In this case, the informative power of the inverse variance is the lowest.

In contrast to inverse variance weighting, the inverse number of effect size estimates per study is a rather 'naïve' weighting scheme because it assigns equal weight to each study. Thereby, it reduces the heavy weight that inverse variance weighting imposes on the studies with many estimates. Hence, this alternative weight is less affected by the imposed dependencies at the study level. Although the inverse number of estimates per study shows higher coverage ratios as compared to inverse variance weighting in the case of study-level fixed heterogeneity, the estimator is not efficient because it has the highest MSE among the three weights, along with the highest values for the bias. Interestingly, the combined weight of the inverse variance times the inverse number of observations per study has, on average, higher coverage ratios than inverse variance weighting while exhibiting the highest efficiency in terms of MSE and the lowest bias when random heterogeneity is study-specific. This combined weight incorporates the advantage of the inverse number of estimates per study as it assigns the same weight to all studies regardless of the number of estimates reported, while accounting for precision of estimates through inverse variance weighting, which in turn makes the weight proportional to the standard error and corrects for heteroscedasticity. Cluster-robust standard errors improve the coverage ratios for all weights, with the combined weight having average coverage proportions of 91% when errors are clustered.

Random heterogeneity at both the effect size level and at the study level are extreme scenarios. In practice, however, it is likely that primary studies contain features of both cases. Therefore, the statistical conclusions could lie somewhere between the results discussed in this chapter. Nevertheless, the previous results imply that inverse variance is not always the superior weighting scheme. The recommendation to be derived from the results of this chapter is that in meta-analyses in which primary studies report multiple effect size estimates and the effect size estimates follow a similar study-level distribution, robustness tests of the meta-regressions using the combined weight of inverse variance and inverse number of estimates per study should be applied. Moreover, reporting the Gini coefficient to show the concentration of the number of estimates in the meta-sample is a useful measure of the unbalancedness in the panel data structure.

Applied meta-analysis has attempted to circumvent the problems of dependence among multiple estimates and an undue influence of studies reporting many effect size estimates in different ways. One method, which is similar to the case of weighing the meta-regression by the inverse number of reported effect size estimates, is to run the meta-regression on the average study estimate (Stanley, 2001: 138). In this case, each study imports an average effect size estimate into the meta-sample. To calculate the average effect size per study, Hunter and Schmidt (2004: 430-432) recommend a weighted mean using the primary study sample size for the weighting. However, this approach is inferior as it results in fewer observations, which leads to a substantial loss of efficiency (Stanley and Doucouliagos, 2012: 72). Therefore, it has become the standard approach in meta-regression to sample all available information and to address the problems of multiple estimates in the meta-regression model (Havranek et al., 2020: 472).

4.6.2. Limitations and Further Research

As with any other simulation study, the design of the Monte Carlo simulation relies on certain assumptions and calibrations. The design decisions and parameter values for the simulations in this chapter are either derived from previous simulation studies of meta-regression methods or they are configured in such a way that they resemble practical conditions in economics and finance research. However, any parameter calibration involves subjectivity. To avoid biased inferences from subjective parameter calibrations, several sensitivity analyses are reported for key parameters, like the Gini coefficient, the level of random heterogeneity, or the degree of sample overlap. Nevertheless, the simulation results are specific to the underlying simulation design that generates an ‘artificial’ data set. In the real world, there is probably an even greater variety of complications than those simulated in this chapter. Any extensions to the simulation design could bring the results closer to reality, but at the same time every model extension comes with new assumptions and additional complexity.

Various extensions of the simulation design and modifications of the model calibrations are possible to reflect the practical conditions in meta-analysis. The reported simulations introduce within-study dependence through sample overlap and study-level heterogeneity, while assuming independence between studies. However, multiple primary studies typically draw samples from similar populations. For example, two studies examine the hedging behavior of all non-financial companies in the S&P 500 for two overlapping time periods, for example, 1990–2010 in study $i = 1$ and 2000–2020 in study $i = 2$. In this scenario, between-study dependence arises because the samples are overlapping and the population of companies, which are the non-financial companies in the S&P 500, is the same in the two studies. The simulation design could reflect such between-study dependence by drawing the samples for each estimate from a global population or from different groups of populations. Similarly, Bom and Rachinger (2020: 820) define either one or multiple groupings of the underlying sample from which the observations for the primary studies are taken. The case of one group, which implies a world sample, would reflect the case that in corporate finance, according to De La Cruz et al. (2019: 7), there are about 41,000 listed companies globally with financial reporting data available for empirical analysis.

Defining random heterogeneity at the study level implies a fixed effects panel structure to the data. Following econometric theory, a pooled OLS model, as applied in the simulation's meta-regression model, might be inferior as it does not explicitly address the fixed effects panel structure. Hence, the model might be misspecified, which could explain the poor performance of the meta-regression estimators when random heterogeneity is defined at study level. In this specific case, panel regression methods are usually preferred (Greene, 2011: 491). By adding study-level fixed effects in the meta-regression model, potential dependencies among the estimates within a given study can be accommodated (Stanley and Doucouliagos, 2012: 113). Further research should evaluate the statistical properties of the three weighting schemes in such a fixed effect panel regression.

In addition to the composition of the sample population with between-study overlaps, random heterogeneity could be defined in such a way that it reflects between-study dependencies. Instead of defining a constant, $\alpha_{3,i}$, at the study level, $\alpha_{3,s}$ could be determined for a specific number of $s = 1, .., S$ groups, with multiple studies drawing their estimates from the same group. This would imply that the mean of the sampling distribution is specific for each group, which might reflect different effects for different countries or industries examined in multiple studies. Moreover, excess random heterogeneity in the simulation is induced by random omitted-variable bias. Following Stanley and Doucouliagos (2017: 26), there are alternative approaches to modeling random heterogeneity. This can either be adding a random disturbance directly to the true mean effect, α_1, or by allowing the true effect of X_1 on z in the primary regression to depend on random variations in some moderator variable.

Related to the design of sample overlap and random heterogeneity is the distribution of the multiple effect size estimates reported in the same study. The distribution of effect size estimates is defined such that the Gini coefficient of the 80 estimates is equal to the proposed thresholds of 0, 0.35, 0.55, and 0.67. However, other distributions of the number of effect size estimates per study could lead to the same Gini coefficient. In particular, when the meta-sample increases from 80 observations to, for example, 512, as in Bom and Rachinger (2020: 825), there are multiple distributions of effect size estimates across the 20 or more studies that produce the desired range of Gini coefficients.

In addition to the design of sample overlap and random heterogeneity, the WLS estimator and the weights could be extended. Bom and Rachinger (2020) present a GW solution that addresses the challenge of overlapping samples by explicitly modeling the variance-covariance matrix that defines the structure of the dependence among the effect size estimates. The GW estimator could outperform the other weights when the sample overlap is large.

Another concern that is not addressed in the simulation design is the impact of precision outliers. When highly precise effect size estimates are biased by incorrect model specification or the application of some idiosyncratic methods, they can produce misleading meta-regression results. Because inverse variance weighting implicitly assigns larger weights to these estimates, it is more affected by precision outliers than the number of estimates reported per study or the combined weight. To examine the impact of precision outliers, the simulation design could be extended by adding a specific ratio of outliers to the random samples.

Another challenge is related to the potential endogeneity of the standard errors (Zigraiova and Havranek, 2016: 973). Such endogeneity might arise when a particular study characteristic or method affects both the effect size estimate and the standard error of the effect. In the case of endogeneity, meta-regression estimates will be biased for any coefficient estimate that includes the standard error of the effect size estimate. As inverse variance weighting considers the standard error of the effect size estimate in all observations, the impact of endogeneity might be stronger for inverse variance weighting than for the alternative weights. The simulation design could be extended to model the effect size estimates and their standard errors such that they are both affected by a specific study characteristic. In the simulation design above, this is partially addressed as a larger value of σ_g has an impact on both the effect size estimate as well as the variation of the effect.

Finally, the simulation investigates weighting schemes in meta-regression. Recent meta-analysis studies use BMA to explore the drivers of heterogeneity, as applied in Section 3.7.6 for the hedging and firm value nexus. The input values for BMA can also be weighted by the inverse variance or the inverse number of estimates per study. The analysis of the impact of weighting schemes on the BMA results could be the subject of further research.

4.7. Summary

In applied meta-analysis in economics and finance, it is quite common for primary studies to report multiple estimates of the effect size under study. Often, the number of estimates varies among studies, forming an unbalanced panel data set. Additionally, study-specific sample characteristics impose a dependency structure among the effect size estimates reported in the same study. Traditional inverse variance weighting, which is the optimal estimator from a statistical point of view, has the disadvantage that studies reporting a large number of estimates may have an *"undue influence"* on the meta-analysis results (Stanley and Doucouliagos, 2012: 72). Therefore, alternative weights have been proposed to address the challenges of inverse variance weighting in unbalanced panel data settings, including the inverse number of estimates per study and the combination of the inverse variance and the inverse number of estimates per study. Nevertheless, the statistical properties and the relative performance of these alternative weighting schemes in comparison to inverse variance weighting have not yet been evaluated in previous simulation studies of meta-regression methods.

The main contribution of this chapter is to examine the statistical properties (i.e., bias, MSE, and coverage) of inverse variance weighting and the alternative weighting schemes in meta-regressions with unbalanced panel data. Thus, the simulation design of previous simulation studies is extended in such a way that primary studies estimate varying numbers of effect sizes. In addition, sample overlap and random heterogeneity at the study level induce a complex structure of dependence among the estimates within the same study. The main findings of this chapter can be summarized as follows:

(1) When random heterogeneity is defined at the effect size level, inverse variance is the superior weighting scheme. This conclusion also applies when many estimates are concentrated in a few studies and the estimates in the same study are sampled from a data set with overlapping observations.

(2) When random heterogeneity is defined at the study level, imposing a mean of the sampling distribution that is the same for all estimates in the same study (fixed effect panel structure), the combined weight of the inverse variance and the inverse number of estimates per study obtains the best relative performance in terms of coverage, MSE, and bias. The statistical properties of inverse variance weighting are particularly poor when sample overlap and unbalancedness of the sample are large, along with high values of random heterogeneity.

(3) Publication selection increases the bias and lowers the efficiency of all three weighting schemes. On average, the combined weight of the inverse variance and the inverse number of estimates per study is preferrable as it has the lowest bias and the lowest MSE. The preference for the combined weight increases when random heterogeneity and unbalancedness are large.

(4) Cluster-robust standard errors increase the power of all three weighting schemes. In the case of study-level random heterogeneity, the average coverage rates of the combined weight are highest with a deviation of 1% from the nominal (95%).

If not accounted for, unbalanced and non-independent panel data can harm the statistical properties of inverse variance weighting. When the meta-data set is dominated by a few studies with a large number of estimates per study and random heterogeneity is study level-specific, the combined weight of the inverse variance times the inverse number of estimates per study provides a valid and practical complement to inverse variance weighting. The recommendation to be derived from the results of this chapter is that if the panel structure is unbalanced, the combined weight of inverse variance times the inverse number of estimates reported per study should be used as a robustness test in addition to conventional inverse variance weighting.

As with any other simulation, the approach presented in this chapter comes with several limitations, especially since the design of the data generating process does not necessarily imply external validity of the results and reality is often more complex than considered in this simulation. Several avenues for further research have been proposed to extend the simulation design by allowing between-study dependence, modeling precision outliers, and explicitly considering an endogenous relation between the effect size estimate and its standard error.

Chapter 5. Conclusion

Good decision-making at both the corporate and policy level should be supported by *"reliable scientific knowledge"* (Ioannidis et al., 2017: F236). On the positive side, empirical research output in finance, as in many other research disciplines, is steadily growing and providing more evidence to inform good decision-making. However, although each additional empirical result sheds more light on the phenomenon under research, the reported findings are not always conclusive and often depend on certain sample characteristics and method choices. In addition, researchers may actively select certain findings for publication and leave unfavorable outcomes unpublished, so that the reported results represent a biased sample of the population effect. Due to the large heterogeneity of empirical results in finance research and the risk of biased effects arising from publication selection, there is an increasing need for meta-research that collects and reviews the existing empirical literature in an objective manner. This thesis examines meta-analysis as a valuable complement to the toolkit used in finance research that allows to integrate, review, and correct previous empirical research findings.

Chapter 2 introduced the general concept of meta-analysis as well as related terms and definitions. The chapter also described the three main meta-analysis methods that have been used in finance to date, including meta-averaging, meta-regression, and meta-analytic structural equation modeling. The introduction of the meta-analysis methods was followed by an overview of the current state of meta-research in finance. A review of 76 previous meta-analyses in finance was presented, showing an increasing acceptance, with most meta-analyses in finance published within the past four years. The previous literature was then used to derive opportunities and challenges for meta-analysis in finance. Among the key strengths of meta-analysis is the ability to accumulate knowledge while increasing the precision of estimated mean effects by combining the statistical power of many studies. Other important strengths include the ability to identify and correct publication selection bias and model misspecification in primary studies, to explore the sources of systematic heterogeneity among the reported effect size estimates, as well as to enable predictions and support evidence-based decision-making. At the same time, meta-analysis is accompanied by challenges in selecting a meaningful effect size measure, assembling the best sample of primary studies, and dealing with econometric problems due to multidimensional dependence structures in the meta-data set. Similarly, the quality of data coding and model specification are critical for a meaningful meta-analysis. In summary, only a carefully conducted meta-analysis can take advantage of the many opportunities offered by meta-research methods.

Chapter 3 applied meta-analysis in the field of corporate finance by integrating the empirical literature that examines the effects of corporate hedging on firm value. The literature on the firm value effects of corporate hedging provides an ideal setting for the application of meta-analysis, given the richness of existing research results and the wide heterogeneity among the countries, time periods, and methods being investigated. Although studies differ in their design, the reported estimates refer to the hedging premium as a common measure of effect size. Meta-regression analysis was used to accumulate a hand-collected data set of 1,016 estimates for the hedging premium reported in 71 primary studies. Using meta-regression analysis, I find that the reported effects of hedging on firm value are systematically higher for FX hedgers than for IR and CP hedgers, for studies published in higher-ranked journals, and for models estimated with firm fixed effects and controls for endogeneity. The results also suggest that hedging premiums increase significantly when a study considers operational hedging strategies in addition to financial hedging. Moreover, the meta-regression results reveal geographical differences, with larger firm value effects through hedging in less developed financial markets and countries with higher tax rates. Considering the existing literature and assuming a best practice study design, meta-regression predicts an overall hedging premium of 1.8% in firm value for FX hedgers as compared to non-hedgers. Moreover, the 'best' practice predictions show a firm value discount of -0.8% for IR hedgers, as well as a discount of -0.6% for CP hedgers. These predictions imply the aggregated hedging premiums/discounts and can be seen as the accumulated empirical evidence in the research field, while accounting for systematic heterogeneity, publication selection, and model misspecification.

Chapter 4 compared the statistical properties of three common weighting schemes applied in WLS meta-regression. This chapter aimed at making further contributions to our understanding of meta-regression when primary studies report more than one effect size estimate. A Monte Carlo simulation was designed in such a way that it considers that primary studies report a varying number of multiple effect size estimates and that samples used for the estimation of multiple estimates may overlap. In addition, the simulation accounts for publication selection and two groups of random heterogeneity: observable and unobservable random heterogeneity. When random heterogeneity is defined at the effect size level, the inverse variance of the effect sizes' variance is the optimal weighting scheme, even when the primary studies report a largely unbalanced number of effect size estimates. When random heterogeneity is study level-specific, imposing a mean of the sampling distribution that is the same for all estimates in the same study, the combined weight of the inverse variance times the inverse number of estimates per study, obtains the best performance in terms of coverage, MSE, and bias. The statistical properties of inverse variance weighting are particularly poor when sample overlap and unbalancedness of the sample are large, along with high values of random heterogeneity at the study level. Introducing publication selection leads to a similar relative performance of the three weighting schemes, with the combined weight being the preferred weighting scheme when random heterogeneity is study-specific. However, publication selection increases the inefficiency and bias of all three weights. Cluster-robust standard errors move the coverage ratios closer to the nominal and, in the case of the combined weight, almost to the nominal of 95%. As a key conclusion, it can be inferred that for unbalanced panel data in a meta-regression, the combined weight of the inverse effect size variance times the inverse number of estimates reported per study should be applied, at least, for robustness checks of the meta-regression results.

In finance, we routinely find large discrepancies in what researchers report on a particular empirical effect. The differences in reported research findings are often coupled with a large amount of genuine heterogeneity resulting from the natural variation that researchers introduce through their choice of methods, specific data characteristics, and their idiosyncratic approach to research. Along with heterogeneity comes the challenge that the credibility of empirical research results diminishes when data mining and p-hacking can be abused to achieve any desired result. Heterogeneity and publication selection bias often make it difficult to assess the true impact of a particular research phenomenon, which lies among a sheer mass of results that may be contaminated by noise, exaggeration, and misinformation. Moreover, the presence of questionable research practices to obtain desired empirical research results can translate into a crisis regarding the credibility of empirical research. It is even worse when corporate or policy decisions are guided by such inaccurate and biased research findings.

Addressing the challenges posed by the wide heterogeneity of reported empirical research findings and strengthening the credibility of scientific evidence requires a multifaceted approach in which meta-analysis can be an important element. The introduction (Chapter 2), application (Chapter 3), and analysis of the methodological challenges (Chapter 4) of meta-analysis show that finance research can benefit from meta-analysis because it provides objective and critical methods to integrate conflicting empirical research findings and to filter out some of the biases that are often found in the reported results of primary empirical studies (these are, in particular, publication selection bias and model misspecification bias). Carefully conducted meta-analyses offer finance researchers the opportunity to summarize and evaluate an area of research while using the same econometric methods and statistical approaches that generated the research. However, meta-analysis also has its weaknesses and limitations that must be thoughtfully considered.

In hundreds of applications in economics, business, and medical research (among others), meta-analysis research *"has proven to be effective at cleansing the often murky and muddy waters of the ever-growing pool of econometric results"* (Stanley and Doucouliagos, 2012: 147). However, despite many advantages, meta-analysis in finance is still a rarely used research method with much untapped potential. Although meta-analysis is still a fairly young discipline in finance, it can adopt proven methods and developments from other research disciplines where meta-analysis is already more widely accepted and developed, especially economics. However, there are also areas where finance needs to develop its own approaches to address discipline-specific nuances. These may include, but are not limited to, strategies to appropriately account for the multilevel dependence structure that is common in financial data with overlapping sample data within and across studies. Finance-specific effect sizes, such as (stock) returns and their standard errors, also require best practices, especially when effect sizes need to be recalculated from other information reported in the primary studies. In addition, best practices for finance-specific moderator variables that explain sources of systematic heterogeneity could provide helpful guidance for future meta-research. As in economics, finance research generally involves observational studies with a wide range of possible model specifications to choose from. However, even the best model cannot eliminate all potentially confounding effects. Dealing with and correcting for misspecification caused by finance-specific model specifications has not yet been fully explored.

I hope that this thesis will make meta-analysis methods more accessible to finance researchers, stimulate its future application in the field, and help meta-analysis become a complement to conventional primary research and replication studies. There are several promising topic research areas in corporate finance, asset pricing, and banking, where a systematic aggregation through meta-analysis could contribute to the progress of the research field. For meta-research to grow in finance, it needs acceptance, not as a niche method, but rather as a powerful addition to primary research approaches and narrative literature reviews. This requires the community and its thought leaders to be receptive towards meta-research. Therefore, this thesis is also a call to editors, reviewers, and conference committees in the finance community to encourage meta-analysis methodology, as it enables researchers to integrate, compare, and correct often conflicting empirical findings. Other disciplines, particularly medicine but increasingly also economics, show that science can benefit from a vigorous culture in which meta-analysis and replication studies become an additional layer in the deductive research process that complements primary research by synthesizing scientific knowledge and inspiring future research.

References

Abdullah, A., Doucouliagos, H., Manning, E., 2015. *Does education reduce income inequality? A meta-regression analysis*. Journal of Economic Surveys 29(2), 301-316. DOI: 10.1111/joes.12056.

Adam, T.R., Fernando, C.S., Salas, J.M., 2017. *Why do firms engage in selective hedging? Evidence from the gold mining industry*. Journal of Banking and Finance 77, 269-282. DOI: 10.1016/j.jbankfin.2015.05.006.

Adam, T.R., Nain, A., 2013. *Strategic Risk Management and Product Market Competition*, in: Batten, J.A., MacKay, P., Wagner, N. (Eds.), Advances in Financial Risk Management - Corporates, Intermediaries and Portfolios. Palgrave Macmillan, Basingstoke, 3-29. DOI: 10.1057/9781137025098.

Afza, T., Alam, A., 2016. *Foreign Currency Derivatives and Firm Value*. European Online Journal of Natural and Social Sciences 5(1), 1-14.

Aguinis, H., Cascio, W.F., Ramani, R.S., 2017. *Science's reproducibility and replicability crisis: International business is not immune*. Journal of International Business Studies 48(6), 653-663. DOI: 10.1057/s41267-017-0081-0.

Ahmed, H., Azevedo, A., Guney, Y., 2013. *The Effect of Hedging on Firm Value and Performance: Evidence from the Nonfinancial UK Firms*. Paper presented at the Multinational Finance Society Annual Conference 2013, Izmir.

Aiello, F., Bonanno, G., 2016. *Efficiency in Banking: A Meta-Regression Analysis*. International Review of Applied Economics 30(1), 112-149. DOI: 10.1080/02692171.2015.1070131.

Aiello, F., Bonanno, G., 2018. *On the Sources of Heterogeneity in Banking Efficiency Literature*. Journal of Economic Surveys 32(1), 194-225. DOI: 10.1111/joes.12193.

Aitken, A.C., 1935. *On Least Squares and Linear Combinations of Observations*. Proceedings of the Royal Society of Edinburgh 55, 42-48. DOI: 10.1017/S0370164600014346.

Akobeng, A.K., 2005. *Understanding Systematic Reviews and Meta-analysis*. Archives of Disease in Childhood 90(8), 845-848. DOI: 10.1136/adc.2004.058230.

Alam, N., Gupta, A., 2018. *Does hedging enhance firm value in good and bad times*. International Journal of Accounting & Information Management 26(1), 132-152. DOI: 10.1108/IJAIM-03-2017-0041.

Alkebäck, P., Hagelin, N., 1999. *Derivative Usage by Nonfinancial Firms in Sweden with an International Comparison*. Journal of International Financial Management and Accounting 10(2), 105–120. DOI: 10.1111/1467-646X.00046.

Allayannis, G., Ihrig, J., Weston, J.P., 2001. *Exchange-Rate Hedging: Financial versus Operational Strategies*. American Economic Review 91(2), 391-395. DOI: 10.1257/aer.91.2.391.

Allayannis, G., Ofek, E., 2001. *Exchange rate exposure, hedging, and the use of foreign currency derivatives*. Journal of International Money and Finance 20(2), 273-296. DOI: 10.1016/S0261-5606(00)00050-4.

Allayannis, G., Ugur, L., Miller, D.P., 2012. *The Use of Foreign Currency Derivatives, Corporate Governance, and Firm Value Around the World*. Journal of International Economics 87(1), 65-79. DOI: 10.1016/j.jinteco.2011.12.003.

Allayannis, G., Weston, J.P., 2001. *The Use of Foreign Currency Derivatives and Firm Market Value*. The Review of Financial Studies 14(1), 243-276. DOI: 10.1093/rfs/14.1.243.

Almeida, H., Hankins, K.W., Williams, R., 2017. *Risk Management with Supply Contracts*. The Review of Financial Studies 30(12), 4179-4215. DOI: 10.1093/rfs/hhx051.

Amberg, N., Friberg, R., 2016. *Three Approaches to Risk Management — and How and Why Swedish Companies Use Them*. Journal of Applied Corporate Finance 28(1), 86-94. DOI: 10.1111/jacf.12161.

Amini, S., Parmeter, C.F., 2020. *A Review of the 'BMS' Package for R with Focus on Jointness*. Econometrics 8(6), 1-20. DOI: 10.3390/econometrics8010006.

Ammon, N., 1998. *Why Hedge? - A Critical Review of Theory and Empirical Evidence*, ZEW Discussion Paper No. 98-18, Mannheim.

Anderson, E., d'Orey, M.A.J., Duvendack, M., Esposito, L., 2018. *Does Government Spending Affect Income Poverty? A Meta-regression Analysis*. World Development 103, 60-71. DOI: 10.1016/j.worlddev.2017.10.006.

