CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Essays on International and Financial Economics

Dissertation thesis

Author: Mgr. Josef Bajzík

Study program: Economics and Finance

Supervisor: Prof. PhDr. Tomáš Havránek, Ph.D.

Year of defense: 2024

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agree with the thesis being used for study and scientific purposes.

Prague, May 14, 2024

Josef Bajzík

Abstract

The dissertation consists of three papers that provide a brief summary of my research. In all of the papers I use meta-analytical approaches to study the publication bias, indicate key heterogeneity drivers in the given areas and suggest the implied values for studied parameters. In the first paper I focused on the Armington elasticity, which measures the elasticity between domestic and foreign goods and suggest the values of the parameters for different countries. In the second paper, I study the effect of changes in capital-based measures on lending for all three: the capital-to-asset ratio, regulatory capital ratio and capital requirements. In the third paper I study the effect of trading volume on stock returns on the financial markets. The general introduction binds the three chapters together, while the detailed abstracts for each paper is presented at the beginning of the respective chapters.

In the first paper, I scrutinize a key parameter in international economics called Armington elasticity. I reflect the context in which researchers obtain their estimates, examine the key drivers of the heterogeneity and account for inherent model uncertainty. I employ the Bayesian model averaging and several newly developed techniques for publication bias detection. I found that there is publication bias against small and insignificant estimates and that the differences between findings from the primary articles are best explained by aggregation, frequency, size, and dimension. Moreover, the mean elasticity implied by the literature corrected for publication bias is 3.

In the second paper, I focus on the relationship between bank capital in its three forms, in the capital-to-asset ratio, regulatory capital ratio and capital requirements; and lending. I show that the relationship between bank capital and lending evolves over time. It reflects the post crisis period accompanied by with demanding bank regulation and reduced profitability. Besides, my findings indicates that the literature fails to provide policymakers with reliable estimates of the effects of capital regulation on lending.

In the third paper, I study if the literature indicates the predictability of the stock returns based on trading volume performances. After correcting for publication bias, which distorts the estimates in primary studies, I found that the predictability of the stock returns vary with different markets and stock types. Besides the results in the primary studies are affected by different data characteristics, structural variations and methodologies, for instance by using monthly data or VAR models.

Acknowledgments

I would like to thank to my supervisor Prof. PhDr. Tomáš Havránek, Ph.D., for his guidance, suggestions and comments. I would like to thank my colleagues, doc. Zuzana Havránková, Phd., Mgr. Jiří Schwarz, Phd., Mgr. Simona Malovaná, Phd., Ing. Zuzana Gric, and doc. Martin Hodula, Phd. for their cooperation, sharing their insights a knowledge of respective fields. My gratitude belongs also my wife Stanka for her patience, our two sons for their motivation, my parents for encouragements, my sisters for inspiration and to Jesus Christ, who gives me the strength to do it all.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record
Bajzík, Josef: *Essays on International and Financial Economics.* Dissertation thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 225. Advisor: Prof. PhDr. Tomáš Havránek, Ph.D.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Meta-analysis is a quantitative review of all empirical results related to a given topic. It is an effective way to draw real inferences from the various findings of primary studies, which are often contradictiory. Moreover, meta-analysis systematically scrutinizes the differences in empirical findings. Since meta-analysis works with many independent variables, it addresses endogeneity problems the primary studies may suffer. It also compares the results qualitatively. Next, meta-analysis identifies study-invariant factors such as sample selection bias and measurement errors that may affect individual studies. It reveals the degree of influence of factors such as time and cross-country variations across the primary studies. Last but not least, it explains the impact of economic fundamentals, such as the degree of market development, on market efficiency (Kim *et al.* 2019).

A meta-analysis addresses publication bias as well as model uncertainty issues. I follow seminal works such as Havranek & Irsova (2017) and employ the most recent techniques for correcting for publication bias together with Bayesian model averaging (Raftery *et al.* 1997) and frequentist model averaging (Amini & Parmeter 2012). To correct for publication bias, I start with the graphical visualization proposed by Egger *et al.* (1997a). Then, I add simple formal tests using ordinary least squares (OLS), the between-effect and weighted least squares (WLS) (Stanley & Doucouliagos 2012). Furthermore, an extension of the formal tests is provided by means of the latest improvement suggested by Bom & Rachinger (2019). Moreover, the newly developed stem-based method (Furukawa 2019) complements the investigation.