Anderson, R.G., Kichkha, A., 2017. *Replication, Meta-Analysis, and Research Synthesis in Economics*. American Economic Review 107(5), 56-59. DOI: 10.1257/aer.p20171033.

Andrews, I., Kasy, M., 2019. *Identification of and correction for publication bias*. American Economic Review 109(8), 2766-2794. DOI: 10.1257/aer.20180310.

Anello, C., Fleiss, J.L., 1995. *Exploratory or Analytic Meta-Analysis: Should We Distinguish Between Them?* Journal of Clinical Epidemiology 48(1), 109-116. DOI: 10.1016/0895-4356(94)00084-4.

Aretz, K., Bartram, S.M., 2010. *Corporate Hedging and Shareholder Value*. The Journal of Financial Research 33(4), 317-371. DOI: 10.1111/j.1475-6803.2010.01278.x.

Arnold, M.M., Rathgeber, A.W., Stöckl, S., 2014. *Determinants of Corporate Hedging: A (statistical) meta-analysis*. The Quarterly Review of Economics and Finance 54(4), 443-458. DOI: 10.1016/j.qref.2014.05.002.

Ashenfelter, O., Harmon, C., Oosterbeek, H., 1999. *A review of estimates of the schooling/earnings relationship, with tests for publication bias*. Labour Economics 6(4), 453-470. DOI: 10.1016/S0927-5371(99)00041-X.

Askarov, Z., Doucouliagos, H., 2013. *Does aid improve democracy and governance? A meta-regression analysis*. Public Choice 157(3-4), 601-628. DOI: 10.1007/s11127-013-0081-y.

Astakhov, A., Havranek, T., Novak, J., 2019. *Firm Size and Stock Returns: A Meta-Analysis*. Journal of Economic Surveys 33(5), 1463-1492. DOI: 10.1111/joes.12335.

Asteriou, D., Hall, S.G., 2011. *Applied Econometrics*, 2nd ed. Palgrave Macmillan, Basingstoke. DOI: 10.12691/jfe-3-1-4.

Ayturk, Y., Gurbuz, A.O., Yanik, S., 2016. *Corporate derivatives use and firm value: Evidence from Turkey*. Borsa Istanbul Review 16(2), 108-120. DOI: 10.1016/j.bir.2016.02.001.

Bachiller, P., 2017. *A meta-analysis of the impact of privatization on firm performance*. Management Decision 55(1), 178-202. DOI: 10.1108/MD-12-2015-0557.

Bachiller, P., Boubaker, S., Mefteh-Wali, S., 2021. *Financial derivatives and firm value: What have we learned?* Finance Research Letters 39. DOI: 10.1016/j.frl.2020.101573.

Bae, S.C., Kim, H.S., Kwon, T.H., 2016. *Foreign Currency Debt Financing, Firm Value, and Risk: Evidence from Korea Surrounding the Global Financial Crisis*. Asia-Pacific Journal of Financial Studies. DOI: 10.1111/ajfs.12123.

Bae, S.C., Kim, H.S., Kwon, T.H., 2018. *Currency derivatives for hedging: New evidence on determinants, firm risk, and performance*. Journal of Future Markets 38(4), 446-467. DOI: 10.1002/fut.21894.

Bai, S., Parajuli, B.S., Ryan, C., 2016. *The impact of hedging on firm value: evidence from US multinational corporations*. International Journal of Services and Standards 11(4), 380-392. DOI: 10.1504/ijss.2016.082433.

Bajzik, J., 2020. *Trading Volume and Stock Returns: A Meta-Analysis*. IES Working Paper 45/2020 at Charles University.

Bajzik, J., 2021. *Trading Volume and Stock Returns: A Meta-Analysis*. International Review of Financial Analysis 78(101923). DOI: 10.1016/j.irfa.2021.101923.

Bandaranayake, S., Das, K.K., Reed, R.W., 2020. *Another Look at 'Bank Competition and Financial Stability: Much Ado About Nothing'*. Journal of Economic Surveys 34(2), 344-371. DOI: 10.1111/joes.12352.

Banks, E., 2004. *Alternative Risk Transfer: Integrated Risk Management through Insurance, Reinsurance, and the Capital Markets*. John Wiley & Sons, Chichester.

Baron, R.M., Kenny, D.A., 1986. *The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations*. Journal of Personality and Social Psychology 51(6), 1173-1182. DOI: 10.1037//0022-3514.51.6.1173.

Bartram, S.M., 2019. *Corporate hedging and speculation with derivatives*. Journal of Corporate Finance 57, 9-34. DOI: 10.1016/j.jcorpfin.2017.09.023.

Bartram, S.M., Brown, G.W., Conrad, J., 2011. *The Effects of Derivatives on Firm Risk and Value*. Journal of Financial and Quantitative Analysis 46(4), 967-999. DOI: 10.1017/S0022109011000275.

Bartram, S.M., Brown, G.W., Fehle, F.R., 2009. *International Evidence on Financial Derivatives Usage*. Financial Management 28(1), 185-206. DOI: 10.1111/j.1755-053X.2009.01033.x.

Bashir, H., Sultan, K., Jghef, O.K., 2013. *Impact of Derivatives Usage on Firm Value: Evidence from Non Financial Firms of Pakistan*. Journal of Management Research 5(4), 108-127. DOI: 10.5296/jmr.v5i4.4050.

Beasley, M.S., Clune, R., Hermanson, D.R., 2005. *Enterprise risk management: An empirical analysis of factors associated with the extent of implementation*. Journal of Accounting and Public Policy 24(6), 521-531. DOI: 10.1016/j.jaccpubpol.2005.10.001.

Beatty, A., Petacchi, R., Zhang, H., 2012. *Hedge commitments and agency costs of debt: evidence from interest rate protection covenants and accounting conservatism*. Review of Accounting Studies 17(3), 700-738. DOI: 10.1007/s11142-012-9189-4.

Becker, B.J., 1992. *Using Results From Replicated Studies to Estimate Linear Models*. Journal of Educational and Behavioral Statistics 17(4), 341-362. DOI: 10.3102/10769986017004341.

Begg, C.B., Berlin, J.A., 1988. *Publication Bias: A Problem in Interpreting Medical Data*. Journal of the Royal Statistical Society. Series A (Statistics in Society) 151(3), 419-463. DOI: 10.2307/2982993.

Belghitar, Y., Clark, E., Judge, A., 2008. *The Value Effects of Foreign Currency and Interest Rate Hedging: The UK Evidence*. International Journal of Business 13(1), 43-50.

Belghitar, Y., Clark, E., Mefteh, S., 2013. *Foreign Currency Derivative Use and Shareholder Value*. International Review of Financial Analysis 29, 283-293. DOI: 10.1016/j.irfa.2012.02.004.

Bentler, P.M., 1990. *Comparative fit indexes in structural models*. Psychological Bulletin 107(2), 238-246. DOI: 10.1037/0033-2909.107.2.238.

Bergh, D.D., Aguinis, H., Heavey, C., Ketchen, D.J., Boyd, B.K., Su, P., Lau, C.L.L., Joo, H., 2016. *Using Meta-Analytic Structural Equation Modeling to Advance Strategic Management Research: Guidelines and an Empirical Illustration via the Strategic Leadership-Performance Relationship*. Strategic Management Journal 37(3), 477-497. DOI: 10.1002/smj.2338.

Berk, J., DeMarzo, P.M., 2014. *Corporate Finance*, 3rd ed. Pearson, Boston.

Berrospide, J.M., Purnanandam, A., Rajan, U., 2008. *Corporate Hedging, Investment and Value*. European Finance Association Annual Meeting 2008, Athens.

Bessembinder, H., 1991. *Forward Contracts and Firm Value: Investment Incentive and Contracting Effects*. Journal of Financial and Quantitative Analysis 26(4), 519-532. DOI: 10.2307/2331409.

Bessler, W., Conlon, T., Huan, X., 2019. *Does corporate hedging enhance shareholder value? A meta-analysis*. International Review of Financial Analysis 61, 222-232. DOI: 10.1016/j.irfa.2018.11.010.

Bhatia, M., Gulati, R., 2021. *Board governance and bank performance: A meta-analysis*. Research in International Business and Finance 58(101425). DOI: 10.1016/j.ribaf.2021.101425.

Bialkowski, J., Perera, D., 2019. *Stock index futures arbitrage: Evidence from a meta-analysis*. International Review of Financial Analysis 61, 284-294. DOI: 10.1016/j.irfa.2018.09.002.

Bijmolt, T.H.A., Pieters, R.G.M., 2001. *Meta-Analysis in Marketing when Studies Contain Multiple Measurements*. Marketing Letters 12(2), 157-169. DOI: 10.1023/A:1011117103381.

BIS, 2016a. *Triennial Central Bank Survey: Foreign exchange turnover in April 2016*. Bank for International Settlements, Basel.

BIS, 2016b. *Triennial Central Bank Survey: Global foreign exchange market turnover in 2016*. Bank for International Settlements, Basel.

BIS, 2016c. *Triennial Central Bank Survey: Interest rate derivatives market turnover in 2016*. Bank for International Settlements, Basel.

BIS, 2019a. *Triennial Central Bank Survey: Foreign exchange turnover in April 2019*. Bank for International Settlements, Basel.

BIS, 2019b. *Triennial Central Bank Survey: Interest rate derivatives market turnover in 2019*. Bank for International Settlements, Basel.

Bodnar, G.M., Gebhardt, G., 1999. *Derivatives Usage in Risk Management by US and German Non-Financial Firms: A Comparative Survey*. Financial Management 10(3), 153-187. DOI: 10.1111/1467-646X.00049.

Bodnar, G.M., Giambona, E., Graham, J.R., Harvey, C.R., 2019. *A View Inside Corporate Risk Management*. Management Science 65(11), 5001-5026. DOI: 10.1287/mnsc.2018.3081.

Bodnar, G.M., Hayt, G.S., Marston, R.C., 1996. *1995 Wharton Survey of Derivatives Usage by US Non-Financial Firms*. Financial Management 25(4), 113-133. DOI: 10.2307/3665595.

Bodnar, G.M., Hayt, G.S., Marston, R.C., 1998. *1998 Wharton Survey of Financial Risk Management by US Non-Financial Firms*. Financial Management 27(4), 70-91.

Bollen, K.A., Jackman, R.W., 1985. *Regression diagnostics: An expository treatment of outliers and influential cases*. Sociological Methods & Research 13(4), 510-542. DOI: 10.1177/0049124185013004004.

Bom, P.R.D., Rachinger, H., 2019. *A kinked meta-regression model for publication bias correction*. Research Synthesis Methods 10(4), 497-514. DOI: 10.1002/jrsm.1352.

Bom, P.R.D., Rachinger, H., 2020. *A generalized-weights solution to sample overlap in meta-analysis*. Research Synthesis Methods 11(6), 812-832. DOI: 10.1002/jrsm.1441.

Booth, A., Sutton, A., Papaioannou, D., 2016. *Systematic Approaches to a Successful Literature Review*, 2nd ed. Sage Publications, London, UK.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. *Introduction to Meta-Analysis*. John Wiley & Sons, Chichester, UK. DOI: 10.1002/9780470743386.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2010. *A basic introduction to fixed-effect and random-effects models for meta-analysis*. Research Synthesis Methods 1(2), 97-111. DOI: 10.1002/jrsm.12.

Boyabatli, O., Toktay, L.B., 2004. *Operational Hedging: A Review with Discussion*. Working Paper at INSEAD Fontainebleau, France.

Brannick, M.T., Yang, L.-Q., Cafri, G., 2011. *Comparison of Weights for Meta-Analysis of r and d Under Realistic Conditions*. Organizational Research Methods 14(4), 587-607. DOI: 10.1177/1094428110368725.

Brealy, R.A., Myers, S.C., Allen, F., 2011. *Principles of corporate finance*, 10th ed. McGraw-Hill, New York.

Brodeur, A., Lé, M., Sangnier, M., Zylberberg, Y., 2016. *Star Wars: The Empirics Strike Back*. American Economic Journal: Applied Economics 8(1), 1-32. DOI: 10.1257/app.20150044.

Brown, G.W., Crabb, P.R., Haushalter, D., 2006. *Are Firms Successful at Selective Hedging?* The Journal of Business 79(6), 2925-2949. DOI: 10.1086/508004.

Brunzell, T., Hansson, M., Liljeblom, E., 2011. *The Use of Derivatives in Nordic Firms*. The European Journal of Finance 17(5-6), 355-376. DOI: 10.1080/1351847X.2010.543836.

Búa, M.V., González, L.O., López, S.F., Santomil, P.D., 2013. *Is Value Creation Consistent with Currency Hedging?* The European Journal of Finance 21(10-11). DOI: 10.1080/1351847X.2013.773262.

Buckley, A., 2004. *Multinational Finance*, 5th ed. Prentice Hall, Harlow.

Buckley, P.J., Devinney, T.M., Tang, R.W., 2013. *Meta-Analytic Research in International Business and International Management*, in: Buckley, P.J. (Ed.), The Multinational Enterprise and the Emergence of the Global Factory. Palgrave Macmillan, London, 100-134. DOI: 10.1057/9781137402387.

Burkhard, B., Siren, C.A., van Essen, M., Grichnik, D., 2018. *A Meta-Analytic Approach to Understanding the Effects of CEO Overconfidence on Firm Performance*. Paper presented at Academy of Management Annual Meeting 2018, Chicago.

CABS, 2018. *Academic Journal Guide 2018*, https://charteredabs.org/academic-journal-guide-2018-view/, Accessed: 2022/01/02.

Cambrea, D.R., La Rocca, M., La Rocca, T., Casciaro, S., 2017. *Cash holding and firm value: a meta-analysis*. Working Paper at Bocconi University. DOI: 10.3280/EI2015-002002.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2011. *Robust Inference With Multiway Clustering*. Journal of Business and Economic Statistics 29(2), 238-249. DOI: 10.1198/jbes.2010.07136.

Cameron, S., Price, D., 2009. *Business Research Methods: A Practical Approach*. Chartered Institute of Personnel and Development, London.

Candolo, C., Davison, A.C., Demétrio, C.G.B., 2003. *A Note on Model Uncertainty in Linear Regression*. Journal of the Royal Statistical Society. Series D (The Statistician) 52(2), 165-177. DOI: 10.1111/1467-9884.00349.

Card, D., Kluve, J., Weber, A., 2010. *Active Labour Market Policy Evaluations: A Meta-Analysis*. The Economic Journal 120(548), 452-477. DOI: 10.1111/j.1468-0297.2010.02387.x.

Card, D., Krueger, A.B., 1995. *Time-Series Minimum-Wage Studies: A Meta-Analysis*. The American Economic Review 85(2), 238-243.

Card, N.A., 2012. *Applied Meta-Analysis for Social Science Research*. Guilford Publications, New York.

Carney, M., Gedajlovic, E.R., Heugens, P.P.M.A.R., van Essen, M., van Oosterhout, J.H., 2011. *Business Group Affiliation, Performance, Context, and Strategy: A Meta-Analysis*. Academy of Management Journal 54(3), 437-460. DOI: 10.5465/AMJ.2011.61967812.

Carter, D.A., Rogers, D.A., Simkins, B.J., 2006. *Does Hedging Affect Firm Value? Evidence from the US Airline Industry*. Financial Management 35(1), 53-86. DOI: 10.1111/j.1755-053X.2006.tb00131.x.

Chang, F.-Y., Hsin, C.-W., Shiah-Hou, S.-R., 2016. *Performance of Financial Hedging and Earnings Management under Diverse Corporate Information Quality*, European Financial Management Annual Conference 2016, Basel.

Chatterjee, S., Hadi, A.S., 2006. *Regression Analysis by Example*. John Wiley & Sons, Hoboken. DOI: 10.1002/0470055464.

Chava, S., Roberts, M.R., 2008. *How Does Financing Impact Investment? The Role of Debt Covenants*. The Journal of Finance 63(5), 2085–2121. DOI: 10.1111/j.1540-6261.2008.01391.x.

Chen, C., Jin, Y., Men, M.-M., 2011. *Executive Compensation, Hedging, and Firm Value*. Working Paper at the California State University.

Chen, J., King, T.-H.D., 2014. *Corporate hedging and the cost of debt*. Journal of Corporate Finance 29, 221-245. DOI: 10.1016/j.jcorpfin.2014.09.006.

Chen, Y., Shao, Q.-q., 2010. *Firm Value Effects of Derivatives Hedging for Risk Exposure: An Empirical Research on Chinese Listed Enterprises*. Paper presented at 17th International Conference on Management Science & Engineering 2010, Melbourne.

Chesher, A., Jewitt, I., 1987. *The bias of a heteroskedasticity consistent covariance matrix estimator*. Econometrica 55(5), 1217-1222. DOI: 10.2307/1911269.

Cheung, M.W.-L., 2015. *Meta-Analysis: A Structural Equation Modeling Approach*. John Wiley & Sons, Chichester. DOI: 10.1002/9781118957813.

Cheung, M.W.-L., Chan, W., 2005. *Meta-Analytic Structural Equation Modeling: A Two-Stage Approach*. Psychological Methods 10(1), 40-64. DOI: 10.1037/1082-989X.10.1.40.

Cheung, M.W.-L., Chan, W., 2009. *A Two-Stage Approach to Synthesizing Covariance Matrices in Meta-Analytic Structural Equation Modeling*. Structural Equation Modeling 16(1), 28-52. DOI: 10.1080/10705510802561295.

Chidambaran, N.K., Fernando, C.S., Spindt, P.A., 2001. *Credit enhancement through financial engineering: Freeport McMoRan's gold-denominated depositary shares*. Journal of Financial Economics 60(2-3), 487-528. DOI: 10.1016/S0304-405X(01)00051-4.

Chod, J., Rudi, N., Van Mieghem, J.A., 2010. *Operational Flexibility and Financial Hedging: Complements or Substitutes?* Management Science 56(6), 1030-1045. DOI: 10.1287/mnsc.1090.1137.

Choi, J.J., Mao, C.X., Upadhyay, A.D., 2013. *Corporate Risk Management under Information Asymmetry*. Journal of Business Finance and Accounting 40(1), 239-271. DOI: 10.1111/jbfa.12008.

Chordia, T., Goyal, A., Saretto, A., 2018. *p-hacking: Evidence from two million trading strategies*. Swiss Finance Institute Research Paper No. 17-37.

Chou, S., Lai, C.-H., 2013. *Overconfident CEOs, Hedging and Performance*. Working Paper at National Yunlin University of Science & Technology.

Christensen, G., Miguel, E., 2018. *Transparency, Reproducibility, and the Credibility of Economics Research*. American Economic Review 56(3), 920-980. DOI: 10.1257/jel.20171350.

Christoffersen, P.F., 2003. *Elements of Financial Risk Management*. Academic Press, Amsterdam.

Chung, K.H., Pruitt, S.W., 1994. *A simple approximation of Tobin's Q*. Financial Management 23(3), 70-74. DOI: 10.2307/3665623.

Clark, E., Judge, A., 2009. *Foreign Currency Derivatives versus Foreign Currency Debt and the Hedging Premium*. European Financial Management 15(3), 606-642. DOI: 10.1111/j.1468-036X.2007.00431.x.

Clark, E., Judge, A., Mefteh, S., 2007. *Corporate Hedging with Forein Currency Derivatives and Firm Value*. Paper presented at European Financial Management Association Annual Meeting 2007, Vienna.

Clark, E., Mefteh, S., 2010. *Foreign Currency Derivatives Use, Firm Value and the Effect of the Exposure Profile: Evidence from France*. International Journal of Business 15(2), 183-196.

Coggin, D.T., Fabozzi, F.J., Rahman, S., 1993. *The Investment Performance of U.S. Equity Pension Fund Managers: An Empirical Investigation*. The Journal of Finance 48(3), 1039-1055. DOI: 10.2307/2329025.

Coggin, D.T., Hunter, J.E., 1983. *Problems in Measuring the Quality of Investment Information: The Perils of the Information Coefficient*. Financial Analysts Journal 39(3), 25-33. DOI: 10.2469/faj.v39.n3.25.

Coggin, D.T., Hunter, J.E., 1987. *A meta-analysis pricing "risk" factors in APT*. The Journal of Portfolio Management 14(1), 35-38. DOI: 10.3905/jpm.1987.409122.

Coggin, D.T., Hunter, J.E., 1993. *A Meta-Analysis of Mutual Fund Perfomance*. Review of Quantitative Finance and Accounting 3(2), 189-201. DOI: 10.1007/BF02407005.

Cohen, J., 1988. *Statistical Power Analysis for the Behavior Sciences*, 2nd ed. Lawrence Erlbaum Associates, Hillsdale. DOI: 10.4324/9780203771587.

Combs, J.G., 2010. *Big Samples and Small Effects: Let's Not Trade Relevance and Rigor for Power*. Academy of Management Journal 53(1), 9-13. DOI: 10.5465/amj.2010.48036305.

Combs, J.G., Crook, T.R., Rauch, A., 2019. *Meta-Analytic Research in Management: Contemporary Approaches, Unresolved Controversies, and Rising Standards*. Journal of Management Studies 56(1), 1-18. DOI: 10.1111/joms.12427.

Dahlquist, M., Pinkowitz, L., Stulz, R.M., Williamson, R., 2003. *Corporate Governance and the Home Bias*. The Journal of Financial and Quantitative Analysis 38(1), 87-110. DOI: 10.2307/4126765.

Daily, C.M., Certo, S.T., Dalton, D.R., Roengpitya, R., 2003. *IPO Underpricing: A Meta-Analysis and Research Synthesis*. Entrepreneurship Theory and Practice 27(3), 271-295. DOI: 10.1111/1540-8520.t01-1-00004.

Dalton, D.R., Daily, C.M., Trevis Certo, S., 2003. *Meta-Analyses of Financial and Equity: Fusion or Confusion?* The Academy of Management Journal 46(1), 13-26. DOI: 10.2307/30040673.

Dao, B.T.T., Ta, T.D.N.T., 2020. *A meta-analysis: Capital structure and firm performance*. Journal of Economics and Development 22(1), 111-129. DOI: 10.1108/JED-12-2019-0072.

Datta, D.K., Pinches, G.E., Narayanan, V.K., 1992. *Factors Influencing Wealth Creation from Mergers and Acquisitions: A Meta-Analysis*. Strategic Management Journal 13(1), 67-84. DOI: 10.1002/smj.4250130106.

de Batz, L., Kočenda, E., 2021. *Financial Crime and Punishment: A Meta-Analysis*. Working Paper at Charles University.

De La Cruz, A., Median, A., Tang, Y., 2019. *Owners of the World's Listed Companies*. OECD Capital Markets Series.

de Mooij, R.A., 2011. *The Tax Elasticity of Corporate Debt: A Synthesis of Size and Variations*. Working Paper IMF WP/11/95.

DeMarzo, P.M., Duffie, D., 1991. *Corporate Financial Hedging with Proprietary Information*. Journal of Economic Theory 53(2), 261-286. DOI: 10.1016/0022-0531(91)90156-X.

DeMarzo, P.M., Duffie, D., 1995. *Corporate Incentives for Hedging and Hedge Accounting*. The Review of Financial Studies 8(3), 743-771. DOI: 10.1093/rfs/8.3.743.

DerSimonian, R., Laird, N., 1986. *Meta-Analysis in Clinical Trials*. Control Clinical Trials 7(3), 177-188. DOI: 10.1016/0197-2456(86)90046-2.

Dhanani, A., Fifield, S., Helliar, C., Stevenson, L., 2007. *Why UK companies hedge interest rate risk*. Managerial Finance 24(1), 72-90. DOI: 10.1108/10867370710737391.

Dionne, G., Chun, O.M., Triki, T., 2013. *Risk Management and Corporate Governance: The Importance of Independence and Financial Knowledge*. HEC Montreal Working Paper No. 05-03. DOI: 10.2139/ssrn.730743.

Disatnik, D., Duchin, R., Schmidt, B., 2014. *Cash Flow Hedging and Liquidity Choices*. Review of Finance 18(2), 715-748. DOI: 10.1093/rof/rft006.

Döhring, B., 2008. *Hedging and invoicing strategies to reduce exchange rate exposure: a euro-area perspective*. European Economy, Economic Papers 299. DOI: 10.2765/22092.

Dolde, W., Mishra, D.R., 2007. *Firm Complexity and FX Derivatives Use*. Quarterly Journal of Business and Economics 46(4), 3-22.

dos Santos, R.B., Lima, F.G., Gatsios, R.C., de Almeida, R.B., 2017. *Risk management and value creation: new evidence for Brazilian non-financial companies*. Applied Economics 49(58), 5815-5827. DOI: 10.1080/00036846.2017.1343451.