In my current research work, I touch through meta-analysis on topics from international trade, monetary and macroprudential policy as well as the topics from financial markets. Even as I approach them usually through meta-analysis, I wrote several articles on the primary data or survey background also. My first-ever published paper was “Estimating the Armington Elasticity: The Importance of Study Design and Publication Bias”, where I, my supervisor, and other authors study the elasticity of substitution between domestic and foreign goods. I collected 3,524 estimates from primary studies, which was the largest meta-analytical dataset collected at that time. I consider this article a great success, as it was published in the Q1 Journal of International Economics and cited by several studies in top-tier journals such as The Economic Journal, Review of Economic Dynamics, and again the Journal of International Economics. Currently, the paper is still being cited, and has thus far yielded over 150 citations in Google Scholar. As a doctoral student, I am quite satisfied of such result.

From the collaborations in Czech National Bank came my second batch of research papers on monetary and macroprudential policies, their effects on economics, and their handling by central banks worldwide. Two out of these four articles are again meta-analytical, but one of the others is based on a survey, for example. I see as my greatest achievement the study “When Does Monetary Policy Sway House Prices? A Meta-Analysis”, published in the IMF Economic Review. The other four articles were also successful: “A Prolonged Period of Low Interest Rates in Europe: Unintended Consequences”, “Bank Capital, Lending, and Regulation: A Meta-Analysis”, and “Borrower-Based Macroprudential Measures and Credit Growth: How Biased is the Existing Literature?” were published in the Journal of Economics Surveys, while “Macroprudential Policy in Central Banks: Integrated or Separate? Survey among Academics and Central Bankers” found its place in the Journal of Financial Stability.

In my other studies, I focused on corporate finance and financial markets. From this field came my two solo-authored papers and one more meta-analysis. My first solo achievement was “Trading Volume and Stock Returns: A Meta-Analysis”, discussing the publication bias and other key drivers of heterogeneity in trading volume and its effect on stock returns. In the area of financial markets, I and my colleagues studied the effect of sentiment of both individual and institutional investors on stock returns in a paper called “Does Sentiment Affect Stock Returns? A Meta-Analysis across Survey-Based Measures”. Both of the latter papers were published in the Q2 journal International Review of Financial Analysis. My last published co-authored paper, called “Retail fund flows and performance: Insights from supervisory data”, was published in the

Q2 journal Emerging Markets Review. It explores inflows and outflows patterns in retail equity mutual funds related to past and future performance derived from detailed monthly security-level holdings of funds in the Czech Republic.

I approach the remainder of this thesis as a sample of my meta-analytical research efforts. Since it is customary to include three papers in the thesis, I have selected one meta-analytical paper each from three different fields – international trade, monetary and macroprudential policy, and financial markets. The last mentioned is a solo-authored paper.

In Chapter 2 I include the aforementioned paper on Armington elasticity that I wrote with my supervisor and other two colleagues from Institute of Economic Studies. Armington elasticity is another name for the elasticity of substitution between domestic and foreign goods, used in honor of Armington (1969), who first formulated a theoretical model of goods distinguished solely by place of origin. The paper focuses on this elasticity because it is a central topic in international trade and macroeconomics, relevant to the welfare effects of globalization (Costinot & Rodriguez-Clare 2014), trade balance adjustments (Imbs & Mejean 2015), and the exchange rate pass-through of monetary policy (Auer & Schoenle 2016). Any attempt to evaluate the effect of tariffs depends crucially on Armington elasticity.

In economic modelling, such a reaction is usually given by the (constant) elasticity of substitution between domestic and foreign goods. So, the magnitude of the elasticity used for calibration drives the models' conclusions. This phenomenon is shown, for example, by Schurenberg-Frosch (2015), who recomputes the results of 50 previously published models using different values of the elasticity. She concludes that with changes in the elasticity, the results change qualitatively in more than half of the cases. As Hillberry & Hummels (2013, p. 1217) put it, "it is no exaggeration to say that [*the elasticity*] is the most important parameter in modern trade theory." Yet, there is no consensus on its magnitude. In diverse countries and contexts, researchers obtain substantially different estimates (Feenstra *et al.* 2018). In this paper, we assign a pattern to these differences that will be useful for calibrating models in international trade and macroeconomics.

In all the models we run, we find evidence of strong publication bias in the estimates of the long-term Armington elasticity. The bias results in an exaggeration of the mean estimate by more than 50%. In contrast, we find no publication bias among the estimates of the short-term elasticity. One explanation consistent with these results is that the short-term elasticity is commonly believed to be small and less critical for policy questions, so there are few incentives to discriminate against insignificant (and even potentially harmful) elasticity estimates. Significant estimates of the long-term elasticity, in contrast, appear intuitive and desirable to many researchers (see, for example, the discussion in McDaniel & Balistreri 2003; Hillberry *et al.* 2005).

Our findings indicate that study characteristics are systematically associated with reported results. Among the 32 variables we constructed, the most important in model averaging are the ones related to the data used in the estimation: data frequency (monthly, quarterly, or annual), data dimension (time series, cross-section, or panel), and dataset size. We also find a systematic correlation between quality measures and the reported magnitude of the elasticity. Studies and estimates of higher quality (as measured by number of citations, publication in a refereed journal, quality of the journal, and preferences of the authors of the primary studies) tend to report larger estimates.

While publication selection creates an upward bias, estimates of lower quality seem to yield a downward bias. We exploit our large dataset and the relationships unearthed by Bayesian model averaging to compute a mean effect corrected for publication bias but conditional on the design of the most reliable studies. We also report implied mean elasticities for individual countries.

In Chapter 3 I and my colleagues focused on the effect of changes to bank capital on the extension of bank credit. The importance of quantifying this relationship has been one of the most pivotal research questions for almost two decades. The topic was given particular attention following the onset of the 2007-2009 Global Financial Crisis (GFC), when the likelihood of a credit crunch was under debate, and again when the first quantitative easing programs were gradually implemented. The question has reemerged more recently with the gradual implementation of Basel III and increasing use of macroprudential policy instruments.

A conspicuous feature of the bank capital - lending literature is that the bank capital ratio may change for various reasons, ranging from regulatory (Peek & Rosengren 1997; De Jonghe *et al.* 2020) to economic and managerial (Houston *et al.* 1997; Berrospide & Edge 2010b; Gambacorta & Marques-Ibanez 2011a). As a result, there is a wide range of possible outcomes when quantifying the impact of changes in bank capital on bank lending. On the one hand, an increase in the bank capital ratio due to the introduction of new capital regulation may dampen bank lending activities, as banks may try to avoid the higher costs of financing loans with capital (De Jonghe *et al.* 2020). On the other hand, a general increase in the bank capital (equity) ratio due to, for example, bank profit accumulation should be reflected in an increase in lending, suggesting a positive effect (Berrospide & Edge 2010b).

In this paper, we conduct a thorough review of the empirical literature on how changes in bank capital affect credit dynamics. Throughout the literature, there are three expressions of the capital ratio: a simple capital-to-asset ratio, a regulatory capital ratio that includes Common Equity Tier 1, Tier 1, and Tier 2 capital over risk-weighted exposures, and a capital requirements ratio that is defined as capital requirements over risk-weighted exposures. We note from Figure 3.1 that the literature exhibits significant fragmentation regarding the estimated coefficients that goes beyond the different expression of capital ratios. To explain the differences, we collect an additional 40 variables reflecting the context in which the estimates were produced. The newly created database allows us not only to derive an "average" effect but also to explain why estimates vary across different studies and describe what the most commonly employed empirical strategy is.

Our findings indicate that various study characteristics are systematically associated with the reported results. Among the 40 variables we construct, the most important for model averaging are those related to data, the estimation technique and cross-country or regional differences. Specifically, we find that single-country studies with larger sample sizes positively correlate with the collected semi-elasticities, while studies shielded from omitted variable bias with more favorable publication characteristics are generally negatively correlated with the reported estimates. Apart from data characteristics, estimates of the effect of changes to the simple capital-to-asset ratio are also dependent on the variables reflecting the macro-financial characteristics of the countries analyzed. The heterogeneity in the estimates based on the regulatory capital ratio can thus mostly be explained by model specification. In the case of the literature on capital requirements, the standard error is the most important variable in explaining the variation in the reported estimates. Large standard errors are associated with more negative estimates, supporting the existence of publication bias in this category.

We perceive the contribution of this paper to be threefold. First, quantifying the effect of changes to bank capital on the supply of credit is of utmost importance to policymakers. Obtaining a comprehensive overview of the literature's findings goes well beyond the scope of individual studies that are, by nature, very selective. Second, we show the caveats associated with modelling the relationship between bank capital and lending and inform about the most commonly employed practices. Third, we present some indications that the relationship is changing over time and discuss the implications this might have for correctly estimating and assessing the impact of capital regulation.

In the Chapter 4 I close with solo-authored article studying the effect of trading volume on stock returns. I investigated this question since it has attracted the traders for decades and recently come to the fore-front again due to the expanding market of online platforms and the COVID crisis. Beaver (1968) said: "An important distinction between the price and volume tests is that the former reflects changes in the expectations of the market as a whole while the latter reflects changes in the expectations of individual investor." Morgan (1976) continues with the suggestion that volume is connected with systematic risk and, thus with stock returns. Thus, the trading volume falls among the possible determinants of the stock returns. Fama & French (1992; 1993; 1996) and Jegadeesh & Titman (1993; 1995).