Doucouliagos, C., 2011. *How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics*, Economics Series at Deakin University, Working Paper 05.

Doucouliagos, C., Stanley, T.D., 2013. *Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity*. Journal of Economic Surveys 27(2), 316-339. DOI: 10.1111/j.1467-6419.2011.00706.x.

Doucouliagos, C., Stanley, T.D., Giles, M., 2012. *Are estimates of the value of a statistical life exaggerated?* Journal of Health Economics 31(1), 197-206. DOI: 10.1016/j.jhealeco.2011.10.001.

Doucouliagos, H., Laroche, P., 2009. *Unions and Profits: A Meta-Regression Analysis*. Industrial Relations: A Journal of Economy and Society 48(1), 146-184. DOI: 10.1111/j.1468-232X.2008.00549.x.

Doucouliagos, H., Stanley, T.D., 2009. *Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis*. British Journal of Industrial Relations 47(2), 406-428. DOI: 10.1111/j.1467-8543.2009.00723.x.

Duan, J., Das, K.K., Meriluoto, L., Reed, R.W., 2020. *Estimating the effect of spillovers on exports: a meta-analysis*. Review of World Economics 156, 219-249. DOI: 10.1007/s10290-020-00377-z.

Duran, P., Kammerlander, N., van Essen, M., Zellweger, T., 2015. *Doing More with Less: Innovation Input and Output in Family Firms*. Academy of Management Journal 59(4), 1224-1264. DOI: 10.5465/amj.2014.0424.

Efendic, A., Pugh, G., Adnett, N., 2011. *Institutions and economic performance: A meta-regression analysis*. European Journal of Political Economy 27(3), 586-599. DOI: 10.1016/j.ejpoleco.2010.12.003.

Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. *Bias in Meta-Analysis Detected by a Simple, Graphical Test*. British Medical Journal 315(7109), 629-634. DOI: 10.1136/bmj.315.7109.629.

Eicher, T.S., Papageorgiou, C., Raftery, A.E., 2011. *Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants*. Journal of Applied Econometrics 26(1), 30-55. DOI: 10.1002/jae.1112.

Eisend, M., 2017. *Meta-Analysis in Advertising Research*. Journal of Advertising 46(1), 21-35. DOI: 10.1080/00913367.2016.1210064.

Elsawaf, N., 2005. *Essays on Currency Risk Management*. Doctoral thesis at Old Dominion University.

Engelen, P.-J., Heugens, P., van Essen, M., Turturea, R., Bailey, N., 2020. *The impact of stakeholders' temporal orientaton on short- and long-term IPO outcomes: A meta-analysis*. Long Range Planning 53(2), 101853. DOI: 10.1016/j.lrp.2018.10.003.

Eshgi, K., 2021. *Are sports sponsorship announcements good news for shareholders? A meta-analysis*. International Journal of Research in Marketing. DOI: 10.1016/j.ijresmar.2021.04.003.

Faulkender, M., 2005. *Hedging or Market Timing? Selecting the Interest Rate Exposure of Corporate Debt*. The Journal of Finance 60(2), 931–962. DOI: 10.1111/j.1540-6261.2005.00751.x.

Fauver, L., Naranjo, A., 2010. *Derivative Usage and Firm Value: The Influence of Agency Costs and Monitoring Problems*. Journal of Corporate Finance 16(5), 719-735. DOI: 10.1016/j.jcorpfin.2010.09.001.

Feld, L.P., Heckemeyer, J.H., Overesch, M., 2013. *Capital Structure Choice and Company Taxation: A Meta-Study*. Journal of Banking and Finance 37(8), 2850-2866. DOI: 10.1016/j.jbankfin.2013.03.017.

Feldkircher, M., Zeugner, S., 2009. *Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging*, IMF Working Paper 09/202.

Ferguson, C.J., Brannick, M.T., 2012. *Publication Bias in Psychological Science: Prevalence, Methods for Identifying and Controlling, and Implications for the Use of Meta-Analyses*. Psychological Methods 17(1), 120-128. DOI: 10.1037/a0024445.

Fernández, C., Ley, E., Steel, M.F.J., 2001. *Benchmark priors for Bayesian model averaging*. Journal of Econometrics 100(2), 381-427. DOI: 10.1016/S0304-4076(00)00076-2.

Fernau, E., Hirsch, S., 2019. *What drives dividend smoothing? A meta regression analysis of the Lintner model*. International Review of Financial Analysis 61, 255-273. DOI: 10.1016/j.irfa.2018.11.011.

Fidrmuc, J., Lind, R., 2020. *Macroeconomic impact of Basel III: Evidence from a meta-analysis*. Journal of Banking and Finance 112, 105359. DOI: 10.1016/j.jbankfin.2018.05.017.

Fok, R.C.W., Carroll, C., Chiou, M.C., 1997. *Determinants of Corporate Hedging and Derivatives: A Revisit*. Journal of Economics and Business 49(6), 569-585. DOI: 10.1016/S0148-6195(97)00040-4.

Frooman, J., 1997. *Socially Irresponsible and Illegal Behavior and Shareholder Wealth*. Business and Society 36(3), 221-249. DOI: 10.1177/000765039703600302.

Froot, K.A., 1989. *Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data*. The Journal of Financial and Quantitative Analysis 24(3), 333-355. DOI: 10.3386/t0062.

Froot, K.A., Scharfstein, D.S., Stein, J.C., 1993. *Risk Management: Coordinating Corporate Investment and Financing Policies*. The Journal of Finance 48(5), 1629-1658. DOI: 10.1111/j.1540-6261.1993.tb05123.x.

Furno, M., 1996. *Small sample behavior of a robust heteroskedasticity consistent covariance matrix estimator*. Journal of Statistical Computation & Simulation 54, 115-128. DOI: 10.1080/00949659608811723.

García-Meca, E., Sánchez-Ballesta, J.P., 2006. *Influences on financial analyst forecast errors: A meta-analysis*. International Business Review 15, 29-52. DOI: 10.1016/j.ibusrev.2005.12.003.

Garg, A.X., Hackam, D., Tonelli, M., 2008. *Systematic Review and Meta-analysis: When One Study Is Just not Enough*. Clinical Journal of the American Society of Nephrology 3(1), 253-260. DOI: 10.2215/CJN.01430307.

Géczy, C.C., Minton, B.A., Schrand, C., 1997. *Why Firms Use Currency Derivatives*. The Journal of Finance 52(4), 1323-1354. DOI: 10.1111/j.1540-6261.1997.tb01112.x.

George, E.I., 2010. *Dilution priors: Compensating for model space redundancy*. Institute of Mathematical Statistics.

Geyer-Klingeberg, J., Hang, M., Rathgeber, A., 2019. *What drives financial hedging? A meta-regression analysis of corporate hedging determinants*. International Review of Financial Analysis 61, 203-221. DOI: 10.1016/j.irfa.2018.11.006.

Geyer-Klingeberg, J., Hang, M., Rathgeber, A., 2021. *Corporate financial hedging and firm value: A meta-analysis*. The European Journal of Finance 27(6), 461-485. DOI: 10.1080/1351847X.2020.1816559.

Geyer-Klingeberg, J., Hang, M., Rathgeber, A., Stöckl, S., 2018a. *What Do We Really Know About Corporate Hedging? A Meta-Analytical Study*. Business Research 11(1), 1-31. DOI: 10.1007/s40685-017-0052-0.

Geyer-Klingeberg, J., Hang, M., Walter, M., Rathgeber, A., 2018b. *Do Stock Markets React to Soccer Games? A Meta-Regression Analysis*. Applied Economics 50(19), 2171-2189. DOI: 10.1080/00036846.2017.1392002.

Geyskens, I., Krishnan, R., Steenkamp, J.-B.E.M., Cunha, P.V., 2009. *A Review and Evaluation of Meta-Analysis Practices in Management Research*. Journal of Management 35(2), 393-419. DOI: 10.1177/0149206308328501.

Giambona, E., Graham, J.R., Harvey, C.R., Bodnar, G.M., 2018. *The Theory and Practice of Corporate Risk Management: Evidence from the Field*. Financial Management 47(4), 783-832. DOI: 10.1111/fima.12232.

Gilje, E.P., Taillard, J.P., 2017. *Does Hedging Affect Firm Value? Evidence from a Natural Experiment*. The Review of Financial Studies 30(12), 4083–4132. DOI: 10.1093/rfs/hhx069.

Glass, G.V., 1976. *Primary, Secondary, and Meta-Analysis of Research*. Educational Researcher 5(10), 3-8.

Gleason, K.C., Kim, Y.S., Mathur, I., 2005. *The Operational and Financial Hedging Strategies of U.S. High Technology Firms*. Working Paper at the Southern Illinois University.

Gómez-Déniz, E., Calderín-Ojeda, E., 2011. *The discrete Lindley distribution: properties and applications*. Journal of Statistical Computation and Simulation 81(11), 1405-1416. DOI: 10.1080/00949655.2010.487825.

Gómez-González, J.E., León Rincón, C.E., Leiton Rodríguez, K.J., 2012. *Does the Use of Foreign Currency Derivatives Affect Firms' Market Value? Evidence from Colombia*. Emerging Markets and Trade 48(1), 50-66. DOI: 10.2753/REE1540-496X480403.

Görg, H., Strobel, E., 2001. *Multinational Companies and Productivity Spillovers: A Meta-Analysis*. The Economic Journal 111(475), 723-739. DOI: 10.1111/1468-0297.00669.

Gormley, T.A., Matsa, D.A., 2014. *Common Errors: How to (and Not to) Control for Unobserved Heterogeneity*. The Review of Financial Studies 27(2), 617-661. DOI: 10.1093/rfs/hht047.

Greene, W.H., 2011. *Econometric Analysis*, 7th ed. Prentice Hall, New York.

Greenhalgh, T., Thorne, S., Malterud, K., 2018. *Time to challenge the spurious hierarchy of systematic over narrative reviews?* European Journal of Clinical Investigation 48(6), e12931. DOI: 10.1111/eci.12931.

Grewal, D., Puccinelli, N., Monroe, K.B., 2018. *Meta-analysis: integrating accumulated knowledge*. Journal of the Academy of Marketing Science 46(1), 9-30. DOI: 10.1007/s11747-017-0570-5.

Grežo, M., 2020. *Overconfidence and financial decision-making: a meta-analysis*. Review of Behavioral Finance 13(3), 276-296. DOI: 10.1108/RBF-01-2020-0020.

Gurevitch, J., Koricheva, J., Nakagawa, S., Stewart, G., 2018. *Meta-analysis and the science of research synthesis*. Nature 5555, 175-182. DOI: 10.1038/nature25753.

Hackshaw, A., 2008. *Small studies: strengths and limitations*. European Respiratory Journal 32(5), 1141-1143. DOI: 10.1183/09031936.00136408.

Hagelin, N., Holmén, M., Knopf, J.D., Pramborg, B., 2007. *Managerial Stock Options and the Hedging Premium*. European Financial Management 13(4), 721-741. DOI: 10.1111/j.1468-036X.2007.00380.x.

Hall, J.A., Rosenthal, R., 1991. *Testing for moderator variables in meta-analysis: Issues and methods*. Communication Monographs 58(4), 437-448. DOI: 10.1080/03637759109376240.

Hang, M., Geyer-Klingeberg, J., Rathgeber, A., Alasseur, C., Wichmann, L., 2021a. *Interaction Effects of Corporate Hedging Activities for a Multi-Risk Exposure: Evidence from a Quasi-Natural Experiment*. Review of Quantitative Finance and Accounting 56, 789-818. DOI: 10.1007/s11156-020-00909-6.

Hang, M., Geyer-Klingeberg, J., Rathgeber, A., Stöckl, S., 2018. *Measurement matters – A meta-study of the determinants of corporate capital structure*. Quarterly Review of Economics and Finance 68, 211-225. DOI: 10.1016/j.qref.2017.11.011.

Hang, M., Geyer-Klingeberg, J., Rathgeber, A., Stöckl, S., 2021b. *Rather complements than substitutes: Firm value effects of capital structure and financial hedging decisions*. International Journal of Finance and Economics 26(4), 4895-4917. DOI: 10.1002/ijfe.2045.

Hankins, K.W., 2009. *How Do Financial Firms Manage Risk? Unraveling the Interaction of Financial and Operational Hedging*. Management Science 57(12), 2197-2212. DOI: 10.1287/mnsc.1090.1068.

Hanousek, J., Kočenda, E., 2011. *Direct and indirect effects of FDI in emerging European markets: A survey and meta-analysis*. Economic Systems 35(3), 301-322. DOI: 10.1016/j.ecosys.2010.11.006.

Hansen, C., Block, J., 2021. *Public family firms and capital structure: A meta-analysis*. Corporate Governance: An International Review 29(3), 297-319. DOI: 10.1111/corg.12354.

Hariton, E., Locascio, J.J., 2018. *Randomised controlled trials—the gold standard for effectiveness research*. BJOG 125(13), 1716. DOI: 10.1111/1471-0528.15199.

Harrison, J.S., Banks, G.C., Pollack, J.M., O'Boyle, E.H., Short, J., 2017. *Publication Bias in Strategic Management Research*. Journal of Management 43(2), 400-425. DOI: 10.1177/0149206314535438.

Harvey, C.R., 2017. *Presidential Address: The Scientific Outlook in Financial Economics*. The Journal of Finance 72(4), 1399-1440. DOI: 10.1111/jofi.12530.

Harvey, C.R., Liu, Y., Zhu, H., 2016. *... and the Cross-Section of Expected Returns*. The Review of Financial Studies 29(1), 5-68. DOI: 10.1093/rfs/hhv059.

Haushalter, D.G., 2000. *Financing Policy, Basis Risk, and Corporate Hedging: Evidence from Oil and Gas Producers*. The Journal of Finance 55(1), 107-152. DOI: 10.1111/0022-1082.00202.

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

Havranek, T., Horvath, R., Zeynalov, A., 2016. *Natural Resources and Economic Growth: A Meta-Analysis*. World Development 88, 134-151. DOI: 10.1016/j.worlddev.2016.07.016.

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), 234-2444. DOI: 10.1016/j.jinteco.2011.07.004.

Havranek, T., Irsova, Z., 2017. *Do Borders Really Slash Trade? A Meta-Analysis*. IMF Economic Review 65, 365-396. DOI: 10.1057/s41308-016-0001-5.

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

Havranek, T., Irsova, Z., Janda, K., Zilberman, D., 2015b. *Selective Reporting and the Social Cost of Carbon*. Energy Economics 51, 394-406. DOI: 10.1016/j.eneco.2015.08.009.

Havranek, T., Rusnak, M., Sokolova, A., 2017. *Habit Formation in Consumption: A Meta-Analysis*. European Economic Review 95, 142-167. DOI: 10.1016/j.euroecorev.2017.03.009.

Havranek, T., Stanley, T.D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, R.W., Rost, K., van Aert, R.C.M., 2020. *Reporting Guidelines for Meta-Analysis in Economics*. Journal of Economic Surveys 34(3), 469-475. DOI: 10.1111/joes.12363.

Hayes, A.F., Cai, L., 2007. *Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation*. Behavior Research Methods 39(4), 709-722. DOI: 10.3758/BF03192961.

Heckman, J.J., 1979. *Sample Selection Bias as a Specification Error*. Econometrica 47(1), 153-161. DOI: 10.2307/1912352.

Hedges, L.V., Olkin, I., 1985. *Statistical Methods for Meta-Analysis*. Academic Press, London. DOI: 10.1016/C2009-0-03396-0.

Hedges, L.V., Tipton, E., Johnson, M.C., 2010. *Robust variance estimation in meta-regression with dependent effect size estimates*. Research Synthesis Methods 1(1), 39-65. DOI: 10.1002/jrsm.5.

Hedges, L.V., Vevea, J.L., 1998. *Fixed- and Random-Effects Models in Meta-Analysis*. Psychological Methods 3(4), 486-504.

Hentschel, L., Kothari, S.P., 2001. *Are Corporations Reducing or Taking Risks with Derivatives?* The Journal of Financial and Quantitative Analysis 36(1), 93-118. DOI: 10.2307/2676199.

Heugens, P.P.M.A.R., van Essen, M., Oosterbeek, J.H., 2009. *Meta-analyzing ownership concentration and firm performance in Asia: Towards a more fine-grained understanding*. Asia Pacific Journal of Management 26(3), 481-512. DOI: 10.1007/s10490-008-9109-0.

Higgins, J.P.T., Thompson, S.G., 2002. *Quantifying Heterogeneity in a Meta-Analysis*. Statistics in Medicine 21(11), 1539-1558. DOI: 10.1002/sim.1186.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. *Measuring inconsistency in meta-analyses*. British Medical Journal 327(7414), 557-560. DOI: 10.1136/bmj.327.7414.557.

Hinne, M., Gronau, Q.F., van den Bergh, D., Wagenmakers, E.-J., 2020. *A Conceptual Introduction to Bayesian Model Averaging*. Advances in Methods and Practices in Psychological Science 3(2), 200-215. DOI: 10.1177/2515245919898657.

Ho, T.H., Nguyen, D.T., Ngo, T., Le, T.D.Q., 2021. *Efficiency in Vietnamese Banking: A Meta-Regression Analysis Approach*. International Journal of Financial Studies 9(3), 1-15. DOI: 10.3390/ijfs9030041.

Hoberg, G., Moon, K., 2017. *Offshore Activities and Financial vs Operational Hedging*. Journal of Financial Economics 125(2), 217-244. DOI: 10.1016/j.jfineco.2017.05.003.

Holderness, C.G., 2018. *Equity issuances and agency costs: The telling story of shareholder approval around the world*. Journal of Financial Economics 129(3), 415-439. DOI: 10.1016/j.jfineco.2018.06.006.

Hopkin, P., 2018. *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Risk Management*, 5th ed. Kogan Page Publishers, London.

Hu, L.-t., Bentler, P.M., 1999. *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives*. Structural Equation Modeling: A Multidisciplinary Journal 6(1), 1-55. DOI: 10.1080/10705519909540118.

Hubler, J., Louargant, C., Laroche, P., Ory, J.-N., 2019. *How do rating agencies' decisions impact stock markets? A meta-analysis*. Journal of Economic Surveys 33(4), 1173-1198. DOI: 10.1111/joes.12317.

Hunter, J.E., Schmidt, F.L., 1990. *Methods of meta-analysis: Correcting error and bias in research findings*, 1st ed. Sage, Newbury Park.

Hunter, J.E., Schmidt, F.L., 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, 2nd ed. Sage Publications, Thousand Oaks, CA.

International Organization for Standardization, 2018. *Risk management — Guidelines (ISO Standard 31000:2018)*, https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en, Accessed: 2022/01/02.

Ioannidis, J., 2016. *The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses*. The Milbank Quarterly 94(3), 485-514. DOI: 10.1111/1468-0009.12210.

Ioannidis, J.P.A., 2005. *Why Most Published Research Findings Are False*. PLoS Med 2(8), 696-701. DOI: 10.1371/journal.pmed.0020124.

Ioannidis, J.P.A., Stanley, T.D., Doucouliagos, H., 2017. *The Power of Bias in Economics Research*. The Economic Journal 127(605), F236-F265. DOI: 10.1111/ecoj.12461.

Irsova, Z., Havranek, T., 2012. *Measuring Bank Efficiency: A Meta-Regression Analysis*. Prague Economic Papers 19(4), 307-328. DOI: 10.18267/j.pep.379.

Iwasaki, I., Kočenda, E., 2017. *Are some owners better than others in Czech privatized firms? Even meta-analysis can't make us perfectly sure*. Economic Systems 41(4), 537-568. DOI: 10.1016/j.ecosys.2017.06.001.

Iwasaki, I., Mizobata, S., 2018. *Post-privatization ownership and firm performance: A large meta-analysis of the transition literature*. Annals of Public and Cooperative Economics 89(2), 263-322. DOI: 10.1111/apce.12180.

Iwasaki, I., Mizobata, S., 2020. *Ownership Concentration and Firm Performance in European Emerging Economies: A Meta-Analysis*. Emerging Markets and Trade 56(1), 32-67. DOI: 10.1080/1540496X.2018.1530107.

Iwasaki, I., Mizobata, S., Muravyev, A., 2018. *Ownership dynamics and firm performance in an emerging economy: a meta-analysis of the Russian literature*. Post-Communist Economies 30(3), 290-333. DOI: 10.1080/14631377.2018.1442036.

Jak, S., 2015. *Meta-Analytic Structural Equation Modelling*. Springer, Cham. DOI: 10.1007/978-3-319-27174-3.

Jankensgård, H., 2015a. *Does Centralisation of FX Derivatives Usage Imapct Firm Value*. European Financial Management 21(2), 309-332. DOI: 10.1111/j.1468-036X.2013.12014.x.

Jankensgård, H., 2015b. *Walking the Walk: Selective hedging and inside ownership*. Working Paper at Lund University.

Jankensgård, H., Hoffmann, K., Rahmat, D., 2014. *Derivative Usage, Risk Disclosure, and Firm Value*. Journal of Accounting and Finance 14(5), 159-174.

Jin, Y., Jorion, P., 2006. *Firm Value and Hedging: Evidence from U.S. Oil and Gas Producers*. The Journal of Finance 61(2), 893-919. DOI: 10.1111/j.1540-6261.2006.00858.x.

Jin, Y., Jorion, P., 2007. *Does Hedging Increase Firm Value? Evidence from the Gold Mining Industry*. Working Paper at California State University.

Johnson, J.L., Mullen, B., Salas, E., 1995. *Comparison of Three Major Meta-Analytic Approaches*. Journal of Applied Psychology 80(1), 94-106. DOI: 10.1037/0021-9010.80.1.94.

Jorge, M.J.d.S., Augusto, M.G.A., 2012. *The Hedging Value Conditional to Corporate Governance: Evidence from European Nonfinancial Firms*. Paper presented at XXXVI Conferência da EnANPAD 2012, Rio de Janeiro.

Joshi, A., Son, J., Roh, H., 2015. *When can women close the gap? A meta-analytic test of sex differences in performance and rewards*. Academy of Management Journal 58(5), 1516-1545. DOI: 10.5465/amj.2013.0721.

Judge, A., 2007. *Why Do Firms Hedge? A Review of the Evidence*, in: McCombie, J., González, C.R. (Eds.), Issues in Finance and Monetary Policy. Palgrave Macmillan, Basingstoke, 128-152. DOI: 10.1057/9780230801493.

Kaldor, N., 1966. *Marginal Productivity and the Macro-Economic Theories of Distribution: Comment on Samuelson and Modigliani*. The Review of Economic Studies 33(4), 309-319. DOI: 10.2307/2974428.

Kapitsinas, S., 2008. *The Impact of Derivatives Usage on Firm Value: Evidence from Greece*. MPRA Working Paper No. 10947.

Kaufmann, D., Kraay, A., Mastruzzi, M., 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues*, World Bank Policy Research Working Paper No. 5430.

Keef, S.P., Roush, M.L., 2007. *A meta-analysis of the international evidence of cloud cover on stock returns*. Review of Accounting and Finance 6(3), 324-338. DOI: 10.1108/14757700710778045.

Kelley, K., Preacher, K.J., 2012. *On Effect Size*. Psychological Methods 17(2), 137-152. DOI: 10.1037/a0028086.

Kepes, S., McDaniel, M.A., Brannick, M.T., Banks, G.C., 2013. *Meta-Analytic Reviews in the Organizational Sciences: Two Meta-analytic Schools on the Way to MARS (the Meta-Analytic Reporting Standards).* Journal of Business and Psychology 28(2), 123-143. DOI: 10.1007/s10869-013-9300-2.

Khediri, K.B., 2010. *Do Investors Really Value Derivative Use? Empirical Evidence from France*. The Journal of Risk Finance 11(1), 62-74. DOI: 10.1108/15265941011012688.

Khediri, K.B., Folus, D., 2010. *Does Hedging Increase Firm Value? Evidence from French Firms*. Applied Economics Letters 17(10), 995-998. DOI: 10.1080/17446540802599697.

Khlif, H., Chalmers, K., 2015. *A review of meta-analytic research in accounting*. Journal of Accounting Literature 35, 1-27. DOI: 10.1016/j.acclit.2015.09.001.

Kim, C., Pantzalis, C., Park, J.C.P., 2014a. *Do Family Owners Use Firm Hedging Policy to Hedge Personal Undiversified Wealth Risk?* Financial Management 43(2), 415-444. DOI: 10.1111/fima.12021.