Other reasons to study the trading-volume effect were found by Karpoff (1987). He noted, first, that this type of research provides insight into financial market structure. Second, he mentioned that such research is seminal for event studies that use price and volume data to draw conclusions. Third, he added that the return-volume relationship has significant implications for futures market research. These suggestions make the findings related to this topic even more valuable. The first studies discussing the price-volume relationship originated in the US in the 1960s (Granger & Morgenstern 1963; Godfrey *et al.* 1964). The focus on the US continued in the following decades (e.g., Crouch 1970; Jain & Joh 1988). By the turn of the millennium, researchers from every continent had begun to show interest in the topic.

I decided to thoroughly investigate the trading volume - stock return relationship through a meta-analysis. I found publication bias, at least in the more recent papers. The mean after correction for publication bias has a negligible value. Moreover, other study-specific aspects affect the corrected mean. The results of both Bayesian model averaging (BMA) and frequentist model averaging (FMA) indicate that data characteristics, structural variation and different methodological approaches explain a large part of the inconsistency in the primary results. For example, usage of monthly data or VAR models makes the effect of trading volume on returns substantially more negative.

Other causes of variation include the type of stocks and country of origin. For instance, this analysis reveals that the trading volume may predict the stock returns of non-financial firms. On the other hand, the effects of trading volume on the stocks of banks or general indices turn out to be insignificant. These conclusions were found worldwide. The same holds for markets in North America, Europe, Australia, and Asia. Only the estimates for developing countries differ. Trading volume predicts stock returns in emerging markets better than in developed ones. So, in developing markets, one can partially predict via trading volume the returns development of any stock. Thus, one should keep in mind the specifics of each stock when forming a portfolio, calibrating a model, preparing a trading strategy, or conducting research.

Chapter 2

Estimating the Armington Elasticity: The Importance of Data Choice and Publication Bias

Josef Bajzik, Tomas Havranek, Zuzana Irsova, Jiri Schwarz

Abstract A key parameter in international economics is the elasticity of substitution between domestic and foreign goods, also called the Armington elasticity. Yet estimates vary widely. We collect 3,524 reported estimates of the elasticity, construct 34 variables that reflect the context in which researchers obtain their estimates, and examine what drives the heterogeneity in the results. To account for inherent model uncertainty, we employ Bayesian and frequentist model averaging. We present the first application of newly developed non-linear techniques to correct for publication bias. Our main results are threefold. First, there is publication bias against small and statistically insignificant elasticities. Second, differences in results are best explained by differences in data: aggregation, frequency, size, and dimension. Third, the mean elasticity implied by the literature after correcting for both publication bias and potential misspecifications is 3.

This paper was published in *Journal of International Economics [2020, 127, pp. 1-23, DOI: 10.1016/j.jinteco.2020.103383].* We thank Chishio Furukawa, Andrés RodrĂguez-Clare, Tom Stanley, and two anonymous referees for their helpful comments. I acknowledge support from Charles University (PRIMUS/17/HUM/16) and the Czech Science Foundation (21-09231S), Havranek and Irsova acknowledge support from the Czech Science Foundation (19-26812X), and Schwarz acknowledges support from the Czech Science Foundation (18-02513S). The views expressed here are ours and not necessarily those of the Czech National Bank.

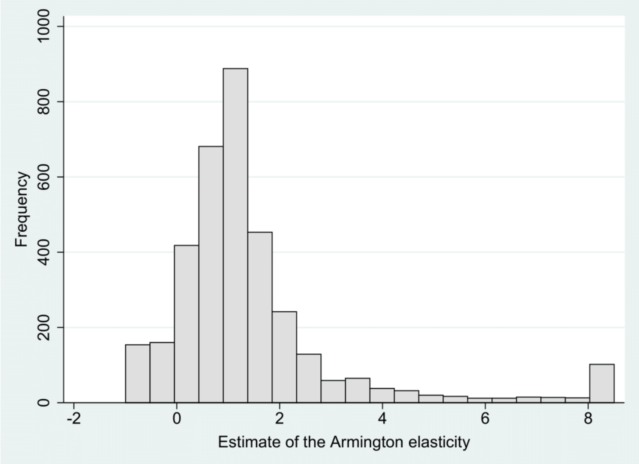
2.1 Introduction

How does the demand for domestic versus foreign goods react to a change in relative prices? The answer is central to a host of research and policy problems in international trade and macroeconomics: the welfare effects of globalization (Costinot & Rodriguez-Clare 2014), trade balance adjustments (Imbs & Mejean 2015), and the exchange rate pass-through of monetary policy (Auer & Schoenle 2016), to name but a few. Any attempt to evaluate the effect of tariffs in particular depends crucially on the assumed reaction of relative demand to relative prices. In most models, the reaction is governed by the (constant) elasticity of substitution between domestic and foreign goods. The size of the elasticity used for calibration often drives the conclusions of the model, as shown by Schurenberg-Frosch (2015), who recomputes the results of 50 previously published models using different values of the elasticity. She finds that, with plausible changes in the elasticity, the results change qualitatively in more than half of the cases. As Hillberry & Hummels (2013, p. 1217) put it, "it is no exaggeration to say that [*the elasticity*] is the most important parameter in modern trade theory."

Yet no consensus on the magnitude of the elasticity exists. In different contexts, researchers tend to obtain substantially different estimates, as observed by Feenstra *et al.* (2018) and many commentators before them. In this paper we assign a pattern to these differences, a pattern that we hope will be useful for calibrating models in international trade and macroeconomics. The elasticity of substitution between domestic and foreign goods is commonly called the Armington elasticity, in honor of Armington (1969), who first formulated a theoretical model featuring goods distinguished solely by the place of origin. The first estimates of the elasticity followed soon afterward, and many thousand have been published since. As the Armington-style literature turns 50, the time is ripe for taking stock. We collect 3,524 estimates of the elasticity of substitution between domestic and foreign goods and construct 34 variables that reflect the context in which researchers produce their estimates.

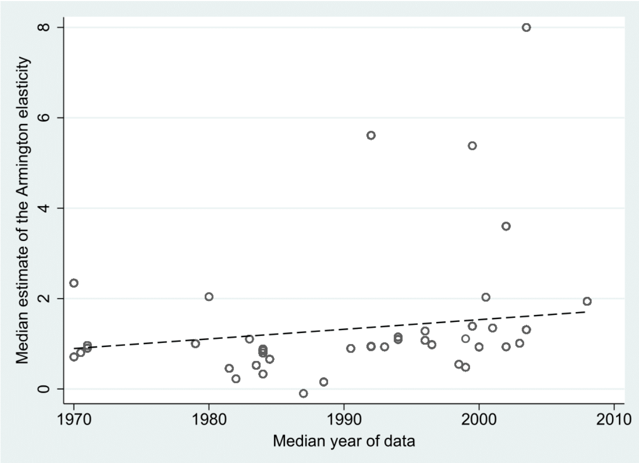
A bird's-eye view of the literature (Figure 2.1 and Figure 2.2) shows four stylized facts, three of which corroborate the common knowledge in the field. First, the estimates of the elasticities vary substantially. A researcher wishing to calibrate her policy model has plenty of degrees of freedom; she can easily find empirical evidence for any value of the elasticity between 0 and 8. Such plausible (that is, justifiable by some empirical evidence) changes in the elas-

Figure 2.1: The reported elasticities are often around 1 but can vary widely



Notes: The figure shows the histogram of the estimates of the macro-level Armington elasticity reported in individual studies. Large values are winsorized for ease of exposition.

Figure 2.2: The mean and variance of reported elasticities increase over time



Notes: The vertical axis measures median estimates of the macro-level Armington elasticity reported in individual studies. The horizontal axis measures the median year of the data used in the corresponding study.

ticity can have decisive effects on the results of the model. For example, Engler & Tervala (2018) show that changing the elasticity from 3 to 8 more than doubles the estimated welfare gains from the Transatlantic Trade and Investment Partnership. Second, the median estimated elasticity in the literature is 1, and many estimates are close to that value. Third, the reported elasticity seems to be increasing in time, but it is not clear whether the apparent trend reflects fundamental changes in preferences or improved data and techniques used by more recent studies.

Finally, the fourth stylized fact is that newer studies show more disagreement on the value of the elasticity of substitution. That is, instead of converging to a consensus value, the literature diverges. The increased variance in the estimated elasticities provides additional rationale for a systematic evaluation of the published results. For this evaluation we use the methods of meta-analysis, which were originally developed in (or inspired by) medical research. Recent applications of meta-analysis in economics include Card *et al.* (2018) on the effectiveness of active labor market programs, Anderson *et al.* (2018) on the impact of government spending on poverty, and Havranek & Irsova (2017) on the border effect in international trade. An important problem inherent in meta-analysis is model uncertainty because for many control variables capturing the study design, little theory exists that can help us determine whether they should be included in the baseline model. To address this issue, we use both Bayesian (Raftery *et al.* 1997; Eicher *et al.* 2011) and frequentist (Hansen 2007; Amini & Parmeter 2012) methods of model averaging (Steel 2020, provides an excellent description of these techniques).