Kim, C.-S., 2019. *Can Socially Responsible Investments Be Compatible with Financial Performance? A Meta-analysis*. Asia-Pacific Journal of Financial Studies 48(1). DOI: 10.1111/ajfs.12244.

Kim, H.T., Papanastassiou, M., Nguyen, Q., 2017. *Multinationals and the impact of corruption on financial derivatives use and firm value: Evidence from East Asia*. Journal of Multinational Financial Management 39, 39-59. DOI: 10.1016/j.mulfin.2017.02.001.

Kim, J.H., Doucouliagos, H., Stanley, T.D., 2014b. *Market Efficiency in Asian and Australasian Stock Markets: A Fresh Look at the Evidence*, in: Chevallier, J., Goutte, S., Guerreiro, D., Saglio, S., Sanhaji, B. (Eds.), International Financial Markets Volume 1. Routledge, New York, 382-420. DOI: 10.4324/9781315162775.

Kim, J.H., Ji, P.I., 2015. *Significance testing in empirical finance: A critical review and assessment*. Journal of Empirical Finance 34, 1-14. DOI: 10.1016/j.jempfin.2015.08.006.

Kim, Y.S., Mathur, I., Nam, J., 2006. *Is Operational Hedging a Substitute for or a Complement to Financial Hedging?* Journal of Corporate Finance 12(4), 834-853. DOI: 10.1016/j.jcorpfin.2005.09.003.

King, D.R., Dalton, D.R., Daily, C.M., Covin, J.G., 2004. *Meta-analyses of post-acquisition performance: indications of unidentified moderators*. Strategic Management Journal 25(2), 187-200. DOI: 10.1002/smj.371.

Klein, P.-O., 2017. *Do Shareholders Value Bond Offerings? A Meta-Analysis*. Working Paper Université de Strasbourg 2017-04.

Knopf, J.D., Nam, J., Thornton, J., 2002. *The Volatility and Price Sensitivities of Managerial Stock Option Portfolios and Corporate Hedging*. The Journal of Finance 57(2), 801-813. DOI: 10.1111/1540-6261.00442.

Kočenda, E., Cerný, A., 2015. *Elements of Time Series Econometrics - An Applied Approach*, 3rd ed. Karolinum Press, Prague.

Kočenda, E., Iwasaki, I., 2022. *Bank survival around the world: A meta-analytic review*. Journal of Economic Surveys 36(1), 108-156. DOI: 10.1111/joes.12451.

Koetse, M.J., Florax, R.J.G.M., de Groot, H.L.F., 2005. *Correcting for Primary Study Misspecifications in Meta-Analysis*. Tinbergen Institute Discussion Paper TI 2005-029/3.

Koetse, M.J., Florax, R.J.G.M., de Groot, H.L.F., 2010. *Consequences of Effect Size Heterogeneity for Meta-Analysis: A Monte Carlo Study*. Statistical Methods and Application 19(2), 217-236. DOI: 10.1007/s10260-009-0125-0.

Korkeamäki, T., Liljeblom, E., Pfister, M., 2016. *Airline fuel hedging and management ownership*. The Journal of Risk Finance 17(5), 492-509. DOI: 10.1108/JRF-06-2016-0077.

Krause, T.A., Tse, Y., 2016. *Risk management and firm value: recent theory and evidence*. International Journal of Accounting & Information Management 24(1), 56-81. DOI: 10.1108/IJAIM-05-2015-0027.

Kwong, L.C., 2016. *How corporate derivatives use impact firm performance?* Pacific-Basin Finance Journal 40, 102-114. DOI: 10.1016/j.pacfin.2016.10.001.

Kysucky, V., Norden, L., 2016. *The Benefits of Relationship Lending in a Cross-Country Context: A Meta-Analysis*. Management Science 62(1), 90-110. DOI: 10.1287/mnsc.2014.2088.

La Rocca, M., Montalto, F., La Rocca, T., Staglianò, R., 2017. *The effect of ownership on firm value: a meta-analysis*. Economics Bulletin 37(4), 2324-2353.

La Rocca, T., 2021. *Do prestigious underwriters shape IPO pricing? A meta-analytic review*. Review of Managerial Science 15(3), 573-609. DOI: 10.1007/s11846-019-00356-1.

Landis, R.S., 2013. *Successfully Combining Meta-analysis and Structural Equation Modeling: Recommendations and Strategies*. Journal of Business and Psychology 28(3), 251-261. DOI: 10.1007/s10869-013-9285-x.

Lazonick, W., O'Sullivan, M., 2000. *Maximizing shareholder value: a new ideology for corporate governance*. Economy and Society 29(1), 13-35. DOI: 10.1080/030851400360541.

Leamer, E.E., 1983. *Let's Take the Con Out of Econometrics*. The American Economic Review 73(1), 31-43.

Leandro, G., 2005. *Meta-analysis in Medical Research.The handbook for the understanding and practice of meta-analysis*, Malden, Mass. DOI: 10.1002/9780470994894.

Lee, D.D., Kirkpatrick-Husk, K., Madhavan, R., 2017. *Diversity in Alliance Portfolios and Performance Outcomes: A Meta-Analysis*. Journal of Management 43(5), 1472-1497. DOI: 10.1177/0149206314556316.

Lee, D.D., Madhavan, R., 2010. *Divestiture and Firm Performance: A Meta-Analysis*. Journal of Management 36(6), 1345-1371. DOI: 10.1177/0149206309360931.

Leland, H.E., 1998. *Agency Costs, Risk Management, and Capital Structure*. The Journal of Finance 53(4), 1213-1243. DOI: 10.1111/0022-1082.00051.

Lewellen, W.G., Badrinath, S.G., 1997. *On the measurement of Tobin's Q*. Journal of Financial Economics 44(1), 77-122. DOI: 10.1016/S0304-405X(96)00013-X.

Ley, E., Steel, M.F.J., 2009. *On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression*. Journal of Applied Econometrics 24(4), 651-674. DOI: 10.1002/jae.1057.

Li, H., Visaltanachoti, N., Luo, R., 2014. *Foreign Currency Derivatives and Firm Value: Evidence from New Zealand*. Journal of Financial Risk Management 3, 96-112. DOI: 10.4236/jfrm.2014.33010.

Lichter, A., Peichl, A., Siegloch, S., 2015. *The Own-Wage Elasticity of Labor Demand: A Meta-Regression Analysis*. European Economic Review 80, 94-119. DOI: 10.1016/j.euroecorev.2015.08.007.

Lievenbrück, M., Schmid, T., 2014. *Why do firms (not) hedge? Novel evidence on cultural influence*. Journal of Corporate Finance 25, 92-106. DOI: 10.1016/j.jcorpfin.2013.10.010.

Lin, C.-M., Philips, R.D., Smith, S.D., 2008. *Hedging, financing, and investment decisions: Theory and empirical tests*. Journal of Banking and Finance 32(8), 1556-1582. DOI: 10.1016/j.jbankfin.2007.11.014.

Lindner, T., Klein, F., Schmidt, S., 2018. *The effect of internationalization on firm capital structure: A meta-analysis and exploration of institutional contingencies*. International Business Review 27(6), 1238-1249. DOI: 10.1016/j.ibusrev.2018.05.005.

Loderer, C., Pichler, K., 2000. *Firms, do you know your currency risk exposure? Survey results*. Journal of Empirical Finance 7(3-4), 317-344. DOI: 10.1016/S0927-5398(00)00014-1.

Lookman, A.A., 2004. *Does Hedging Increase Firm Value? Evidence from Oil and Gas Producing Firms*. Working Paper at the Carnegie Mellon University.

Luo, R.H., 2016. *Foreign Currency Derivatives and Corporate Value: Evidence from China*. Paper presented at Australia-Middle East Conference on Business and Social Sciences 2016, Dubai.

MacKay, P., Moeller, S.B., 2007. *The Value of Corporate Risk Management*. The Journal of Finance 62(3), 1379-1419. DOI: 10.1111/j.1540-6261.2007.01239.x.

Magee, S., 2013. *Foreign Currency Hedging and Firm Value: A Dynamic Panel Approach*, in: Batten, J.A., MacKay, P., Wagner, N. (Eds.), Advances in Financial Risk Management. Palgrave Macmillan, London, 57-80. DOI: 10.1057/9781137025098_3.

Manchiraju, H., Pierce, S., Sridharan, S.S., 2014. *Do firms use derivatives for hedging or non-hedging purposes? Evidence based on SFAS 161 disclosures*. Working Paper at Northwestern University. DOI: 10.2139/ssrn.2417194.

Marami, A., Dubois, M., 2013. *Interest Rate Derivatives and Firm Value: Evidence from Mandatory versus Voluntary Hedging*. Working Paper at University of Neuchatel. DOI: 10.2139/ssrn.2336094.

Marano, V., Arregle, J.-L., Hitt, M.A., Spadafora, E., van Essen, M., 2016. *Home Country Institutions and the Internationalization-Performance Relationship: A Meta-Analytic Review*. Journal of Management 42(5). DOI: 10.1177/0149206315624963.

Marin-Martinez, F., Sanchez-Meca, J., 2010. *Weighting by Inverse Variance or by Sample Size in Random-Effects Meta-Analysis*. Educational and Psychological Measurement 70(1), 56-73. DOI: 10.1177/0013164409344534.

Markowitz, H., 1952. *Portfolio Selection*. The Journal of Finance 7(1), 77-91. DOI: 10.2307/2975974.

Marshall, A.P., Weetman, P., 2007. *Modelling Transparency in Disclosure: The Case of Foreign Exchange Risk Management*. Journal of Business Finance and Accounting 34(5-6), 705-739. DOI: 10.1111/j.1468-5957.2007.02007.x.

Matousek, J., Havranek, T., Irsova, Z., 2022. *Individual Discount Rates: A Meta-Analysis of Experimental Evidence*. Experimental Economics 25, 318-358. DOI: 10.1007/s10683-021-09716-9.

Mayers, D., Smith, C.W., 1982. *On the Corporate Demand for Insurance*. The Journal of Business 55(2), 281-296. DOI: 10.1086/296165.

McAfee, A., Brynjolfsson, E., 2012. *Big Data: The Management Revolution*. Harvard Business Review 90(10), 60-68.

McNeil, A.J., Rüdiger, F., Embrechts, P., 2015. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton.

Meredith, J.H., 2002. *The Corporate Use of Derivatives to Hedge Energy Price Risk*. Dissertation at the University of Mississippi.

Mian, S.L., 1996. *Evidence on Corporate Hedging Policy*. Journal of Financial and Quantitative Analysis 31(3), 419-439. DOI: 10.2307/2331399.

Mikolajewicz, N., Komarova, S.V., 2019. *Meta-Analytic Methodology for Basic Research: A Practical Guide*. Frontiers in Psychology 10, 1-20. DOI: doi.org/10.3389/fphys.2019.00203.

Mitton, T., 2019. *Methodological Variation in Empirical Corporate Finance*. Working Paper at Marriott School of Business, Brigham Young University.

Mitton, T., 2022. *Methodological Variation in Empirical Corporate Finance*. The Review of Financial Studies 35(2), 527-575. DOI: 10.1093/rfs/hhab030.

Modigliani, F., Miller, M.H., 1958. *The Cost of Capital, Corporation Finance and the Theory of Investment*. The American Economic Review 48(3), 261-297.

Mohammad, K., 2014. *Value of Hedging in US Airline Industry: A Perspective on Firm Value & Accounting Performance*. International Journal of Business and Management 1(1), 35-51.

Moral-Benito, E., 2012. *Bayesian posterior prediction and meta-analysis: an application to the value of travel time savings*. The Annals of Regional Science 48(3), 801-817. DOI: 10.1007/s00168-010-0407-3.

Moreno, S.G., Sutton, A.J., Stanley, T.D., Abrams, K.R., Peters, J.L., Cooper, N.J., 2009. *Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study*. BMC Medical Research Methodology 9, 1-17. DOI: 10.1186/1471-2288-9-2.

Morey, M.R., Yadav, S., 2018. *Documentation of the File Drawer Problem in Academic Finance Journals*. The Journal of Investing 27(1), 143-147. DOI: 10.3905/joi.2018.27.1.143.

Mosteller, F., Tukey, J.W., 1977. *Data Analysis and Regression*. Addison-Wesley, Reading.

Myers, S.C., 1977. *Determinants of Corporate Borrowing*. Journal of Financial Economics 5(2), 147-175. DOI: 10.1016/0304-405X(77)90015-0.

Nagendrababu, V., Dilokthornsakul, P., Jinatongthai, P., Veettil, S.K., Pulikkotil, S.J., Duncan, H.F., Dummer, P.M.H., 2020. *Glossary for systematic reviews and meta-analyses*. International Endodontic Journal 53(2), 232-249. DOI: 10.1111/iej.13217.

Nain, A., 2005. *Corporate Risk Management in an Industry Setting: An Empirical Investigation*. Paper presented at American Finance Association Annual Meeting 2005, Philadelphia.

Nance, D.R., Smith Jr, C.W., Smithson, C.W., 1993. *On the Determinants of Corporate Hedging*. The Journal of Finance 48(1), 267-284. DOI: 10.1111/j.1540-6261.1993.tb04709.x.

Nehrebecka, N., Dzik-Walczak, A., 2018. *The dynamic model of partial adjustment of the capital structure. Meta-analysis and a case of Polish enterprises*. Zbornik Radova Ekonomskog Fakultet au Rijeci 36(1), 53-79. DOI: 10.18045/zbefri.2018.1.55.

Nelson, J.P., Kennedy, P.E., 2009. *The Use and (Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment*. Environmental and Resource Economics 42(3), 345-377. DOI: 10.1007/s10640-008-9253-5.

Newey, W.K., West, K.D., 1987. *A simple, positive semi-definite, heteroscedastic and autocorrelation consistent covariance matrix*. Econometrica 55(3), 703-708. DOI: 10.2307/1913610.

Nguyen, H., Faff, R., 2007. *Are Financial Derivatives Really Value Enhancing? Australian Evidence*. Deakin University School Working Paper in Accounting/Finance Series 2007.

Nguyen, H., Faff, R., 2010. *Does the type of derivative instrument used by companies impact firm value?* Applied Economics Letters 17(7), 681-683. DOI: 10.1080/13504850802297822.

Nocco, B.W., Stulz, R.M., 2006. *Enterprise Risk Management: Theory and Practice*. Journal of Applied Corporate Finance 18(4), 8-20. DOI: 10.1111/j.1745-6622.2006.00106.x.

Noel, C., Farley, J.U., Hoenig, S., 1990. *Determinants of Financial Performance: A Meta-Analysis*. Management Science 36(10), 1143-1159. DOI: 10.1287/mnsc.36.10.1143.

Nova, M., Cerqueira, A., Brandao, E., 2015. *Hedging with Derivatives and Firm Value: Evidence for the nonfinancial firms listed on the London Stock Exchange*. FEP Working Paper Series Bo. 568.

Oxford English Dictionary, 2022. *Risk*, www.oed.de, Accessed: 2022/01/02.

Pae, C.-U., 2015. *Why Systematic Review rather than Narrative Review?* Psychiatry Investigation 12(3), 417-419. DOI: 10.4306/pi.2015.12.3.417.

Panaretou, A., 2014. *Corporate risk management and firm value: evidence from the UK market*. The European Journal of Finance 20(12), 1161-1186. DOI: 10.1080/1351847X.2013.766625.

Papenfuß, P., Schischke, A., Zender, V., Geyer-Klingeberg, J., Rathgeber, A., 2021. *Is the negative impact of interest rate changes on mineral commodity prices still valid? Evidence from a Meta-Regression Analysis*. Paper presented at the Meta-Analysis of Economics Research Network Colloquium 2021, Athens.

Papke, L.E., Wooldridge, J.M., 2005. *A computational trick for delta-method standard errors*. Economics Letters 86(3). DOI: 10.1016/j.econlet.2004.07.022.

Park, S., Beretvas, N.S., 2018. *Using Total Sample Size Weights in Meta-Analysis of Log-Odds Ratios*. The Journal of Experimental Education 87(3), 400-414. DOI: 10.1080/00220973.2018.1451295.

Pérez-González, F., Yun, H., 2013. *Risk Management and Firm Value: Evidence from Whether Derivatives*. The Journal of Finance 68(5), 2143-2176. DOI: 10.1111/jofi.12061.

Perfect, S.B., Wiles, K.W., 1994. *Alternative Constructions of Tobin's Q: An Empirical Comparison*. Journal of Empirical Finance 1(3-4), 313-341. DOI: 10.1016/0927-5398(94)90007-8.

Petersen, M.A., 2008. *Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches*. The Review of Financial Studies 22(1), 324-480. DOI: 10.1093/rfs/hhn053.

Phan, D., Nguyen, H., Faff, R., 2014. *Uncovering the asymmetric linkage between financial derivatives and firm value - The case of oil and gas exploration and production companies*. Energy Economics 45, 340-352. DOI: 10.1016/j.eneco.2014.07.018.

Phillips, A.L., 1995. *1995 Derivatives Practices and Instruments Survey*. Financial Management 24(2), 115-125. DOI: 10.2307/3665539.

Pierce, S., 2015. *Does the Accounting for Derivatives Affect Risk and Value?* Working Paper at Florida State University. DOI: 10.2139/ssrn.2685896.

Post, C., Byron, K., 2015. *Women on Boards and Firm Financial Performance: A Meta-Analysis*. Academy of Management Journal 58(5), 1546-1571. DOI: 10.5465/amj.2013.0319.

Pramborg, B., 2004. *Derivatives hedging, geographical diversification, and firm market value*. Journal of Multinational Financial Management 14(2), 117-133. DOI: 10.1016/j.mulfin.2003.07.002.

Prevost, A.K., Rose, L.C., Miller, G., 2000. *Derivatives Usage and Financial Risk Management in Large and Small Economies: A Comparative Analysis*. Journal of Business Finance and Accounting 27(5-6), 733–759. DOI: 10.1111/1468-5957.00332.

PRS Group, 2015. *International Country Risk Guide (ICRG)*. The PSR Group, New York.

Quintana, D.S., 2015. *From pre-registration to publication: a non-technical primer for conducting a meta-analysis to synthesize correlational data*. Frontiers in Psychology 6(1549), 1-9. DOI: 10.3389/fpsyg.2015.01549.

Raftery, A.E., Madigan, D., Hoeting, J.A., 1997. *Bayesian Model Averaging for Linear Regression Models*. Journal of the American Statistical Association 92(437), 179-191. DOI: 10.1080/01621459.1997.10473615.

Rahim, N.A., Goodacre, A., Veld, C., 2014. *Wealth effects of convertible-bond and warrant-bond offerings: a meta-analysis*. The European Journal of Finance 20(4), 380-398. DOI: 10.1080/1351847X.2012.712920.

Ramlall, I., 2010. *Reasons Motivating Firms to Hedge: A Review of the Empirical Literature*. The IUP Journal of Financial Economics 8(1&2), 67-81.

Ratcliffe, C., Dimovsdki, W., 2012. *REIT mergers and acquisitions: A meta-analysis*. Journal of Property Investment & Finance 30(3), 241-256. DOI: 10.1108/14635781211223815.

Reed, R.W., Florax, R.J.G.M., Poot, J., 2015. *A Monte Carlo Analysis of Alternative Meta-Analysis Estimators in the Presence of Publication Bias*. Economics: The Open-Access, Open-Assessment E-Journal 9(2015-30), 1-40. DOI: 10.5018/economics-ejournal.ja.2015-30.

Reed, R.W., Sidek, N.N., 2015. *A Replication of "Meta-Analysis of the Effect of Fiscal Policies on Long-run Growth" (European Journal of Political Economy 2004)*. Public Finance Review 44(3), 397-405. DOI: 10.1177/1091142114568659.

Reilly, D., Lucey, B., Gurdgiev, 2010. *Real estate and the stock market: a meta-regression analysis*. Working Paper at Trinity College Dublin. DOI: 10.2139/ssrn.1679303.

Revelli, C., Viviani, J.-L., 2015. *Financial Performance of Socially Responsible Investing (SRI): What Have We Learned? A Meta-Analysis*. Business Ethics: A European Review 24(2). DOI: 10.1111/beer.12076.

Riley, R.D., Higgins, J.P.T., Deeks, J.J., 2011. *Interpretation of random effects meta-analyses*. British Medical Journal 342, 964-967. DOI: 10.1136/bmj.d549.

Roberts, C.J., Stanley, T.D., 2005. *Meta-Regression Analysis: Issues of Publication Bias in Economics*. Blackwell, Malden.

Rosenbusch, N., Brinckmann, J., Müller, V., 2013. *Does acquiring venture capital pay off for the funded firms? A meta-analysis on the relationship between venture capital investment and funded firm financial performance*. Journal of Business Venturing 28(3), 335-353. DOI: 10.1016/j.jbusvent.2012.04.002.

Rosietta, H., Oktavia, D., 2011. *Motivation for the Use of Derivatives in Public Companies in Indonesia*. Paper presented at British Accounting and Finance Association Annual Conference 2011, Birmingham.

Rossi, J.L., Laham, J., 2008. *The Impact of Hedging on Firm Value: Evidence from Brazil*. Working Paper at Ibmec São Paulo.

Rothstein, H.R., Sutton, A.J., Borenstein, M., 2005. *Publication bias in meta-analysis. Prevention, assessment and adjustments*. John Wiley & Sons, Chichester.

Rusnak, M., Havranek, T., Horvath, R., 2013. *How to Solve the Price Puzzle? A Meta-Analysis*. Journal of Money, Credit and Banking 45(1), 37-70. DOI: 10.1111/j.1538-4616.2012.00561.x.

Sahoo, A., 2015. *Corporate Risk Management through Financial Derivatives: A Review of the Literature*. International Journal of Advanced Research in Managment and Social Sciences 4(8), 200-233.

Sanchez-Meca, J., Marin-Martinez, F., 1998. *Weighting by Inverese Variance or by Sample Size in Meta-Analysis: A Simulation Study*. Educational and Psychological Measurement 58(2), 211-220. DOI: 10.1177/0013164498058002005.

Schock, C., Geyer-Klingeberg, J., Rathgeber, A., 2021. *Corporate Social Responsibility as a Strategic Instrument of Risk Mitigation? A Meta-Regression Analysis on the Impact of CSR on Corporate Risk*. Paper presented at the Meta-Analysis of Economics Research Network Colloquium 2021, Athens.

Schommer, M., Richter, A., Karna, A., 2019. *Does the Diversification–Firm Performance Relationship Change Over Time? A Meta-Analytical Review*. Journal of Management Studies 56(1), 270-298. DOI: 10.1111/joms.12393.

Seidens, S., 2019. *A Meta-Study on the Relation between Idiosyncratic Volatility and Expected Returns*. Working Paper WHU - Otto Beisheim School of Management.

Shadish, W.R., Sweeney, R.B., 1991. *Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes*. Journal of Consulting and Clinical Psychology 59(6), 883-893. DOI: 10.1037//0022-006x.59.6.883.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., 2005. *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation*. British Medical Journal 349(g7647). DOI: 10.1136/bmj.g7647.

Shao, R., Rupp, D.E., Skarlicki, D.P., Jones, K.S., 2013. *Employee Justice Across Cultures: A Meta-Analytic Review*. Journal of Management 39(1). DOI: 10.1177/0149206311422447.

Siddiqui, S.S., 2015. *The association between corporate governance and firm performance – a meta-analysis*. International Journal of Accounting and Information Management 23(3), 218-237. DOI: 10.1108/IJAIM-04-2014-0023.

Silver, N.C., Dunlap, W.P., 1987. *Averaging correlation coefficients: Should Fisher's z transformation be used?* Journal of Applied Psychology 72(1), 146-148. DOI: 10.1037/0021-9010.72.1.146.

Singh, H.P., Kumar, S., Colombage, S., 2017. *Working capital management and firm profitability: a meta-analysis*. Quantitative Research in Financial Markets 9(1), 34-47. DOI: 10.1108/QRFM-06-2016-0018.

Slavin, R.E., 1986. *Best-Evidence Synthesis: An Alternative to Meta-Analytic and Traditional Reviews*. Educational Researcher 15(9), 5-11. DOI: 10.3102/0013189X015009005.

Slavin, R.E., 1995. *Best evidence synthesis: An intelligent alternative to meta-analysis*. Journal of Clinical Epidemiology 48(1), 9-18. DOI: 10.1016/0895-4356(94)00097-A.