Meta-analysis also allows us to correct for potential publication bias in the literature. Publication bias arises when, holding other aspects of study design constant, some results (for example, those that are statistically insignificant at standard levels or have the "wrong" sign) have a lower probability of publication than other results (Stanley 2001).[1] For example, in the context of the elasticity of substitution, it is safe to assume that its sign is positive: a negative value is not compatible with any commonly applied model of preferences. Similarly, it is difficult to interpret a zero elasticity. Thus, from the point of view of an individual study, it makes sense not to report such unintuitive estimates—or find a specification where the elasticity is positive—because

[1]Publication bias has recently been discussed, among others, by Brodeur *et al.* (2016), Bruns & Ioannidis (2016), Christensen & Miguel (2018), Brodeur *et al.* (2020), and Blanco-Perez & Brodeur (2020).

non-positive elasticity suggests that something is wrong with the data or the estimation technique. Nevertheless, non-positive estimates will occur from time to time simply because of sampling error; for the same reason, researchers will sometimes obtain estimates much larger than the true value. If large estimates (which are still intuitive) are kept but non-positive ones are omitted, an upward bias arises. Paradoxically, publication bias can thus improve inferences drawn from individual studies (if they avoid making central conclusions based on negative or zero elasticities) but inevitably bias inferences drawn from the literature as a whole. Ioannidis *et al.* (2017b) shows that, in economics, the effects of publication selection are dramatic and exaggerate the mean reported estimate twofold.

To correct for publication bias, we use meta-regression techniques based on Egger *et al.* (1997b) and their extensions and three new non-linear techniques developed specifically for meta-analysis in economics. The first one is due to Ioannidis *et al.* (2017b) and relies on estimates that are adequately powered. The second technique was developed by Andrews & Kasy (2019) and employs a selection model that estimates the probability of publication for results with different p-values. The third non-linear technique is the so-called stem-based method by Furukawa (2019), a non-parametric estimator that exploits the variance-bias trade-off. As far as we know, the latter two estimators have not been applied so far apart from illustrative examples outlined by Andrews & Kasy (2019) and Furukawa (2019).

In all the models we run, linear or non-linear, Bayesian or frequentist, we find evidence of strong publication bias in the estimates of the long-run Armington elasticity. The bias results in an exaggeration of the mean estimate by more than 50%. In contrast, we find no publication bias among the estimates of the short-run elasticity. One explanation consistent with these results is that the short-run elasticity is commonly believed to be small and less important for policy questions, so there are few incentives to discriminate against insignificant (and even potentially negative) estimates of the elasticity. Large estimates of the long-run elasticity, in contrast, appear intuitive and desirable to many researchers (see, for example, the discussion in McDaniel & Balistreri 2003; Hillberry *et al.* 2005).

Our findings indicate that the study characteristics are systematically associated with the reported results. Among the 34 variables we construct, the most important are the ones related to the data used in the estimation. We find that, ceteris paribus, using more aggregated data yields smaller estimates of the elasticity. Annual data bring substantially smaller elasticities compared to monthly and quarterly data. If a study uses cross-sectional data, it is more likely to report larger estimates of the elasticity than if time-series data are used. Our results also suggest that employing a small number of observations and ignoring endogeneity in the estimation yields a downward bias. Finally, we find systematic correlation between measures of quality and the magnitude of the reported elasticity. Studies of higher quality (as measured by the number of citations, publication in a refereed journal, and the RePEc impact factor of the outlet) tend to report larger estimates.

Therefore, while publication selection creates an upward bias, many questionable method choices seem to create a downward bias. We exploit the relationships unearthed by Bayesian model averaging to compute a mean effect corrected for publication bias, misspecification biases, and conditional on the maximum quality defined based on the peer-review status, the publication outlet, and the number of citations. The resulting elasticity reaches 3, and we interpret the number as our best guess (based on the available empirical literature published during the last five decades) for how to calibrate a model that allows for only one parameter to govern the aggregate elasticity of substitution between domestic and foreign goods—for example, an open economy dynamic stochastic general equilibrium model of the type used in many central banks. We also report these aggregate elasticities for individual countries and provide information in the online appendix that allows other researchers to use our data to compute the elasticities for individual industries.

The remainder of the paper is structured as follows. Section 2.2 briefly describes how the Armington elasticity is estimated and how we collect data from primary studies. Section 2.3 tests for publication bias in the literature. Section 2.4 explores heterogeneity and computes the aggregate elasticity corrected for publication and misspecification biases. Section 2.5 concludes the paper. An online appendix at meta-analysis.cz/armington provides the data and codes.