Smith, C.W., Stulz, R.M., 1985. *The Determinants of Firms' Hedging Policies*. Journal of Financial and Quantitative Analysis 20(4), 391-405. DOI: 10.2307/2330757.

Smith, G., 2018. *Step away from stepwise*. Journal of Big Data 5(32), 1-12. DOI: 10.1186/s40537-018-0143-6.

Sollis, R., 2012. *Empirical Finance for Finance and Banking*, 1st ed. Wiley, Chichester.

Stanley, T.D., 2001. *Wheat from Chaff: Meta-Analysis as Quantitative Literature Review*. Journal of Economic Perspectives 15(3), 131-150. DOI: 10.1257/jep.15.3.131.

Stanley, T.D., 2005. *Beyond Publication Bias*. Journal of Economic Surveys 19(3), 309-345. DOI: 10.1111/j.0950-0804.2005.00250.x.

Stanley, T.D., 2008. *Meta-regression methods for detecting and estimating empirical effect in the presence of publication bias*. Oxford Bulletin of Economics and Statistics 70(1), 103-127. DOI: 10.1111/j.1468-0084.2007.00487.x.

Stanley, T.D., 2015. *Reader's comments on "A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias"*. Economics: The Open-Access, Open-Assessment E-Journal. Available from: http://www.economics-ejournal.org/dataset/comments/comment.2015-04-23.8287099025.pdf, Accessed: 2022/01/02.

Stanley, T.D., 2017. *Limitations of PET-PEESE and Other Meta-Analysis Methods*. Social Psychological and Personality Science 8(5), 581-591. DOI: 10.1177/1948550617693062.

Stanley, T.D., Doucouliagos, C., 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), 390-394. DOI: 10.1111/joes.12363.

Stanley, T.D., Doucouliagos, C., Jarrell, S.B., 2008. *Meta-Regression Analysis as the Socio-Economics of Economics Research*. The Journal of Socio-Economics 37(1), 276-292. DOI: 10.1016/j.socec.2006.12.030.

Stanley, T.D., Doucouliagos, H., 2010. *Picture This: A Simple Graph that Reveals much Ado About Research*. Journal of Economic Surveys 24(1), 170-191. DOI: 10.1111/j.1467-6419.2009.00593.x.

Stanley, T.D., Doucouliagos, H., 2012. *Meta-Regression Analysis in Economics and Business*, 1st ed. Routledge, Oxon. DOI: 10.4324/9780203111710

Stanley, T.D., Doucouliagos, H., 2014. *Meta-Regression Approximations to Reduce Publication Selection Bias*. Research Synthesis Methods 5(1), 60-78. DOI: 10.1002/jrsm.1095.

Stanley, T.D., Doucouliagos, H., 2015. *Neither Fixed nor Random: Weighted Least Squares Meta-Analysis*. Statistics in Medicine 34(13), 2116-2127. DOI: 10.1002/sim.6481.

Stanley, T.D., Doucouliagos, H., 2017. *Neither Fixed nor Random: Weighted Least Squares Meta-Regression*. Research Synthesis Methods 8(1), 19-42. DOI: 10.1002/jrsm.1211.

Stanley, T.D., Jarrell, S.B., 1989. *Meta-Regression Analysis: A Quantitative Method of Literature Surveys*. Journal of Economic Surveys 3(2), 161-170. DOI: 10.1111/j.1467-6419.1989.tb00064.x.

Stanley, T.D., Jarrell, S.B., Doucouliagos, H., 2010. *Could It Be Better to Discard 90% of the Data? A Statistical Paradox*. The American Statistician 64(1), 70-77. DOI: 10.1198/tast.2009.08205.

Steiger, J.H., Lind, J.C., 1980. *Statistically based tests for the number of common factors*. Paper presented at that Annual meeting of the Psychometric Society, Iowa City.

Sterck, O., 2019. *Beyond the stars*. Journal of Economic Surveys 33(5), 1409-1436. DOI: 10.1111/joes.12330.

Stulz, R.M., 1996. *Rethinking Risk Management*. Journal of Applied Corporate Finance 9(3), 8-25. DOI: 10.1111/j.1745-6622.1996.tb00295.x.

Sundaramurthy, C., Rhoades, D.L., Rechner, P.L., 2005. *A Meta-analysis of the Effects of Executive and Institutional Ownership on Firm Performance*. Journal of Managerial Issues 17(4), 494-510.

Tanda, A., Manzi, G., 2020. *Underpricing of venture backed IPOs: a meta-analysis approach*. Economics of Innovation and New Technology 29(4), 331-348. DOI: 10.1080/10438599.2019.1625154.

Taub, A.J., 1975. *Determinants of the Firm's Capital Structure*. The Review of Economics and Statistics 57(4), 410-416. DOI: 1935900.

Thornton, A., Lee, P., 2000. *Publication Bias in Meta-Analysis: Its Causes and Consequences*. Journal of Clinical Epidemiology 53(2), 207-216. DOI: 10.1016/s0895-4356(99)00161-4.

Treanor, S.D., Carter, D.D., Rogers, D.A., Simkins, B.J., 2013. *Operational and Financial Hedging: Friend or Foe? Evidence from the U.S. Airline Industry*. Journal of Accounting and Finance 13(6), 64-91.

Treanor, S.D., Rogers, D.A., Carter, D.A., Simkins, B.J., 2014. *Exposure, hedging, and value: New evidence from the U.S. airline industry*. International Review of Financial Analysis 34, 200-211. DOI: 10.1016/j.irfa.2014.04.002.

Trikalinos, T.A., Hoaglin, D.C., Schmid, C.H., 2013. *Empirical and Simulation-Based Comparison of Univariate and Multivariate Meta-Analysis for Binary Outcomes*. Methods Research Reports by Agency for Healthcare Research and Quality (US). Available from: https://www.ncbi.nlm.nih.gov/books/NBK132562/.

Triki, T., 2005. *Research on Corporate Hedging Theories: A Critical Review of the Evidence to Date*, HEC Montreal Working Paper No. 05-04.

Tufano, P., 1996. *Who Manages Risk? An Empirical Examination of Risk Management Practices in the Gold Mining Industry*. The Journal of Finance 51(4), 1097-1137. DOI: 10.1111/j1540-6261.1996.tb04064.x.

Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison-Wesley, Reading.

Valentine, T.J., 1979. *Hypothesis tests and confidence intervals for mean elasticities calculated from linear regression equations*. Economics Letters 4(4), 363-367. DOI: 10.1016/0165-1765(79)90186-1.

Valickova, P., Havranek, T., Horvath, R., 2015. *Financial Development and Economic Growth: A Meta-Analysis*. Journal of Economic Surveys 29(3), 506-526. DOI: 10.1111/joes.12068.

van Aert, R.C.M., van Assen, M.A.L.M., 2020. *Correcting for Publication Bias in a Meta-Analysis with the P-Uniform* Method*. Working Paper at Tilburg University & Utrecht University.

van Essen, M., Otten, J., Carberry, E.J., 2015a. *Assessing Managerial Power Theory: A Meta-Analytic Approach to Understanding the Determinants of CEO Compensation*. Journal of Management 41(1), 164-202. DOI: 10.1177/0149206311429378.

van Essen, M., Otten, J., Carbery, E.J., 2015b. *Assessing Managerial Power Theory: A Meta-Analytic Approach to Understanding the Determinants of CEO Compensation*. Journal of Management 41(1), 164-202. DOI: 10.1177/0149206311429378.

van Essen, M., van Oosterhout, J.H., Carney, M., 2012. *Corporate boards and the performance of Asian firms: A meta-analysis*. Asia Pacific Journal of Management 29, 873-905. DOI: 10.1007/s10490-011-9269-1.

van Ewijk, C., de Groot, H.L.F., Santing, A.J.C., 2012. *A Meta-Analysis of the Equity Premium*. Journal of Empirical Finance 19(5), 819-830. DOI: 10.1016/j.jempfin.2012.07.002.

VanderWerf, P.A., Mahon, J.F., 1992. *Meta-Analysis of the Impact of Research Methods on Findings of First-Mover Advantage*. Management Science 43(11), 1469-1608. DOI: 10.1287/mnsc.43.11.1510.

Veld, C., Verwijmeren, P., Zabolotnyuk, Y., 2020. *Wealth Effects of Seasoned Equity Offerings: A Meta-Analysis*. International Review of Finance 20(1), 77-131. DOI: 10.1111/irfi.12218.

Vernimmen, P., Quiry, P., Dallocchio, M., Fur, Y.L., Salvi, A., 2009. *Corporate Finance - Theory and Practice*, 2nd ed. John Wiley & Sons, Chichester.

Viechtenbauer, W., 2005. *Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model*. Journal of Educational and Behavioral Statistics 30(3), 261-293. DOI: 10.3102/10769986030003261.

Viswesvaran, C., Ones, D.S., 1995. *Theory Testing: Combining Psychometric Meta-Analysis and Structural Equations Modeling*. Personnel Psychology 48(4), 865-885. DOI: 10.1111/j.1744-6570.1995.tb01784.x.

Wang, K., Shailer, G., 2015. *Ownership concentration and firm performance in emerging markets: a meta-analysis*. Journal of Economic Surveys 29(2), 199-229. DOI: 10.1111/joes.12048.

Wang, K.T., Shailer, G., 2018. *Does Ownership Identity Matter? A Meta-analysis of Research on Firm Financial Performance in Relation to Government versus Private Ownership*. Abacus 54(1), 1-35. DOI: 10.1111/abac.12103.

Wang, P.-f., Li, S., Zhou, J., 2010. *Financial risk management and enterprise value creation: Evidence from non-ferrous metal listed companies in China*. Nankai Business Review International 1(1), 5-19. DOI: 10.1108/20408741011032836.

Wasserstein, R.L., Lazar, N.A., 2016. *The ASA's Statement on p-Values: Context, Process, and Purpose*. The American Statistician 70(2), 129-133. DOI: 10.1080/00031305.2016.1154108.

Weidemann, J.F., 2016. *The Determinants of Cash Holdings: Evidence from Meta-Regression Analysis*, European Financial Management Annual Meeting 2016, Basel.

Weiying, J., Jian, Z., 2010. *Risk Management and its Impacts on Accounting Performance and Firm Value in China*. Paper presented at International Conference on Management and Service Science 2010, Wuhan. DOI: 10.1109/ICMSS.2010.5576910.

White, H., 1980. *A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity*. Econometrica 48(4), 817-830. DOI: 10.2307/1912934.

Wimmer, T., Geyer-Klingeberg, J., Hütter, M., Schmid, F., Rathgeber, A., 2021. *The Impact of Speculation on Commodity Prices: A Meta-Granger Analysis*. Journal of Commodity Markets 22(100148). DOI: 10.1016/j.jcomm.2020.100148.

Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge.

Wooldridge, J.M., 2012. *Introductory Econometrics. A Modern Approach*, 5th ed. South Western Cengage Learning, Mason.

World Bank, 2017a. *Doing Business Project*, www.doingbusiness.org, Accessed: 2022/01/02.

World Bank, 2017b. *World Bank national accounts data, and OECD National Accounts data files*. World Bank, Washington.

World Bank, 2017c. *World Federation of Exchanges Database*. World Bank, Washington.

World Bank, 2017d. *Worldwide Governance Indicators*. World Bank, Washington.

Wruck, K.H., 1990. *Financial distress, reorganization, and organizational efficiency*. Journal of Financial Economics 27(2), 419-444. DOI: 10.1016/0304-405X(90)90063-6.

Xiang, C., Bi, C., 2015. *Empirical Research on Influence of China Listed Companies Using the Derivative Financial Instruments on Its Own Value*, in: Xu, J., Nickel, S., Machado, V.C., Hajiyev, A. (Eds.), Advances in Intelligent Systems and Computing, 711-722. DOI: 10.1007/978-3-662-47241-5.

Xue, X., Cheng, M., Zhang, W., 2020. *Does Education Really Improve Health? A Meta-Analysis*. Journal of Economic Surveys 35(1), 71-105. DOI: 10.1111/joes.12399.

Yitzhaki, S., Schechtman, E., 2013. *The Gini Methodology*. Springer, New York. DOI: 10.1007/978-1-4614-4720-7.

Zhou, H., Wang, L., Wu, J., 2012. *Impacts of Derivatives on Firms' Value Empirical Results from Chinese Companies*. Paper presented at International Conference on Emerging Computation and Information Technologies for Education 2012, Hangzhou. DOI: 10.1007/978-3-642-28466-3.

Zigraiova, D., Havranek, T., 2016. *Bank Competition and Financial Stability: Much Ado about Nothing?* Journal of Economic Surveys 30(5), 944-981. DOI: 10.1111/joes.12131.

Zigraiova, D., Havranek, T., Irsova, Z., Novak, J., 2021. *How Puzzling Is the Forward Premium Puzzle? A Meta-Analysis*. European Economic Review 134(103714). DOI: 10.1016/j.euroecorev.2021.103714.

Appendix A. Appendices to Chapter 2

A.1. Overview of Existing Meta-Analyses in Finance

Topic Field	Authors	Topic Area	Publication Outlet	AJG
Asset Pricing	Astakhov et al. (2019)	Cross-Section of Returns	Journal of Economic Surveys	2
	Bajzik (2021)	Information, Market Microstructure, or Trading Frequency	International Review of Financial Analysis	3
	Bialkowski and Perera (2019)	Derivatives, Market Mispricing	International Review of Financial Analysis	3
	Coggin and Hunter (1987)	Market Risk Factors	Journal of Portfolio Management	2
	Eshgi (2021)	Analysts, News, Media, and Market Sentiment	International Journal of Research in Marketing	4
	Frooman (1997)	Analysts, News, Media, and Market Sentiment	Business and Society	3
	García-Meca and Sánchez-Ballesta (2006)	Analysts, News, Media, and Market Sentiment	International Business Review	3
	Geyer-Klingeberg et al. (2018b)	Analysts, News, Media, and Market Sentiment	Applied Economics	2
	Grežo (2020)	Psychology, Investor Behavior, and Household Finance	Review of Behavioral Finance	1
	Engelen et al. (2020)	Asymmetric Information and Pricing	Long Range Planning	3
	Hubler et al. (2019)	Analysts, News, Media, and Market Sentiment	Journal of Economic Surveys	2
	Kim et al. (2014b)	Frictions and Market Efficiency	Book Chapter	–
	Papenfuß et al. (2021)	Information, Market Microstructure, or Trading Frequency	Working Paper	–
	Reilly et al. (2010)	Return Dynamics	Working Paper	–
	Revelli and Viviani (2015)	Mutual Funds	Business Ethics: A European Review	2
	Seidens (2019)	Market Risk Factors; Volatility, Tail Risk	Working Paper	–
	van Ewijk et al. (2012)	Market Risk Factors	Journal of Empirical Finance	3
	Wimmer et al. (2021)	Market Mispricing	Journal of Commodity Markets	–
	Zigraiova et al. (2021)	Market Mispricing; Market Risk Factors	European Economic Review	3
Corporate Finance	Arnold et al. (2014)	Risk Management	Quarterly Review of Economics and Finance	2
	Bachiller (2017)	Corporate Governance	Management Decision	2
	Bachiller et al. (2021)	Risk Management	Finance Research Letters	2
	Bessler et al. (2019)	Risk Management	International Review of Financial Analysis	3
	Burkhard et al. (2018)	Behavioral; CEOs/CFOs	Working Paper	–
	Cambrea et al. (2017)	Liquidity and Cash Management*	Working Paper	–
	Noel et al. (1990)	Financial Performance*	Management Science	4*
	Daily et al. (2003)	Raising Capital (incl. IPOs/SEOs)	Entrepreneurship, Theory and Practice	4
	Dalton et al. (2003)	Corporate Governance	Academy of Management Journal	4*
	Dao and Ta (2020)	Capital Structure	Journal of Economics and Development	–

Topic Field	Authors	Topic Area	Publication Outlet	AJG
Corporate Finance	Datta et al. (1992)	Mergers and Acquisitions	Strategic Management Journal	4*
	de Mooij (2011)	Capital Structure	Working Paper	–
	Feld et al. (2013)	Capital Structure	Journal of Banking and Finance	3
	Fernau and Hirsch (2019)	Dividend and Payout Policy	International Review of Financial Analysis	3
	Geyer-Klingeberg et al. (2018a)	Risk Management	Business Research	–
	Geyer-Klingeberg et al. (2019)	Risk Management	International Review of Financial Analysis	3
	Geyer-Klingeberg et al. (2021)	Risk Management	The European Journal of Finance	3
	Hang et al. (2018)	Capital Structure	The Quarterly Review of Economics and Finance	2
	Hang et al. (2021a)	Capital Structure; Risk Management	International Journal of Finance and Economics	3
	Hansen and Block (2021)	Capital Structure	Corporate Governance: An International Review	3
	Heugens et al. (2009)	Corporate Governance	Asia Pacific Journal of Management	3
	Holderness (2018)	Corporate Governance; Raising Capital (incl. IPOs/SEOs)	Journal of Financial Economics	4*
	Iwasaki and Mizobata (2018)	Corporate Governance	Annals of Public and Cooperative Economics	2
	Iwasaki and Mizobata (2020)	Corporate Governance	Journal of Economics and Business	1
	Iwasaki et al. (2018)	Corporate Governance	Post-Communist Economies	1
	Iwasaki and Mizobata (2020)	Corporate Governance	Emerging Markets Finance and Trade	2
	Kim (2019)	Corporate Culture or Social Responsibility	Asia-Pacific Journal of Financial Studies	–
	King et al. (2004)	Mergers and Acquisitions	Strategic Management Journal	4*
	Klein (2017)	Raising Capital (incl. IPOs/SEOs)	Working Paper	–
	La Rocca (2021)	Raising Capital (incl. IPOs/SEOs)	Review of Managerial Science	2
	La Rocca et al. (2017)	Corporate Governance	Economics Bulletin	–
	Lee and Madhavan (2010)	Mergers and Acquisitions	Journal of Management	4*
	Lindner et al. (2018)	Capital Structure	International Business Review	3
	Nehrebecka and Dzik-Walczak (2018)	Capital Structure	Zb. rad. Ekon. fak. Rij.	–
	Post and Byron (2015)	Corporate Governance	Academy of Management Journal	4*
	Rahim et al. (2014)	Raising Capital (incl. IPOs/SEOs)	European Journal of Finance	3
	Ratcliffe and Dimovsdki (2012)	Mergers and Acquisitions	Journal of Property Investment and Finance	1
	Rosenbusch et al. (2013)	Raising Capital (incl. IPOs/SEOs)	Journal of Business Venturing	4
	Schock et al. (2021)	Corporate Culture or Social Responsibility	Working Paper	–
	Schommer et al. (2019)	Corporate Strategy*	Journal of Management Studies	4
	Siddiqui (2015)	Corporate Governance	Int. Journal of Accounting and Information Management	2
	Singh et al. (2017)	Liquidity and Cash Management*	Quantitative Research in Financial Markets	–
	Sundaramurthy et al. (2005)	Corporate Governance	Journal of Managerial Issues	–
	Tanda and Manzi (2020)	Raising Capital (incl. IPOs/SEOs)	Economics of Innovation and New Technology	2

Topic Field	Authors	Topic Area	Publication Outlet	AJG
Corporate Finance	van Essen et al. (2012)	Corporate Governance	Asia Pacific Journal of Management	3
	van Essen et al. (2015a)	CEOs/CFOs; Compensation and Agency	Journal of Management	4*
	Veld et al. (2020)	Raising Capital (incl. IPOs/SEOs)	International Review of Finance	3
	Wang and Shailer (2015)	Corporate Governance	Journal of Economic Surveys	2
	Wang and Shailer (2018)	Corporate Governance	Abacus	3
	Weidemann (2016)	Liquidity and Cash Management*	Working Paper	–
Financial Intermediation	Aiello and Bonanno (2016)	Financial Intermediation: Banking	International Review of Applied Economics	1
	Aiello and Bonanno (2018)	Financial Intermediation: Banking	Journal of Economic Surveys	2
	Bhatia and Gulati (2021)	Banking	Research in International Business and Finance	2
	Ho et al. (2021)	Banking	International Journal of Financial Studies	–
	Irsova and Havranek (2012)	Financial Intermediation: Banking	Prague Economic Papers	–
	Kysucky and Norden (2016)	Bank Lending Behavior	Management Science	4*
	Zigraiova and Havranek (2016)	Banking	Journal of Economic Surveys	2

Notes: This table presents an overview of existing meta-analytical studies in finance research, as of 31/01/2022. Topic fields and topic areas are the same as in the list area assignments of the Annual Meeting of the American Finance Association (AFA). Multiple topic areas are separated by a semicolon. AJG is the ranking in the 2018 Academic Journal Guide published by the Chartered Association of Business Schools (https://charteredabs.org/academic-journal-guide-2018-view/): 4* is the top category, followed by 4, 3, 2, and 1 (lowest ranking). AJG rankings are not available for all journals, also not for working papers and book chapter. The hypen in the last column indicates that no ranking is available. * Marks additional topic areas not included in the AFA list area assignments.

Appendix B. Appendices to Chapter 3

B.1. List of Primary Studies on Hedging and Firm Value

ID	Author(s)	Start Year	End Year	No. of Estimates	Mean	Median	Min.	Max.	Std. Dev.
1	Adam and Nain (2013)	1999	1999	9	0.188	0.209	-0.054	0.412	0.168
2	Afza and Alam (2016)	2004	2010	13	0.120	0.058	0.007	0.525	0.165
3	Ahmed et al. (2013)	2005	2012	66	0.003	0.009	-0.181	0.145	0.068
4	Alam and Gupta (2018)	2008	2015	12	0.159	0.124	-0.067	0.525	0.141
5	Allayannis and Weston (2001)	1990	1995	35	0.041	0.042	-0.063	0.108	0.039
6	Ayturk et al. (2016)	2007	2013	18	0.003	0.002	-0.122	0.294	0.085
7	Bae et al. (2016)	2002	2010	2	-0.004	-0.004	-0.005	-0.003	0.001
8	Bae et al. (2018)	2005	2010	2	0.283	0.283	0.052	0.514	0.327
9	Bai et al. (2016)	2008	2015	3	0.361	0.412	0.183	0.487	0.159
10	Bashir et al. (2013)	2006	2010	9	-0.007	-0.014	-0.263	0.220	0.150
11	Belghitar et al. (2008)	1995	1995	10	0.129	0.141	-0.024	0.204	0.062
12	Belghitar et al. (2013)	2002	2005	10	-0.024	-0.021	-0.095	0.022	0.037
13	Berrospide et al. (2008)	1997	2005	4	0.132	0.129	0.123	0.147	0.011
14	Brunzell et al. (2011)	2007	2007	3	0.462	0.463	0.398	0.525	0.064
15	Carter et al. (2006)	1992	2003	36	0.060	0.062	-0.066	0.230	0.062
16	Chang et al. (2016)	2001	2010	3	-0.130	-0.121	-0.224	-0.044	0.090
17	Chen and King (2014)	1994	2009	2	0.170	0.170	0.029	0.311	0.200
18	Chen and Shao (2010)	2007	2009	8	0.122	0.116	0.056	0.202	0.062
19	Chen et al. (2011)	1998	2001	6	-0.063	-0.041	-0.263	0.025	0.105
20	Choi et al. (2013)	2001	2006	16	0.224	0.159	-0.030	0.525	0.192
21	Chou and Lai (2013)	2005	2010	3	-0.078	-0.083	-0.260	0.109	0.184
22	Clark and Judge (2009)	1995	1995	34	0.186	0.146	0.116	0.411	0.076
23	Clark and Mefteh (2010)	2004	2004	7	0.133	0.089	0.060	0.385	0.113
24	Clark et al. (2007)	2004	2004	6	0.033	0.053	-0.054	0.064	0.046
25	Dionne et al. (2013)	1993	1999	8	0.130	0.133	0.096	0.161	0.024
26	Disatnik et al. (2014)	2002	2007	3	0.040	0.045	-0.008	0.084	0.046
27	dos Santos et al. (2017)	2006	2014	45	0.046	0.029	-0.053	0.299	0.067
28	Elsawaf (2005)	1993	2000	82	0.144	0.101	-0.096	0.525	0.140
29	Fauver and Naranjo (2010)	1991	2000	16	-0.130	-0.143	-0.263	0.001	0.085
30	Gleason et al. (2005)	1998	1998	5	0.034	0.034	0.033	0.034	0.000
31	Hagelin et al. (2007)	1997	2001	6	0.152	0.112	-0.029	0.525	0.207
32	Jankensgård (2015a)	2009	2009	6	0.148	0.154	0.069	0.206	0.045
33	Jankensgård (2015b)	2000	2008	1	-0.166	-	-	-	-
34	Jankensgård et al. (2014)	2009	2009	4	-0.016	-0.018	-0.021	-0.006	0.007
35	Jin and Jorion (2006)	1998	2001	18	-0.018	-0.021	-0.098	0.045	0.038
36	Jin and Jorion (2007)	1991	2000	4	-0.116	-0.108	-0.189	-0.057	0.060

(Continued on next page)

37	Jorge and Augusto (2012)	2007	2007	2	-0.081	-0.081	-0.084	-0.078	0.004
38	Kapitsinas (2008)	2004	2006	16	0.117	0.089	-0.090	0.438	0.131
39	Khediri (2010)	2000	2002	6	-0.021	-0.019	-0.132	0.087	0.072
40	Khediri and Folus (2010)	2001	2001	5	-0.057	-0.047	-0.085	-0.031	0.024
41	Kim et al. (2006)	1998	1998	3	0.008	0.008	0.000	0.017	0.009
42	Kim et al. (2014a)	1992	2004	22	0.161	0.155	-0.127	0.525	0.157
43	Kim et al. (2017)	2003	2013	15	0.035	0.067	-0.128	0.205	0.103
44	Korkeamäki et al. (2016)	2000	2015	4	0.029	0.028	0.021	0.041	0.008
45	Li et al. (2014)	2007	2007	5	-0.020	0.003	-0.263	0.135	0.147
46	Lievenbrück and Schmid (2014)	1995	2005	6	-0.102	-0.139	-0.263	0.115	0.166
47	Lookman (2004)	1992	2000	28	0.006	-0.004	-0.058	0.105	0.048
48	Luo (2016)	2007	2013	6	0.230	0.289	0.073	0.332	0.118
49	MacKay and Moeller (2007)	1985	2004	26	-0.006	-0.032	-0.113	0.142	0.088
50	Magee (2013)	1996	2000	11	0.031	0.021	-0.096	0.120	0.057
51	Manchiraju et al. (2014)	2007	2012	1	0.027	0.027	0.027	0.027	-
52	Marami and Dubois (2013)	1998	2005	12	0.124	0.065	-0.001	0.335	0.128
53	Meredith (2002)	1996	1998	7	-0.050	-0.053	-0.129	0.051	0.053
54	Mohammad (2014)	2006	2010	1	0.080	-	-	-	-
55	Nain (2005)	1999	1999	4	0.142	0.108	0.008	0.344	0.143
56	Nguyen and Faff (2007)	1999	2000	22	-0.149	-0.177	-0.263	0.005	0.094
57	Nguyen and Faff (2010)	1999	2000	12	-0.051	-0.029	-0.197	0.106	0.084
58	Nova et al. (2015)	2005	2013	12	-0.006	-0.006	-0.226	0.080	0.077
59	Panaretou (2014)	2003	2010	30	0.070	0.063	-0.084	0.195	0.061
60	Pérez-González and Yun (2013)	1997	2007	58	0.154	0.095	-0.166	0.525	0.145
61	Phan et al. (2014)	1998	2009	64	-0.020	-0.030	-0.205	0.224	0.099
62	Pierce (2015)	2008	2008	1	0.004	-	-	-	-
63	Pramborg (2004)	1997	2001	12	0.080	0.133	-0.086	0.163	0.095
64	Rosietta and Oktavia (2011)	2001	2009	1	0.051	-	-	-	-
65	Rossi and Laham (2008)	1996	2005	61	0.127	0.116	0.004	0.525	0.095
66	Treanor et al. (2013)	1994	2006	27	0.023	0.046	-0.084	0.096	0.058
67	Treanor et al. (2014)	1994	2008	5	0.046	0.055	0.025	0.066	0.018
68	Wang et al. (2010)	2002	2008	13	0.182	0.189	0.091	0.239	0.049
69	Weiying and Jian (2010)	2007	2007	1	-0.107	-	-	-	-
70	Xiang and Bi (2015)	2009	2013	1	0.024	-	-	-	-
71	Zhou et al. (2012)	2007	2010	1	0.095	-	-	-	-
	Overall	**1985**	**2015**	**1,016**	**0.064**	**0.053**	**-0.263**	**0.525**	**0.134**

Notes: This table reports the overview of the sample of 71 empirical primary studies examining the relationship between corporate financial hedging and firm value. The start year and the end year are the maximum sample periods examined in the primary studies. The summary statistics (mean, median, minimum, maximum, and standard deviation) refer to the estimates of the hedging premiums that are calculated from the results reported in the primary studies. Extreme observations of the hedging premiums are winsorized at the 1% and 99% quantiles.

B.2. Moderating Factors of Publication Selection (Alternative Weights)

	(1) Journal Quality	(2) Focus on FX Exposure	(3) Correction for Endogeneity	(4) No. of Control Variables Included	(5) All Variables
***Panel A. Weights:* $1/m$**					
Intercept: $\hat{\beta}_0$	0.014	0.017*	0.014	0.008	0.017
	(1.45)	(1.83)	(1.56)	(0.86)	(1.77)
SE: $\hat{\beta}_1$	0.567***	0.082	0.511***	1.138***	0.532
	(3.82)	(0.66)	(2.77)	(3.40)	(1.46)
$SE \times$ Top journal: $\hat{\delta}_1$	-0.350***				-0.201*
	(-2.81)				(-1.76)
$SE \times$ FX hedgers: $\hat{\delta}_2$		0.561***			0.663***
		(5.43)			(4.05)
$SE \times$ Control for endogeneity: $\hat{\delta}_3$			0.100		0.300
			(0.42)		(1.26)
$SE \times$ No. of control variables: $\hat{\delta}_4$				-0.053**	-0.061**
				(-2.00)	(-2.16)
***Panel B. Weights:* $1/(SE(HP)^2 \times m)$**					
Intercept: $\hat{\beta}_0$	0.008*	0.009*	0.008*	0.008*	0.008*
	(1.89)	(1.85)	(1.89)	(1.91)	(1.88)
SE: $\hat{\beta}_1$	0.663***	0.044	0.698***	0.781*	0.078
	(2.81)	(0.12)	(3.00)	(1.78)	(0.16)
$SE \times$ Top journal: $\hat{\delta}_1$	-0.173				-0.001
	(-1.19)				(-0.01)
$SE \times$ FX hedgers: $\hat{\delta}_2$		0.859***			0.865**
		(2.61)			(2.51)
$SE \times$ Control for endogeneity: $\hat{\delta}_3$			-0.222		-0.246
			(-0.98)		(-0.94)
$SE \times$ No. of control variables: $\hat{\delta}_4$				-0.014	0.002
				(-0.52)	(0.07)
No. of studies	71	71	71	71	71
No. of obs.	1,016	1,016	1,016	1,016	1,016

Notes: This table reports the results of Eq. (14) including additional interaction variables: $HP_{ij} = \beta_0 + \beta_1 SE(HP_{ij}) + \delta_k SE(HP_{ij}) M_{ij} + v_{ij}$, where HP_{ij} is the ith estimate of the hedging premium reported in the jth study. M_{ij} is a moderator variable. The meta-regression models in Panel A are estimated with inverse of the number of estimates reported per study as weights. The meta-regression models in Panel B are estimated with the inverse of the estimates' squared standard errors times the inverse of the number of effect size estimates as weights. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level. Top journal = Hedging premium is calculated from the estimates reported in a journal with a SJR score larger than 1.00. FX hedgers = Hedging premium refers to hedgers of FX risk. Control for endogeneity = Primary study explicitly controls for endogeneity. No. of control variables = Total number of variables included as controls in the primary study regression.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

B.3. Data and Method Choices as Drivers of Hedging Premiums (Alternative Weights)

Weights	(1) Alternative Weights I $\frac{1}{m}$	(2) Alternative Weights I $\frac{1}{SE(HP)^2 \times m}$
Journal Quality		
Top journal	0.005	-0.008*
	(0.16)	(-1.73)
Sample Year		
After 2001	0.026	0.017**
	(1.48)	(2.47)
Geographical Region		
Europe vs. North America	0.007	-0.020**
	(0.21)	(-2.27)
East Asia & Pacific vs. North America	-0.016	-0.019***
	(-0.50)	(-3.55)
South Asia vs. North America	0.006	-0.017
	(0.17)	(-0.96)
Latin America vs. North America	0.050***	0.018***
	(2.96)	(4.83)
Measurement of Hedging		
IR vs. FX hedgers	-0.017	-0.016***
	(-0.97)	(-5.58)
CP vs. FX hedgers	-0.052***	-0.032***
	(-5.81)	(-6.51)
Hedging dummy variable	0.007	0.013**
	(0.30)	(2.41)
Derivatives users	0.026	-0.007
	(0.35)	(-0.43)
Focus on specific instruments	-0.070***	-0.033***
	(-3.49)	(-4.14)
Estimation Characteristics		
Control for endogeneity	0.017	-0.010*
	(0.82)	(-1.94)
Interaction term	-0.051	-0.017*
	(-1.43)	(-1.79)
Control Variables		
Control for other risk exposures	-0.002	0.010**
	(-0.07)	(2.03)
Control for managerial ownership	-0.005	-0.012
	(-0.14)	(-0.83)
Publication Selection		
SE	0.541***	0.603***
	(2.60)	(3.73)
Constant	-0.008	0.028**
	(-0.10)	(2.01)
No. of studies	71	71
No. of observations	1,016	1,016

Notes: This table extends Table 14 (Model 5), which is a reduced model based on the general-to-specific approach. Definitions of explanatory variable can be found in Table 7. Model (1) uses the inverse number of estimates per study as weights. Model (2) uses the interaction between the inverse of the estimates' squared standard errors times the inverse number of estimates per study as weights. The t-statistics in parentheses are based on robust errors, clustered at study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

B.4. Country-Level Drivers of Hedging Premiums (Alternative Weights I)

	(1)	(2)	(3)	(4)
Panel A. Financial and Economic Development				
Derivatives market volume	0.012			
	(0.26)			
Stock trading volume		-0.013		
		(-0.59)		
Trade magnitude			0.033	
			(1.46)	
OECD member				-0.032
				(-1.12)
Constant	-0.008	0.037	-0.122	0.106
	(-0.06)	(0.23)	(-0.94)	(0.08)
	(5)	(6)	(7)	(8)
Panel B. Legality and Governance				
Rule-of-law	-0.010			
	(-0.63)			
Shareholder rights		0.017***		
		(2.74)		
Creditor rights			-0.009**	
			(-2.19)	
Ownership concentration				-0.005
				(-0.10)
Constant	0.001	-0.090	0.069	-0.005
	(0.01)	(-0.72)	(0.60)	(-0.04)
	(9)	(10)	(11)	(12)
Panel C. Financial Distress and Taxes				
Time to resolve insolvency	0.018			
	(1.12)			
Financial risk		-0.060		
		(-1.32)		
Composite risk			-0.061	
			(-1.29)	
Tax rate				0.049**
				(2.46)
Constant	0.065*	0.300*	0.344	-0.106
	(1.84)	(1.73)	(1.64)	(-1.11)
Other controls from Table 14 included	Yes	Yes	Yes	Yes
No. of studies	71	71	71	71
No. of observations	1,016	1,016	1,016	1,016

Notes: This table presents the results of the same meta-regression model as shown in Table 14 (Model 1), but the dummy variables for geographical regions are substituted by the country-level variables defined in Table 8. These variables are assigned to the hedging premiums based on the sample year and country reported in each study. Meta-regressions are estimated using the inverse of the number of effect sizes reported per study as weights in the meta-regression. The *t*-statistics in parentheses are based on robust errors, clustered at study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

B.5. Country-Level Drivers of Hedging Premiums (Alternative Weights II)

	(1)	(2)	(3)	(4)
Panel A. Financial and Economic Development				
Derivatives market volume	-0.044**			
	(-2.08)			
Stock trading volume		-0.005		
		(-0.80)		
Trade magnitude			-0.021***	
			(-2.72)	
OECD member				-0.012
				(-0.93)
Constant	0.046*	0.062*	0.120**	0.326
	(1.74)	(1.76)	(2.55)	(1.45)
	(5)	(6)	(7)	(8)
Panel B. Legality and Governance				
Rule-of-law	-0.002			
	(-0.43)			
Shareholder rights		-0.001		
		(-0.30)		
Creditor rights			-0.001	
			(-0.23)	
Ownership concentration				-0.008
				(-0.39)
Constant	0.048*	0.050*	0.051**	0.054*
	(1.89)	(1.91)	(2.29)	(1.88)
	(9)	(10)	(11)	(12)
Panel C. Financial Distress and Taxes				
Time to resolve insolvency	0.078**			
	(2.32)			
Financial risk		0.043		
		(0.45)		
Composite risk			-0.135	
			(-0.94)	
Tax rate				-0.054
				(-0.96)
Constant	-0.065	-0.158	0.567	0.197
	(-0.47)	(-0.45)	(0.92)	(0.84)
Other controls from Table 14 included	Yes	Yes	Yes	Yes
No. of studies	71	71	71	71
No. of observations	1,016	1,016	1,016	1,016

Notes: This table presents the results of the same meta-regression model as shown in Table 14 (Model 1), but the dummy variables for geographical regions are substituted by the country-level variables defined in Table 8. These variables are assigned to the hedging premiums based on the sample year and country reported in each study. Meta-regressions are estimated using the inverse of the estimates' squared standard errors times the inverse of the number of effect sizes reported per study as weights. The t-statistics in parentheses are based on robust errors, clustered at study level and country level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

B.6. FX Hedging Premiums after Matching on IR and CP Hedging (Alternative Weights I)

	(1) Interaction Term	(2) Subsample
FX hedgers × Control for other risk exposures	0.020	
	(0.43)	
FX hedgers	0.027	-0.011
	(1.05)	(-0.50)
Control for other risk exposures	-0.009	
	(-0.51)	
Constant	-0.068	0.152*
	(-0.55)	(1.92)
Other controls from Tab. 14 included	Yes	Yes
No. of studies	71	18
No. of observations	1,016	307

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2). Unreported control variables are identical to those in Table 14. Meta-regressions are estimated using the inverse of the number of effect sizes reported per study as weights. Model (1) includes an interaction term between the dummy variable for FX hedgers and another dummy variable indicating whether multiple risk exposures are estimated in the same primary regression (suggesting that the reported hedging premiums do not suffer from a bias due to the omission of other hedging exposures). Model (2) is based on a reduced sample of all estimates observed from models with multiple risk exposures estimated in the same primary regression. The *t*-statistics in parentheses are based on robust errors, clustered at the study level and country level.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

B.7. FX Hedging Premiums after Matching on IR and CP Hedging (Alternative Weights II)

	(1) Interaction Term	(2) Subsample
FX hedgers × Control for other risk exposures	0.006	
	(0.73)	
FX hedgers	0.015**	0.014***
	(2.51)	(4.78)
Control for other risk exposures	-0.013**	
	(-2.44)	
Constant	0.042	0.293***
	(1.59)	(5.66)
Other controls from Tab. 14 included	Yes	Yes
No. of studies	71	18
No. of observations	1,016	307

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2). Unreported control variables are identical to those in Table 14. Meta-regressions are estimated using the inverse of the estimates' squared standard errors times the inverse of the number of effect sizes reported per study as weights. Model (1) includes an interaction term between the dummy variable for FX hedgers and another dummy variable indicating whether multiple risk exposures are estimated in the same primary regression (suggesting that the reported hedging premiums do not suffer from a bias due to the omission of other hedging exposures). Model (2) is based on a reduced sample of all estimates observed from models with multiple risk exposures estimated in the same primary regression. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level.

$^{***}\, p < 0.01$, $^{**}\, p < 0.05$, $^{*}\, p < 0.10$

B.8. Accounting Rules and Measurement of Continuous Hedging (Alternative Weights I)

	(1) Accounting Changes (Issue Dates)	(2) Accounting Changes (Effective Dates)	(3) Measurement of Continuous Hedging Variable
Issue of FAS 133 in 1998	0.057***		
	(2.75)		
Issue of IAS 39 in 2003	0.048**		
	(2.35)		
Effective date of FAS 133 in 2000		0.072***	
		(4.23)	
Effective date of IAS 39 in 2005		0.053***	
		(2.73)	
Fair values vs. notional amounts			-0.040**
			(-2.32)
Actual hedge ratios vs. notional amounts			-0.056***
			(-3.01)
Other measures vs. notional amounts			0.032
			(1.39)
Constant	0.037	0.018	-0.028
	(0.86)	(0.51)	(-1.15)
Other controls from Tab. 14 included	Yes	Yes	Yes
No. of studies	40	40	40
No. of observations	326	326	326

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2) using the inverse of the number of effect sizes reported per study as weights. Unreported control variables are identical as in Table 14. Reported coefficients refer to the alternative variables included for robustness analysis. Model (1) includes two breakpoint variables referring to the issuance year of major accounting changes relevant for the reporting of hedging instruments (FAS 133 and IAS 39). Model (2) refers to the year in which the accounting changes became effective. The omitted base category is the time period before 1998 (Model 1) and before 2000 (Model 2). The breakpoint is assigned to the hedging premiums from the primary studies based on the average sample year examined in each study. Model (3) breaks down the continuous hedging variable in the different categories of measuring the extent of hedging. The omitted base group are notional amounts of hedging instruments reported in annual reports. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

B.9. Accounting Rules and Measurement of Continuous Hedging (Alternative Weights II)

	(1) Accounting Changes (Issue Dates)	(2) Accounting Changes (Effective Dates)	(3) Measurement of Continuous Hedging Variable
Issue of FAS 133 in 1998	0.029***		
	(3.90)		
Issue of IAS 39 in 2003	0.018***		
	(2.94)		
Effective date of FAS 133 in 2000		0.042***	
		(6.73)	
Effective date of IAS 39 in 2005		0.029***	
		(6.12)	
Fair values vs. notional amounts			-0.022***
			(-4.05)
Actual hedge ratios vs. notional amounts			-0.018
			(-1.44)
Other measures vs. notional amounts			0.038***
			(5.08)
Constant	0.067***	0.061***	0.026**
	(2.92)	(3.29)	(2.10)
Other controls from Tab. 14 included	Yes	Yes	Yes
No. of studies	40	40	40
No. of observations	326	326	326

Notes: This table reports the results for the same regression model as reported in Table 14 (Model 2) using the inverse of the estimates' squared standard errors times the inverse of the number of effect sizes reported per study as weights. Unreported control variables are identical as in Table 14. Reported coefficients refer to the alternative variables included for robustness analysis. Model (1) includes two breakpoint variables referring to the issuance year of major accounting changes relevant for the reporting of hedging instruments (FAS 133 and IAS 39). Model (2) refers to the year in which the accounting changes became effective. The omitted base category is the time period before 1998 (Model 1) and before 2000 (Model 2). The breakpoint is assigned to the hedging premiums from the primary studies based on the average sample year examined in each study. Model (3) breaks down the continuous hedging variable in the different categories of measuring the extent of hedging. The omitted base group are notional amounts of hedging instruments reported in annual reports. The t-statistics in parentheses are based on robust errors, clustered at the study level and country level.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Appendix C. Appendices to Chapter 4

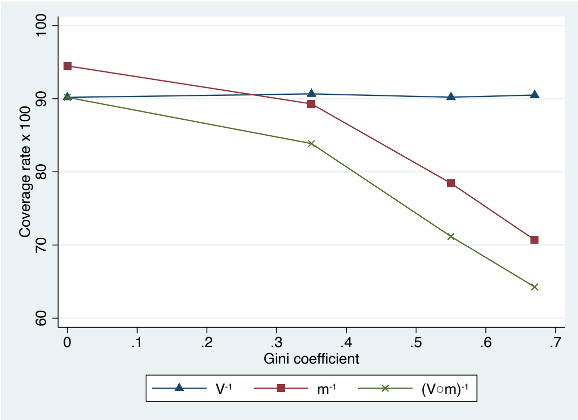
C.1. Simulation Results ($d = 0.25$, $f = 0$, Random Heterogeneity at ES Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.64	0	95.63	93.47	95.63	1.03	2.61	1.03	0.08	1.38	0.08
0	0.125	28.20	0	93.03	93.99	93.03	1.71	2.92	1.71	0.24	0.01	0.24
0	0.25	60.96	0	90.77	94.11	90.77	3.70	4.18	3.70	0.04	0.98	0.04
0	0.5	85.63	0	90.19	94.49	90.19	10.35	8.95	10.35	0.67	1.16	0.67
0	1	95.41	0	92.03	94.72	92.03	28.13	28.27	28.13	1.31	2.82	1.31
0	2	98.45	0	94.24	95.29	94.24	67.39	102.85	67.39	1.30	3.23	1.30
0	4	99.40	0	95.37	95.13	95.37	154.32	420.60	154.32	1.02	3.12	1.02
0.35	0	4.45	0	95.29	87.97	90.03	1.05	3.69	1.47	0.83	0.14	0.61
0.35	0.125	28.19	0	93.02	88.62	87.58	1.74	4.27	2.43	0.13	0.34	0.29
0.35	0.25	60.88	0	91.23	88.35	83.91	3.77	6.11	5.30	0.57	0.61	0.99
0.35	0.5	85.60	0	90.66	89.30	83.87	10.12	12.71	14.37	0.18	1.25	1.10
0.35	1	95.42	0	91.50	90.42	85.62	29.37	39.12	40.20	0.76	2.44	1.49
0.35	2	98.44	0	93.63	90.01	88.22	70.13	148.66	98.29	3.43	5.59	5.98
0.35	4	99.40	0	96.06	89.77	91.05	145.58	578.56	207.37	7.68	3.38	7.40
0.55	0	4.41	0	95.38	76.25	79.03	1.04	6.35	2.53	0.04	0.55	0.13
0.55	0.125	28.09	0	92.76	76.39	76.14	1.76	7.40	4.10	0.29	1.01	0.72
0.55	0.25	60.84	0	91.28	77.41	72.99	3.66	9.99	8.55	0.22	1.33	0.42
0.55	0.5	85.61	0	90.21	78.44	71.16	10.44	21.72	24.12	1.12	1.97	0.87
0.55	1	95.42	0	91.61	77.78	72.79	29.39	68.45	69.29	1.31	0.52	2.31
0.55	2	98.45	0	94.07	77.62	75.98	69.34	261.86	168.90	4.18	2.89	1.59
0.55	4	99.40	0	95.37	77.90	79.50	153.94	1034.26	376.43	4.63	10.43	5.66
0.67	0	4.47	0	94.78	68.36	71.74	1.07	8.50	3.45	0.32	0.37	0.37
0.67	0.125	28.31	0	93.09	67.90	67.28	1.73	9.87	5.67	0.62	1.50	0.70
0.67	0.25	60.88	0	91.38	69.33	65.93	3.63	13.38	11.30	0.44	0.76	0.41
0.67	0.5	85.58	0	90.50	70.70	64.27	10.12	29.48	32.18	2.72	1.43	1.79
0.67	1	95.41	0	92.19	70.44	65.63	27.62	91.96	91.43	1.72	3.81	4.14
0.67	2	98.45	0	93.85	70.61	68.96	69.26	345.19	232.58	0.14	0.49	0.09
0.67	4	99.40	0	95.61	71.43	72.33	150.29	1321.74	535.93	7.09	17.71	15.42
0	0	4.56	1	95.04	94.00	95.04	1.05	2.55	1.05	0.07	0.57	0.07
0	0.125	28.31	1	93.00	94.00	93.00	1.74	2.97	1.74	0.06	0.43	0.06
0	0.25	60.92	1	90.69	93.87	90.69	3.79	4.31	3.79	0.18	0.55	0.18
0	0.5	85.62	1	90.15	94.43	90.15	10.50	9.13	10.50	0.64	1.94	0.64
0	1	95.41	1	91.50	94.60	91.50	28.67	28.43	28.67	3.48	3.07	3.48
0	2	98.45	1	94.41	95.04	94.41	67.64	103.52	67.64	2.72	2.86	2.72
0	4	99.40	1	95.08	95.15	95.08	155.50	414.84	155.50	1.25	13.11	1.25
0.35	0	4.46	1	95.21	88.51	90.61	1.06	3.63	1.46	0.14	0.53	0.28
0.35	0.125	28.12	1	93.30	88.11	87.30	1.73	4.30	2.44	0.00	0.21	0.18
0.35	0.25	60.78	1	90.67	88.94	84.46	3.76	5.92	5.22	0.06	0.50	0.04
0.35	0.5	85.60	1	90.77	89.51	84.21	10.28	12.85	14.56	0.88	2.16	1.41
0.35	1	95.41	1	91.84	89.72	85.34	28.86	39.87	40.78	2.60	0.59	1.74
0.35	2	98.45	1	94.21	90.20	88.84	67.99	146.22	95.28	4.73	5.41	4.46
0.35	4	99.40	1	95.28	90.08	90.74	153.25	583.03	216.44	4.43	2.80	5.17
0.55	0	4.38	1	94.92	76.55	79.46	1.07	6.29	2.49	0.34	0.16	0.29
0.55	0.125	28.16	1	93.03	76.83	75.79	1.75	7.27	4.07	0.41	1.46	0.14
0.55	0.25	60.88	1	91.03	76.82	71.89	3.73	10.21	8.74	0.21	0.25	0.48
0.55	0.5	85.58	1	90.06	77.96	71.52	10.44	22.04	24.16	0.03	0.23	1.00
0.55	1	95.41	1	92.09	77.60	72.66	28.56	70.07	68.81	1.39	2.18	2.63
0.55	2	98.45	1	94.05	78.98	76.60	69.22	249.75	167.04	1.03	2.36	0.25
0.55	4	99.40	1	95.72	78.39	80.35	150.92	1016.24	372.71	3.90	20.17	12.04
0.67	0	4.47	1	95.27	68.36	71.47	1.05	8.42	3.54	0.09	0.76	0.70
0.67	0.125	28.07	1	93.27	70.12	68.28	1.72	9.43	5.44	0.13	0.76	0.47
0.67	0.25	60.87	1	91.11	69.13	64.56	3.61	13.81	11.76	0.30	0.70	1.09
0.67	0.5	85.55	1	90.06	69.86	62.40	10.36	29.46	32.89	0.91	2.54	3.97
0.67	1	95.41	1	91.76	70.98	65.25	28.81	92.30	94.55	2.46	2.46	0.06
0.67	2	98.45	1	94.45	71.91	69.34	66.30	334.27	225.35	1.19	4.92	1.37
0.67	4	99.40	1	95.14	70.81	71.87	151.01	1347.98	534.38	5.43	5.60	1.17
Average				**93.00**	**82.80**	**80.86**	**37.97**	**163.80**	**78.63**	**1.47**	**2.69**	**1.87**

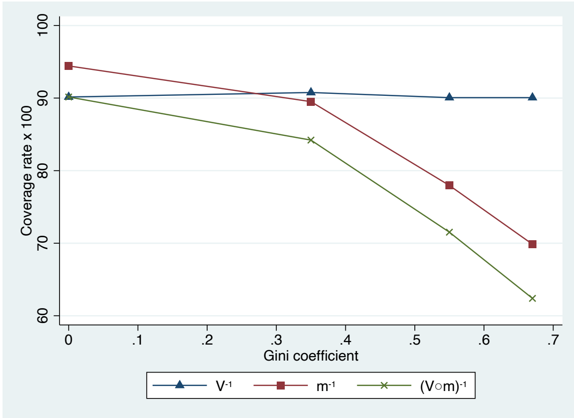
Notes: See Table 23.