2.2 Collecting the Elasticity Dataset

The derivation of the Armington elasticity follows a two-stage optimization process (please refer to Hillberry & Hummels 2013; Feenstra *et al.* 2018, for a more detailed treatment than we have the space to offer here): in the first stage, the consumer with a CESutility function:

$$u(Q_D, Q_M) = \left(\beta \cdot Q_D^{(\sigma-1)/\sigma} + (1-\beta) \cdot Q_M^{(\sigma-1)/\sigma}\right)^{\sigma/(\sigma-1)}, \tag{2.1}$$

allocates her total spending to various product categories following her budget constraint with a given general price index. The consumer thus chooses a quantity of the composite good $Q_D + Q_M$, her aggregate demand for goods produced in her home country (D) and foreign countries (M). In the second stage, the consumer decides what proportion of domestic and foreign goods to consume while minimizing her expenditures $Q_D \cdot P_D + Q_M \cdot P_M$ or maximizing her utility. Utility maximization subject to the budget constraint or cost minimization subject to the utility function both imply that the marginal rate of substitution between domestic and foreign goods should equal the corresponding price ratio (Welsch 2008). The first-order condition follows:

$$\frac{Q_M}{Q_D} = \left[\frac{\beta}{1-\beta} \cdot \frac{P_D}{P_M}\right]^{\sigma}, \tag{2.2}$$

where the quantity of domestic goods Q_D and foreign goods Q_M is related to the corresponding domestic price P_D and import price P_M. β is a distribution parameter between the domestic and the foreign good, and σ denotes the Armington elasticity. For estimation, the first-order condition is commonly log-linearized:

$$\log\left(\frac{Q_M}{Q_D}\right) = \underbrace{\sigma \log\left(\frac{\beta}{1-\beta}\right)}_{\text{Constant}} + \underbrace{\sigma}_{\text{Armington elasticity}} \log\left(\frac{P_D}{P_M}\right) + e. \tag{2.3}$$

As the main building block of our dataset, we collect estimates of σ from the literature. Several recent papers, such as Aspalter (2016) or Feenstra *et al.* (2018), call this type of Armington elasticity a *macro-elasticity*. A macro-elasticity governs the substitution between home and foreign goods, where varieties from different foreign countries are aggregated into one composite good. A *micro-elasticity*, on the other hand, governs the substitution among the varieties of foreign goods and thus differentiates among the specific countries of origin (Balistreri *et al.* 2010). For comparability, in this paper, we focus on macro-elasticities.

We need each study to report a measure of uncertainty of its estimates. Such a measure, which is necessary to test for the potential presence of publication bias in the literature, can be either the standard error or other metrics

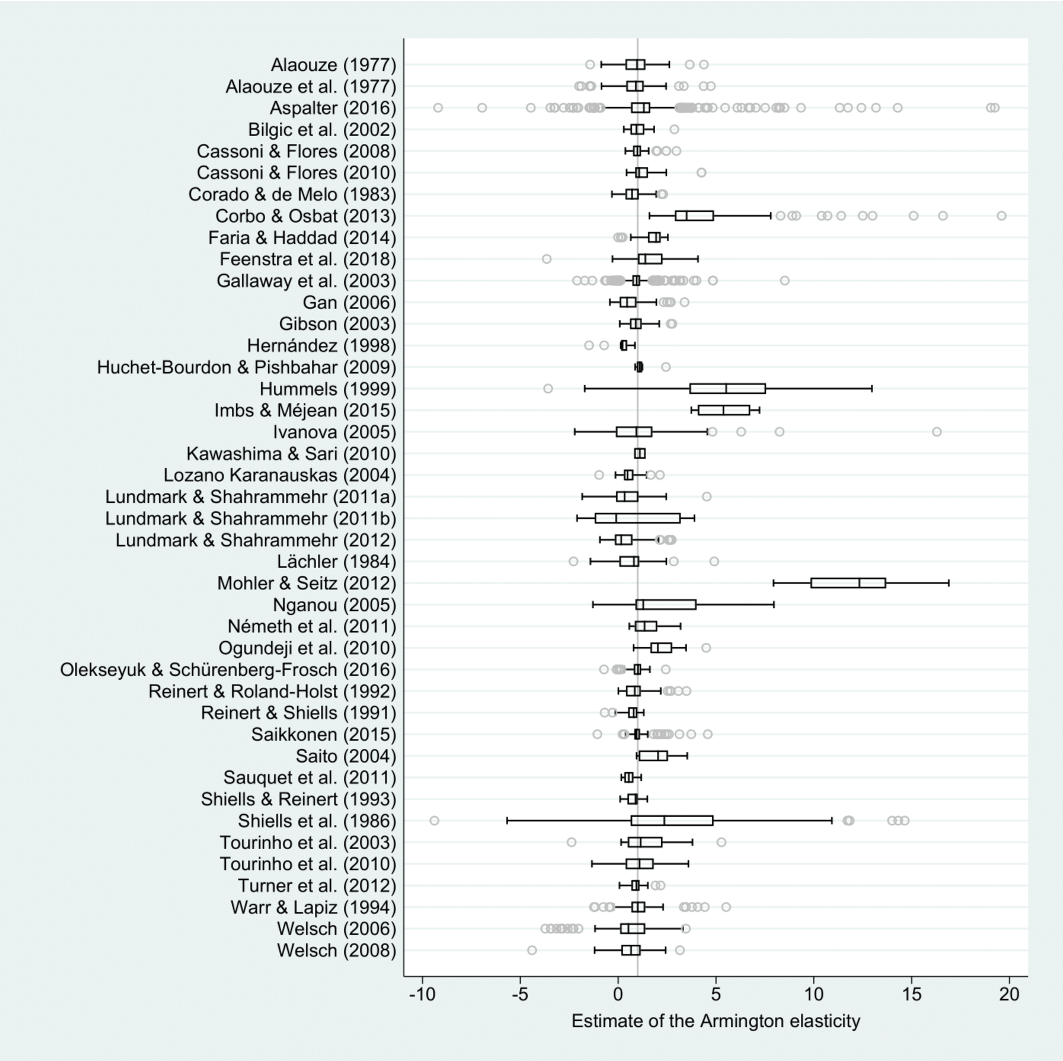
Table 2.1: Studies included in the meta-analysis