C.2. Coverage, MSE, Bias ($d = 0.25$, $f = 0$, Random Heterogeneity at ES Level)

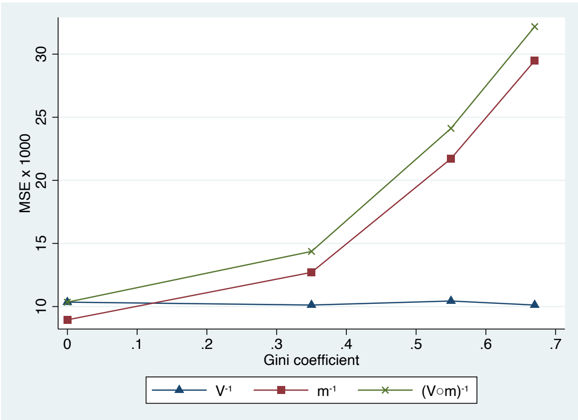
(a) Coverage; $\alpha_1 = 0$



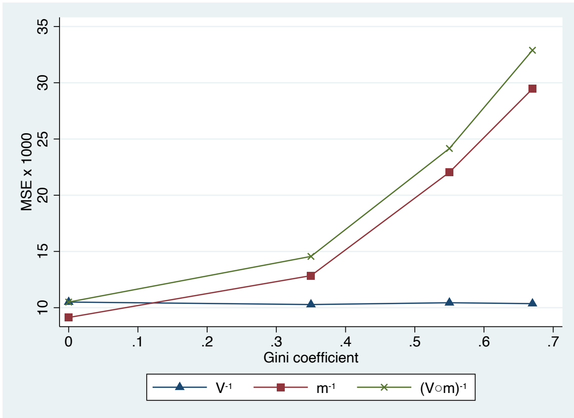
(b) Coverage; $\alpha_1 = 1$



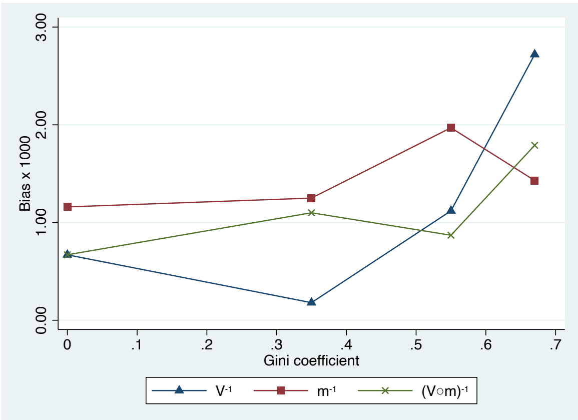
(c) MSE; $\alpha_1 = 0$



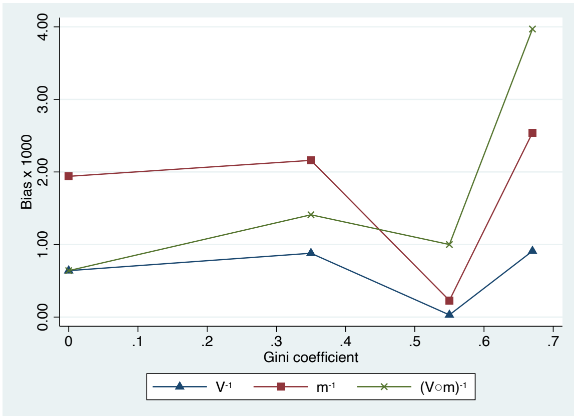
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: See Figure 35.

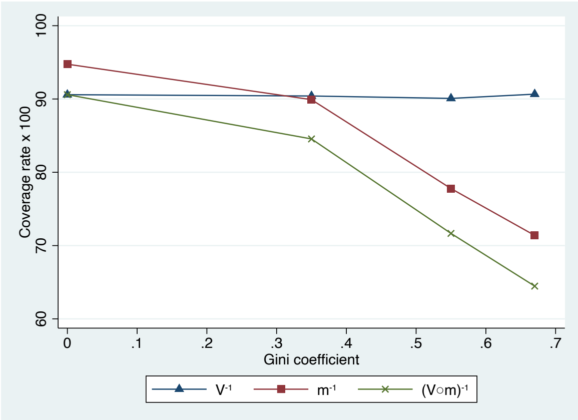
C.3. Simulation Results ($d = 0.5$, $f = 0$, Random Heterogeneity at ES Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.51	0	94.47	93.99	94.47	1.07	2.54	1.07	0.39	0.21	0.39
0	0.125	28.09	0	92.82	93.65	92.82	1.79	3.05	1.79	0.06	0.21	0.06
0	0.25	60.86	0	90.73	94.65	90.73	3.77	4.13	3.77	0.88	0.05	0.88
0	0.5	85.56	0	90.59	94.75	90.59	10.08	8.93	10.08	0.20	0.43	0.20
0	1	95.41	0	92.27	94.82	92.27	28.21	28.24	28.21	0.73	2.25	0.73
0	2	98.45	0	93.68	94.78	93.68	68.76	106.03	68.76	1.72	5.52	1.72
0	4	99.40	0	95.54	95.08	95.54	152.04	416.78	152.04	3.35	4.02	3.35
0.35	0	4.51	0	95.10	88.47	90.48	1.05	3.64	1.48	0.02	0.00	0.09
0.35	0.125	28.40	0	93.26	88.42	87.88	1.71	4.28	2.38	0.03	0.20	0.09
0.35	0.25	60.87	0	90.81	88.74	84.31	3.76	6.02	5.29	0.04	0.25	0.32
0.35	0.5	85.55	0	90.41	89.91	84.55	10.31	12.39	14.00	0.99	0.40	1.04
0.35	1	95.41	0	92.63	89.98	86.26	27.58	40.06	39.20	1.88	2.06	2.47
0.35	2	98.45	0	94.14	89.70	88.54	68.16	150.13	96.66	3.14	0.15	5.29
0.35	4	99.40	0	95.45	90.22	90.69	151.59	582.74	214.55	0.31	1.67	0.05
0.55	0	4.45	0	95.02	76.48	78.79	1.06	6.22	2.52	0.01	0.21	0.24
0.55	0.125	28.16	0	93.34	76.23	75.23	1.72	7.40	4.20	0.06	0.56	0.26
0.55	0.25	60.83	0	90.54	77.09	72.33	3.81	10.25	8.67	0.44	0.41	0.12
0.55	0.5	85.56	0	90.08	77.78	71.66	10.45	21.80	23.82	1.41	2.23	0.26
0.55	1	95.42	0	91.67	77.96	73.28	29.07	69.62	68.34	2.28	0.54	0.27
0.55	2	98.45	0	93.94	78.44	76.80	67.52	253.24	163.93	0.07	3.19	0.29
0.55	4	99.40	0	95.61	77.86	78.99	153.00	1013.48	376.81	5.73	5.16	3.08
0.67	0	4.43	0	95.15	68.07	70.92	1.03	8.65	3.51	0.24	1.04	0.59
0.67	0.125	28.19	0	93.27	69.02	68.69	1.73	9.73	5.62	0.67	2.70	2.22
0.67	0.25	60.87	0	91.04	69.85	64.32	3.69	13.50	11.96	0.38	0.10	0.39
0.67	0.5	85.60	0	90.66	71.39	64.47	10.15	28.14	31.61	0.11	0.05	0.39
0.67	1	95.41	0	91.86	69.89	64.30	28.64	94.63	94.34	2.79	2.65	3.39
0.67	2	98.44	0	93.89	70.33	67.88	68.96	348.74	231.11	2.20	8.55	0.12
0.67	4	99.39	0	95.08	70.91	71.20	151.44	1352.14	553.56	2.81	5.30	5.56
0	0	4.53	1	94.90	93.84	94.90	1.04	2.60	1.04	0.33	0.16	0.33
0	0.125	28.20	1	93.33	94.11	93.33	1.72	3.01	1.72	0.28	0.71	0.28
0	0.25	60.90	1	91.23	94.80	91.23	3.65	4.08	3.65	0.76	0.76	0.76
0	0.5	85.57	1	90.41	94.29	90.41	10.48	9.20	10.48	1.12	0.09	1.12
0	1	95.42	1	92.22	95.38	92.22	27.64	27.24	27.64	1.78	0.10	1.78
0	2	98.45	1	93.71	94.45	93.71	69.58	107.21	69.58	0.03	2.02	0.03
0	4	99.40	1	95.44	95.10	95.44	149.40	411.00	149.40	6.25	7.35	6.25
0.35	0	4.46	1	94.90	88.34	90.05	1.05	3.59	1.46	0.11	0.22	0.23
0.35	0.125	28.45	1	93.31	89.16	87.75	1.72	4.06	2.38	0.58	0.23	1.02
0.35	0.25	60.77	1	90.88	88.54	84.54	3.79	6.04	5.21	0.07	0.28	0.01
0.35	0.5	85.59	1	90.09	89.62	84.23	10.38	12.76	14.28	1.02	0.12	0.39
0.35	1	95.40	1	91.63	89.47	85.47	29.53	39.76	40.52	0.32	1.47	0.80
0.35	2	98.45	1	93.54	90.20	88.30	70.23	147.01	98.19	0.07	3.92	3.06
0.35	4	99.40	1	95.58	89.58	90.52	149.48	584.26	214.49	5.00	4.48	5.69
0.55	0	4.47	1	95.08	76.12	78.88	1.05	6.34	2.51	0.09	0.04	0.50
0.55	0.125	28.07	1	92.98	77.04	75.43	1.75	7.13	4.12	0.34	0.31	0.34
0.55	0.25	60.95	1	90.78	77.71	72.96	3.77	9.91	8.40	0.22	0.74	0.44
0.55	0.5	85.56	1	90.14	77.25	70.88	10.60	22.29	24.33	0.07	0.17	0.21
0.55	1	95.40	1	91.80	78.84	73.32	28.73	67.51	67.56	0.06	0.66	0.02
0.55	2	98.45	1	94.37	78.45	77.03	67.64	254.59	163.51	2.62	2.82	4.98
0.55	4	99.40	1	95.48	78.13	79.65	150.91	1014.17	370.41	3.25	4.66	0.68
0.67	0	4.39	1	95.22	69.12	71.42	1.05	8.30	3.51	0.18	1.33	0.30
0.67	0.125	28.18	1	92.93	69.08	67.83	1.76	9.65	5.64	0.26	0.21	0.22
0.67	0.25	60.88	1	90.61	68.41	63.99	3.84	14.12	11.93	0.57	1.70	0.15
0.67	0.5	85.63	1	90.46	71.20	63.46	10.26	29.50	33.30	0.18	0.26	0.50
0.67	1	95.41	1	91.78	70.84	64.29	28.68	91.50	94.59	1.59	3.21	2.21
0.67	2	98.44	1	93.49	70.86	68.09	69.84	339.82	230.09	1.01	7.57	4.56
0.67	4	99.40	1	95.55	71.41	71.37	150.67	1349.56	545.45	2.42	5.83	11.75
Average				**92.94**	**82.85**	**80.76**	**37.90**	**164.16**	**79.01**	**1.13**	**1.81**	**1.47**

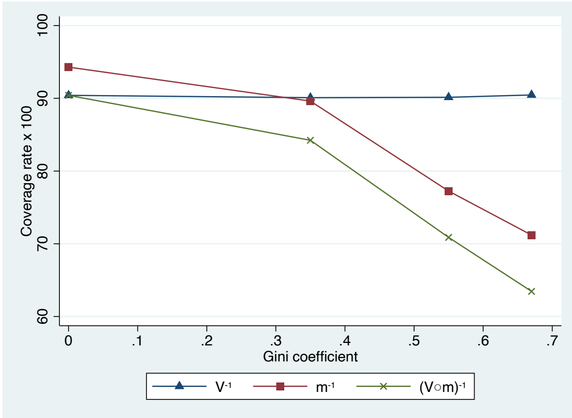
Notes: See Table 23.

C.4. Coverage, MSE, Bias ($d = 0.5$, $f = 0$, Random Heterogeneity at ES Level)

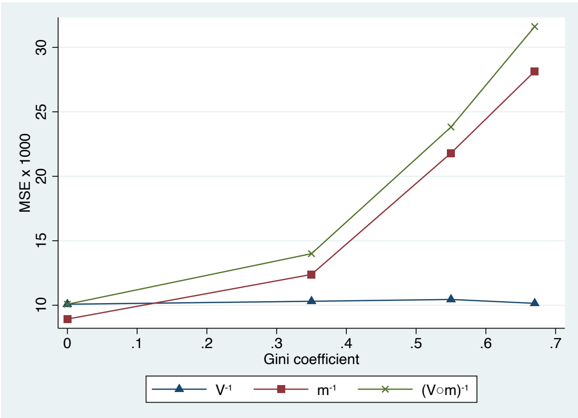
(a) Coverage; $\alpha_1 = 0$



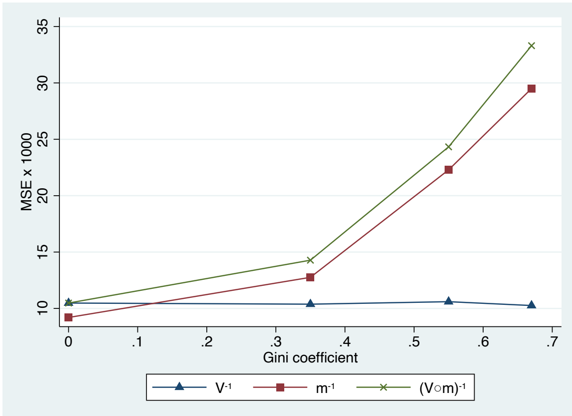
(b) Coverage; $\alpha_1 = 1$



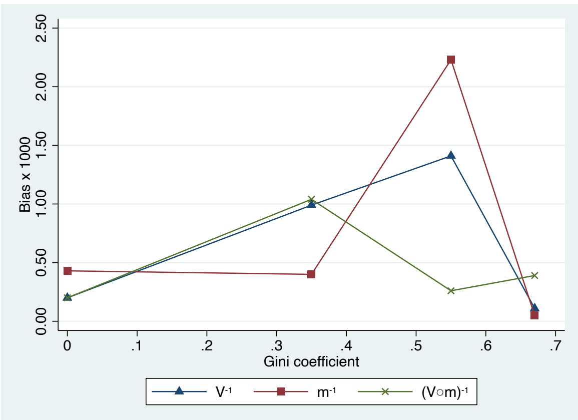
(c) MSE; $\alpha_1 = 0$



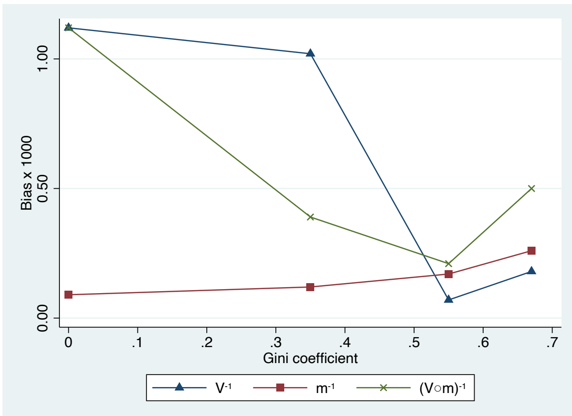
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: See Figure 35.

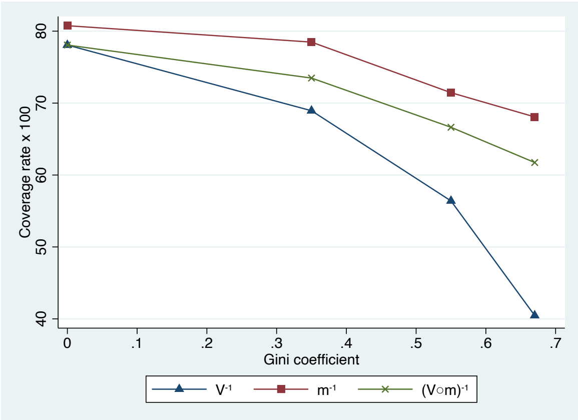
C.5. Simulation Results ($d = 0.25$, $f = 0$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.47	0	93.66	90.46	93.66	1.18	3.13	1.18	0.11	0.53	0.11
0	0.125	27.04	0	87.38	89.31	87.38	2.42	4.00	2.42	0.55	0.78	0.55
0	0.25	58.67	0	80.82	85.60	80.82	6.04	6.95	6.04	0.50	0.45	0.50
0	0.5	84.21	0	78.07	80.76	78.07	18.23	18.65	18.23	0.13	0.75	0.13
0	1	94.90	0	80.01	78.92	80.01	51.09	64.73	51.09	3.54	5.13	3.54
0	2	98.24	0	82.86	78.04	82.86	127.71	252.13	127.71	3.20	3.02	3.20
0	4	99.30	0	85.55	77.19	85.55	283.70	1004.75	283.70	4.63	1.51	4.63
0.35	0	4.41	0	92.93	86.35	88.32	1.21	4.00	1.61	0.65	0.01	0.42
0.35	0.125	25.93	0	82.90	84.35	83.19	2.99	5.12	2.99	0.06	0.41	0.43
0.35	0.25	57.00	0	72.93	81.26	76.46	8.16	8.37	7.22	1.29	1.93	1.37
0.35	0.5	83.01	0	68.94	78.48	73.48	25.21	20.69	20.87	0.73	0.49	1.85
0.35	1	94.52	0	69.77	76.44	75.37	75.76	72.30	60.99	4.82	6.83	5.02
0.35	2	98.10	0	71.81	75.52	78.69	194.42	278.31	153.38	3.57	5.30	0.21
0.35	4	99.23	0	76.02	75.14	82.14	436.47	1097.58	344.51	3.32	3.92	5.48
0.55	0	4.37	0	91.80	75.08	77.65	1.35	6.66	2.71	0.13	0.34	0.61
0.55	0.125	24.13	0	76.55	75.20	73.62	3.95	7.70	4.45	0.70	0.17	0.21
0.55	0.25	54.30	0	63.06	73.47	68.44	11.85	11.66	10.00	1.79	1.06	1.61
0.55	0.5	81.31	0	56.40	71.45	66.64	39.22	27.56	29.09	0.91	1.71	1.24
0.55	1	93.69	0	56.46	71.52	68.30	117.96	87.34	80.19	2.44	1.64	2.97
0.55	2	97.82	0	58.23	70.60	72.27	315.97	340.93	204.41	7.38	4.10	3.91
0.55	4	99.09	0	60.50	70.77	74.96	730.31	1316.89	472.10	12.01	18.63	2.54
0.67	0	4.26	0	90.15	68.04	71.57	1.51	8.77	3.56	0.11	0.61	0.37
0.67	0.125	20.29	0	64.58	68.79	66.98	6.19	9.89	5.84	0.18	0.93	0.26
0.67	0.25	47.36	0	48.23	67.77	63.24	19.59	14.53	12.56	1.60	1.49	1.85
0.67	0.5	75.96	0	40.47	68.08	61.74	68.34	31.74	34.83	0.14	1.32	0.15
0.67	1	91.48	0	39.63	66.70	61.82	216.17	103.83	102.49	2.77	6.16	8.72
0.67	2	96.97	0	41.47	68.01	67.05	616.18	381.34	252.25	4.46	1.83	3.56
0.67	4	98.71	0	44.49	67.66	69.61	1528.93	1514.69	596.46	7.25	8.45	6.46
0	0	4.29	1	93.80	91.51	93.80	1.16	3.02	1.16	0.37	0.37	0.37
0	0.125	26.92	1	87.47	89.74	87.47	2.41	3.93	2.41	0.08	0.16	0.08
0	0.25	58.73	1	81.18	85.59	81.18	6.01	6.96	6.01	0.30	0.40	0.30
0	0.5	84.14	1	77.93	81.09	77.93	18.21	18.59	18.21	0.93	0.70	0.93
0	1	94.90	1	79.61	78.79	79.61	51.29	64.98	51.29	1.53	0.27	1.53
0	2	98.25	1	82.85	77.60	82.85	126.29	256.67	126.29	6.61	10.04	6.61
0	4	99.30	1	85.88	77.65	85.88	285.11	984.82	285.11	5.89	16.82	5.89
0.35	0	4.40	1	93.51	86.82	88.83	1.21	4.00	1.59	0.22	0.42	0.51
0.35	0.125	25.82	1	82.54	84.14	82.42	3.02	5.16	3.08	0.08	0.12	0.50
0.35	0.25	57.19	1	73.77	81.68	76.63	8.24	8.26	7.27	0.03	0.38	0.42
0.35	0.5	83.14	1	68.60	78.15	74.06	26.17	21.52	21.52	2.32	1.62	2.21
0.35	1	94.47	1	68.92	76.00	75.45	75.76	72.13	60.44	3.21	2.16	2.01
0.35	2	98.10	1	72.88	75.27	78.96	191.29	282.38	148.80	6.98	12.49	5.91
0.35	4	99.23	1	76.18	75.50	81.94	436.11	1100.71	342.67	2.67	5.14	2.21
0.55	0	4.31	1	92.03	74.81	78.20	1.33	6.70	2.61	0.24	0.57	0.71
0.55	0.125	24.44	1	75.53	75.48	73.54	4.09	7.70	4.51	0.44	0.85	0.06
0.55	0.25	54.09	1	62.69	73.39	68.25	11.90	11.66	9.96	0.53	0.61	0.67
0.55	0.5	81.13	1	57.01	72.50	67.10	38.82	26.54	28.04	0.42	1.68	1.29
0.55	1	93.75	1	55.93	71.38	67.87	118.73	87.24	82.08	3.25	6.36	6.48
0.55	2	97.80	1	58.65	70.18	71.78	306.68	345.63	206.29	6.81	6.80	0.77
0.55	4	99.09	1	61.73	71.39	75.01	736.54	1323.71	473.93	3.63	10.18	5.68
0.67	0	4.28	1	89.97	68.22	71.69	1.52	8.54	3.49	0.10	0.55	0.00
0.67	0.125	20.58	1	64.12	68.10	66.70	6.37	9.97	5.88	0.47	0.77	0.58
0.67	0.25	47.45	1	48.01	68.34	62.63	20.70	14.54	12.61	0.91	1.78	2.42
0.67	0.5	75.93	1	40.55	67.85	60.78	68.98	31.52	34.54	1.92	0.64	1.93
0.67	1	91.47	1	39.64	67.11	62.74	219.60	104.67	100.79	0.51	4.87	2.76
0.67	2	96.98	1	42.75	67.75	67.96	606.52	379.55	245.06	1.52	4.47	2.13
0.67	4	98.71	1	44.86	67.17	69.51	1546.96	1530.21	604.79	12.67	6.86	8.83
Average				**69.90**	**75.79**	**75.37**	**175.59**	**239.74**	**103.27**	**2.39**	**3.17**	**2.24**

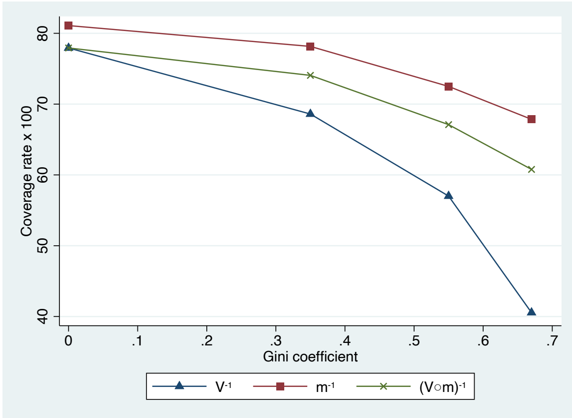
Notes: See Table 26.

C.6. Coverage, MSE, Bias ($d = 0.25$, $f = 0$, Random Heterogeneity at Study Level)

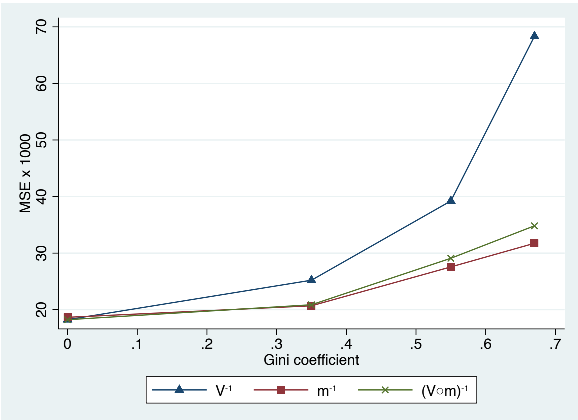
(a) Coverage; $\alpha_1 = 0$



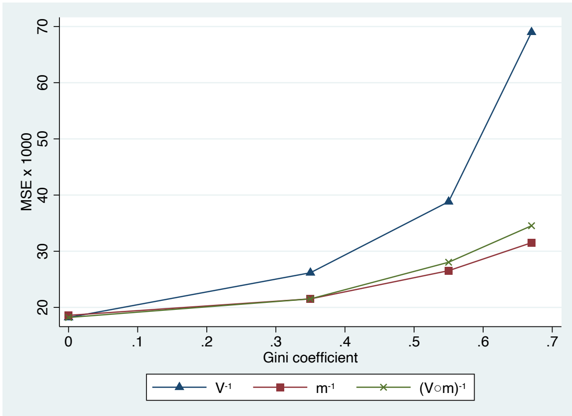
(b) Coverage; $\alpha_1 = 1$



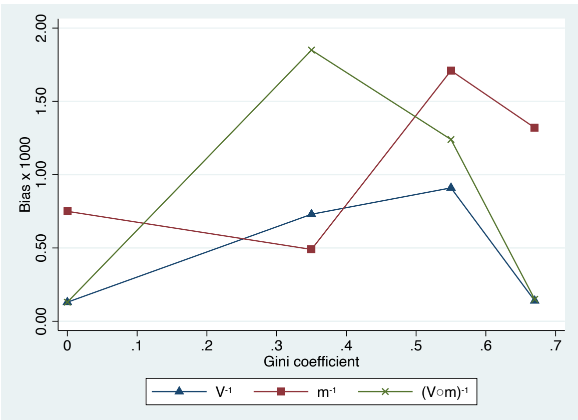
(c) MSE; $\alpha_1 = 0$



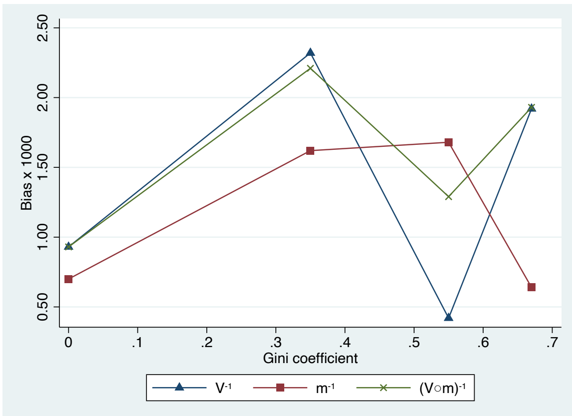
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: See Figure 37.

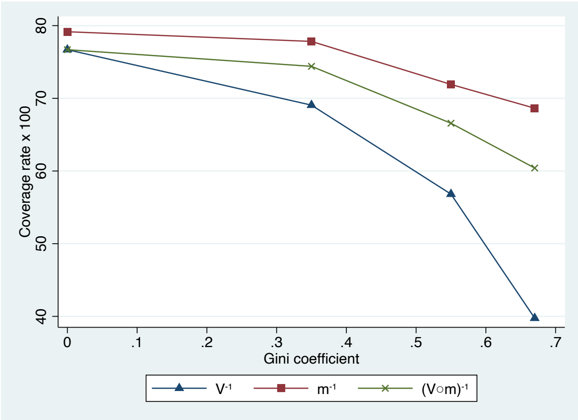
C.7. Simulation Results ($d = 0.5$, $f = 0$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.39	0	92.61	89.12	92.61	1.27	3.46	1.27	0.33	0.00	0.33
0	0.125	26.93	0	86.96	87.84	86.96	2.52	4.30	2.52	0.32	0.14	0.32
0	0.25	58.70	0	80.93	84.44	80.93	6.11	7.28	6.11	0.46	0.72	0.46
0	0.5	84.11	0	76.69	79.16	76.69	19.06	19.67	19.06	0.73	1.55	0.73
0	1	94.90	0	79.78	77.79	79.78	52.74	68.25	52.74	3.51	3.80	3.51
0	2	98.25	0	83.25	78.33	83.25	124.92	244.40	124.92	0.12	3.10	0.12
0	4	99.29	0	85.28	77.47	85.28	292.93	1003.02	292.93	2.85	3.77	2.85
0.35	0	4.37	0	91.22	84.63	87.50	1.36	4.39	1.69	0.29	0.10	0.42
0.35	0.125	25.65	0	81.64	82.85	81.49	3.18	5.47	3.16	0.59	0.93	0.57
0.35	0.25	56.68	0	72.34	80.61	75.95	8.44	8.63	7.28	1.25	0.65	0.56
0.35	0.5	83.10	0	69.06	77.83	74.40	26.01	21.42	21.47	3.03	1.63	1.83
0.35	1	94.47	0	69.75	75.87	75.64	75.17	72.32	61.40	1.92	1.98	2.83
0.35	2	98.09	0	73.24	75.80	79.25	184.54	275.25	147.68	4.48	9.53	4.50
0.35	4	99.23	0	75.54	74.53	81.59	440.19	1116.61	346.32	4.50	13.84	7.84
0.55	0	4.38	0	89.56	74.55	78.18	1.57	6.69	2.68	0.06	0.02	0.02
0.55	0.125	23.99	0	73.93	74.47	72.48	4.46	7.97	4.62	0.46	0.57	0.29
0.55	0.25	53.70	0	62.45	73.26	68.72	11.97	11.87	9.82	0.30	0.31	0.18
0.55	0.5	81.08	0	56.82	71.91	66.57	39.26	27.56	28.50	2.71	1.58	3.43
0.55	1	93.71	0	55.86	70.70	68.33	121.90	89.21	80.20	4.27	2.32	2.48
0.55	2	97.81	0	57.37	70.55	71.57	314.36	339.70	203.36	3.99	1.34	5.34
0.55	4	99.09	0	60.91	71.16	75.24	722.19	1315.05	457.48	19.63	22.78	12.51
0.67	0	4.06	0	85.11	68.17	70.76	1.98	8.66	3.56	0.37	1.46	0.22
0.67	0.125	19.98	0	61.70	67.85	66.85	6.84	10.16	5.90	0.58	0.12	0.98
0.67	0.25	46.68	0	47.05	67.98	63.32	20.97	14.61	12.40	1.43	1.97	2.61
0.67	0.5	75.33	0	39.76	68.64	60.41	70.82	32.27	36.04	0.51	0.07	0.31
0.67	1	91.39	0	39.00	67.41	62.29	222.08	103.75	102.29	4.20	1.64	1.10
0.67	2	96.95	0	42.54	67.50	65.75	599.20	386.47	257.30	5.24	6.79	0.00
0.67	4	98.70	0	44.86	66.97	69.32	1534.25	1544.72	603.58	16.56	9.40	4.37
0	0	4.43	1	92.35	89.23	92.35	1.29	3.46	1.29	0.34	0.52	0.34
0	0.125	26.62	1	86.18	87.23	86.18	2.53	4.48	2.53	0.35	0.12	0.35
0	0.25	58.46	1	80.59	84.34	80.59	6.17	7.38	6.17	1.30	1.45	1.30
0	0.5	84.10	1	77.61	80.40	77.61	18.21	18.72	18.21	1.66	1.15	1.66
0	1	94.90	1	79.70	78.26	79.70	51.98	66.18	51.98	1.98	3.69	1.98
0	2	98.25	1	82.16	77.26	82.16	128.33	256.32	128.33	1.09	2.62	1.09
0	4	99.30	1	85.54	77.16	85.54	292.61	1014.48	292.61	6.15	13.77	6.15
0.35	0	4.38	1	91.76	83.99	87.83	1.36	4.41	1.69	0.27	0.74	0.20
0.35	0.125	25.66	1	81.04	82.20	80.66	3.24	5.69	3.25	0.92	1.40	1.28
0.35	0.25	56.81	1	72.71	80.92	76.86	8.24	8.60	7.19	0.97	0.76	0.00
0.35	0.5	82.95	1	69.04	77.91	74.42	25.65	21.38	21.30	1.12	0.59	0.59
0.35	1	94.48	1	69.42	76.13	75.64	76.04	73.36	61.45	0.96	1.12	0.89
0.35	2	98.10	1	72.76	76.05	79.63	188.93	272.63	147.92	4.89	5.84	4.12
0.35	4	99.23	1	75.33	74.80	82.13	436.57	1126.70	343.98	4.00	7.69	0.67
0.55	0	4.29	1	89.19	75.22	78.14	1.58	6.61	2.64	0.41	0.01	0.11
0.55	0.125	23.72	1	73.36	73.78	73.17	4.34	8.08	4.47	0.55	0.22	0.16
0.55	0.25	53.67	1	62.13	72.85	68.75	12.36	12.21	10.14	0.71	0.09	0.83
0.55	0.5	80.97	1	55.89	71.76	67.08	39.50	27.21	28.26	1.23	0.06	1.79
0.55	1	93.65	1	56.16	70.36	68.24	118.07	91.06	80.88	1.05	0.05	1.02
0.55	2	97.79	1	57.99	70.91	71.50	316.08	336.81	203.30	12.82	1.98	6.39
0.55	4	99.09	1	61.46	70.09	74.71	719.51	1349.70	477.66	7.91	5.19	3.20
0.67	0	3.96	1	84.88	67.11	70.84	1.95	8.76	3.63	0.19	1.40	0.42
0.67	0.125	20.06	1	61.94	66.72	66.63	6.61	10.32	6.06	1.34	0.77	1.11
0.67	0.25	47.08	1	48.25	68.04	63.18	20.07	14.26	12.56	1.24	0.96	0.78
0.67	0.5	75.62	1	40.06	67.86	61.14	68.92	32.54	35.21	3.34	2.22	4.16
0.67	1	91.39	1	38.32	68.17	62.10	222.84	101.66	103.23	8.83	6.02	0.98
0.67	2	96.97	1	41.75	67.24	66.77	615.76	392.19	257.14	4.31	5.66	1.47
0.67	4	98.70	1	45.23	66.83	70.29	1534.92	1544.66	595.09	2.27	3.18	3.48
Average				**69.07**	**75.18**	**75.09**	**175.61**	**242.38**	**103.65**	**2.80**	**2.88**	**1.91**

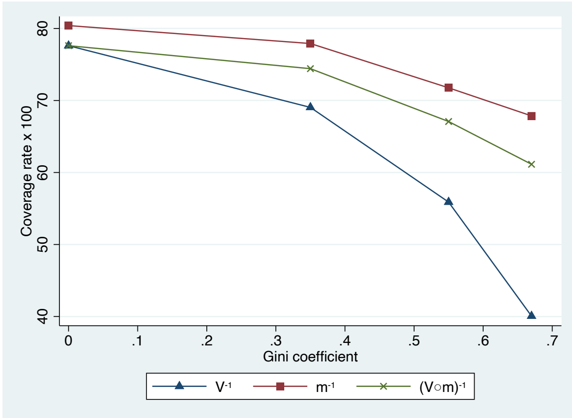
Notes: See Table 26.

C.8. Coverage, MSE, Bias ($d = 0.5$, $f = 0$, Random Heterogeneity at Study Level)

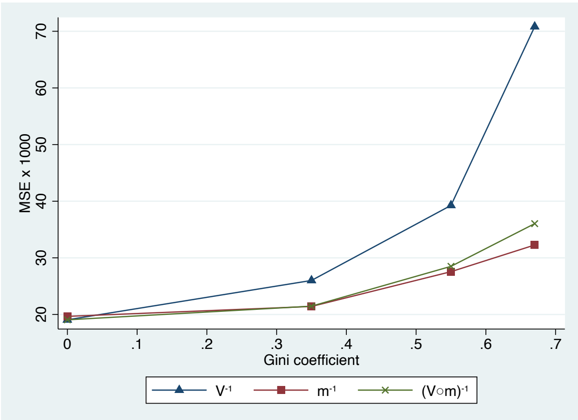
(a) Coverage; $\alpha_1 = 0$



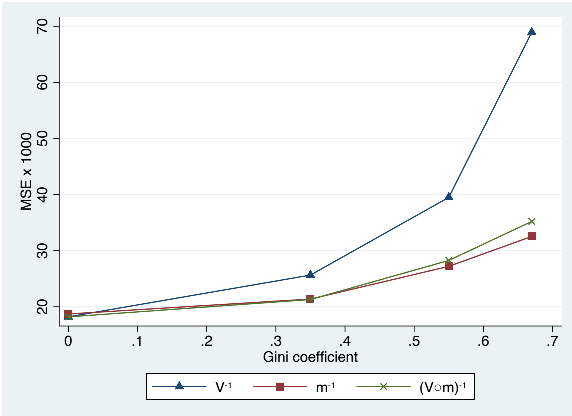
(b) Coverage; $\alpha_1 = 1$



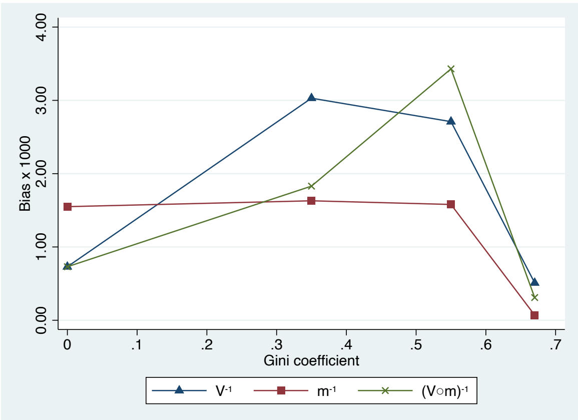
(c) MSE; $\alpha_1 = 0$



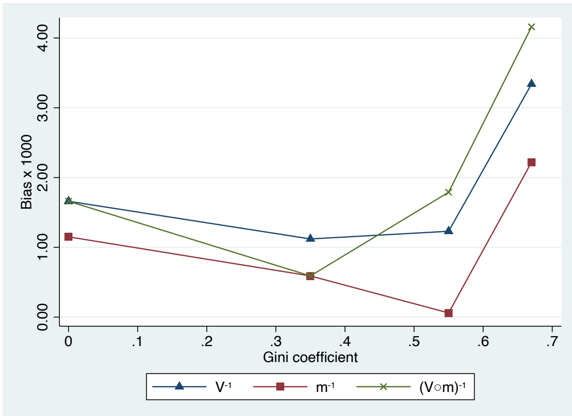
(d) MSE; $\alpha_1 = 1$



(e) Bias; $\alpha_1 = 0$



(f) Bias; $\alpha_1 = 1$



Notes: See Figure 37.

C.9. Simulation Results ($d = 0.5$, $f = 0.25$, Random Heterogeneity at Study Level)

Simulation				Coverage x 100			MSE x 1,000			Bias x 1,000		
G	σ_g	I^2	α_1	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$	V^{-1}	m^{-1}	$(V \circ m)^{-1}$
0	0	4.39	0	91.30	94.31	91.30	6.26	11.57	6.26	25.95	39.38	25.95
0	0.125	26.93	0	79.96	88.94	79.96	12.31	17.50	12.31	70.37	79.74	70.37
0	0.25	58.70	0	83.14	88.79	83.14	18.30	23.75	18.30	72.51	88.48	72.51
0	0.5	84.11	0	85.33	87.71	85.33	42.46	48.94	42.46	95.20	115.70	95.20
0	1	94.90	0	89.06	88.89	89.06	100.76	132.30	100.76	126.37	156.88	126.37
0	2	98.25	0	90.33	84.74	90.33	238.85	489.34	238.85	113.61	194.40	113.61
0	4	99.29	0	88.86	83.23	88.86	521.63	1521.09	521.63	78.36	270.12	78.36
0.35	0	4.37	0	88.56	89.72	85.03	7.51	15.68	8.49	32.70	45.26	29.97
0.35	0.125	25.65	0	78.13	84.83	74.94	13.67	21.62	14.53	69.40	79.26	67.49
0.35	0.25	56.68	0	80.00	83.93	78.38	21.59	29.53	22.50	71.80	86.69	69.23
0.35	0.5	83.10	0	81.41	82.70	80.04	50.47	61.02	52.43	90.97	114.94	89.77
0.35	1	94.47	0	84.68	83.78	83.99	124.86	167.91	130.34	111.26	151.26	116.69
0.35	2	98.09	0	85.56	81.15	86.68	299.86	556.87	284.79	104.21	173.40	104.52
0.35	4	99.23	0	82.53	79.52	86.20	703.14	1802.39	609.78	61.25	293.71	59.41
0.55	0	4.38	0	69.87	76.10	67.13	15.41	27.23	15.70	79.62	77.57	66.94
0.55	0.125	23.99	0	72.79	75.58	68.67	18.08	31.34	19.61	72.15	79.25	62.49
0.55	0.25	53.70	0	73.10	75.35	68.43	29.42	42.96	32.83	74.55	87.36	69.80
0.55	0.5	81.08	0	73.22	74.73	70.07	70.57	86.58	77.68	93.68	116.56	91.72
0.55	1	93.71	0	75.78	74.38	73.06	183.87	246.70	197.23	119.24	149.22	122.21
0.55	2	97.81	0	75.26	72.88	76.38	457.91	793.63	447.42	103.56	205.23	135.57
0.55	4	99.09	0	71.67	72.05	78.19	1097.55	2455.58	910.98	88.22	327.11	82.06
0.67	0	4.06	0	59.18	68.84	63.24	25.38	34.00	17.90	90.82	71.40	58.32
0.67	0.125	19.98	0	59.78	69.01	62.88	31.50	38.87	23.61	88.50	74.56	57.22
0.67	0.25	46.68	0	58.76	68.90	62.44	52.39	52.26	40.77	93.42	84.14	63.09
0.67	0.5	75.33	0	57.71	66.79	61.95	129.06	109.17	101.80	105.05	116.64	89.72
0.67	1	91.39	0	58.22	66.48	65.11	331.24	316.40	263.75	115.41	164.39	129.07
0.67	2	96.95	0	57.73	65.47	68.90	826.81	1001.84	589.07	126.56	217.63	129.79
0.67	4	98.70	0	54.01	67.17	72.96	1982.57	2837.99	1129.83	146.42	337.90	85.23
0	0	4.43	1	93.60	94.54	93.60	3.68	7.23	3.68	11.84	18.66	11.84
0	0.125	26.62	1	90.37	93.72	90.37	6.34	9.06	6.34	11.33	17.48	11.33
0	0.25	58.46	1	87.99	92.94	87.99	13.27	14.19	13.27	9.55	15.30	9.55
0	0.5	84.10	1	88.05	91.40	88.05	33.66	32.80	33.66	2.59	1.51	2.59
0	1	94.90	1	91.21	90.70	91.21	81.08	106.77	81.08	29.31	46.62	29.31
0	2	98.25	1	91.97	85.47	91.97	198.52	429.77	198.52	27.13	82.97	27.13
0	4	99.30	1	89.16	83.39	89.16	472.62	1430.30	472.62	1.66	179.25	1.66
0.35	0	4.38	1	91.44	90.54	88.40	4.28	9.70	5.15	11.49	19.41	13.16
0.35	0.125	25.66	1	88.24	90.05	86.20	7.10	11.77	8.07	7.50	16.23	9.79
0.35	0.25	56.81	1	84.22	88.01	82.22	15.61	18.53	17.31	7.29	14.00	11.21
0.35	0.5	82.95	1	84.53	87.15	83.17	39.87	42.75	44.56	9.74	3.13	4.02
0.35	1	94.48	1	86.96	86.14	85.85	103.78	136.49	109.59	30.23	45.38	23.49
0.35	2	98.10	1	86.75	80.99	87.62	266.96	530.58	252.85	33.71	74.79	18.69
0.35	4	99.23	1	82.91	79.02	86.39	641.98	1665.16	562.04	13.30	200.56	6.78
0.55	0	4.29	1	87.62	80.44	77.22	5.64	15.80	8.46	5.69	16.59	11.50
0.55	0.125	23.72	1	81.63	79.58	74.61	10.01	19.41	13.49	5.21	19.12	14.15
0.55	0.25	53.67	1	77.86	78.61	71.82	20.16	28.99	27.55	0.41	17.21	11.74
0.55	0.5	80.97	1	77.24	77.36	71.83	54.91	67.68	71.12	18.79	3.23	2.75
0.55	1	93.65	1	79.02	76.47	75.09	144.96	210.05	179.14	41.22	35.05	23.42
0.55	2	97.79	1	78.22	72.96	78.78	382.48	705.99	386.21	41.07	62.87	5.44
0.55	4	99.09	1	73.35	71.51	79.59	943.11	2278.46	778.51	34.58	213.98	5.84
0.67	0	3.96	1	76.61	73.57	70.18	10.83	20.80	11.46	5.04	16.30	11.72
0.67	0.125	20.06	1	68.72	71.98	66.51	17.00	25.40	18.26	9.76	17.93	12.73
0.67	0.25	47.08	1	62.85	70.93	63.95	34.94	39.88	37.32	19.75	15.60	10.03
0.67	0.5	75.62	1	59.80	70.45	63.90	94.27	87.84	95.37	48.84	8.74	7.19
0.67	1	91.39	1	63.49	69.35	67.68	260.86	271.05	242.22	67.65	29.90	15.06
0.67	2	96.97	1	61.90	67.52	71.64	686.61	895.93	525.48	89.04	55.11	3.91
0.67	4	98.70	1	57.80	67.56	73.06	1702.06	2594.67	1053.41	114.47	154.87	42.37
Average				**77.67**	**79.61**	**78.23**	**244.11**	**440.73**	**199.79**	**57.67**	**96.43**	**48.59**

Notes: See Table 28.