Alaouze (1977)	Lundmark & Shahrammehr (2011b)
Alaouze *et al.* (1977)	Lundmark & Shahrammehr (2012)
Aspalter (2016)	Imbs & Mejean (2015)
Bilgic *et al.* (2002)	Mohler & Seitz (2012)
Cassoni & Flores (2008)	Nemeth *et al.* (2011)
Cassoni & Flores (2010)	Nganou (2005)
Corado & de Melo (1983)	Ogundeji *et al.* (2010)
Corbo & Osbat (2013)	Olekseyuk & Schurenberg-Frosch (2016)
Faria & Haddad (2014)	Reinert & Roland-Holst (1992)
Feenstra *et al.* (2018)	Reinert & Shiells (1991)
Gallaway *et al.* (2003)	Saikkonen (2015)
Gan (2006)	Saito (2004)
Gibson (2003)	Sauquet *et al.* (2011)
Hernandez (1998)	Shiells & Reinert (1993)
Huchet-Bourdon & Pishbahar (2009)	Shiells *et al.* (1986)
Hummels (1999)	Tourinho *et al.* (2003)
Ivanova (2005)	Tourinho *et al.* (2010)
Kawashima & Sari (2010)	Turner *et al.* (2012)
Lachler (1984)	Warr & Lapiz (1994)
Lozano Karanauskas (2004)	Welsch (2006)
Lundmark & Shahrammehr (2011a)	Welsch (2008)

recomputable to the standard error. This requirement prevents us from using a dozen empirical papers, including the highly cited contribution by Broda & Weinstein (2006). For similar reasons, we drop a few estimates for which uncertainty measures are incorrectly reported (for example, when the reported standard errors are negative or when the reported confidence intervals do not include the point estimate). The final dataset is an unbalanced one because some studies report more estimates than other studies. We choose to include all the reported estimates because it is often unclear which estimate is the one preferred by the author; moreover, including more estimates obtained using alternative methods or datasets increases the variation we can exploit by meta-analysis.

The first step in a meta-analysis is the search for relevant studies. Building on the comprehensive surveys by McDaniel & Balistreri (2003) and Cassoni & Flores (2008), we design our search query in Google Scholar in a way that shows the well-known studies estimating the Armington elasticity among the first hits. The final query along with the dataset is available online at meta-analysis.cz/armington. We also go through the references of the most recent studies and obtain other papers that might provide empirical estimates of the elasticity. While the keywords we use are specified in English, we do not exclude any study based on the language of publication: several papers written in Spanish (e.g. Hernandez 1998; Lozano Karanauskas 2004) and Portuguese (Faria & Haddad 2014) are included. We add the last study in March 2018 and terminate the literature search. The final set of studies that fulfill all requirements for meta-analysis is reported in Table 2.1; our sample consists of 3,524 estimates from 42 papers.

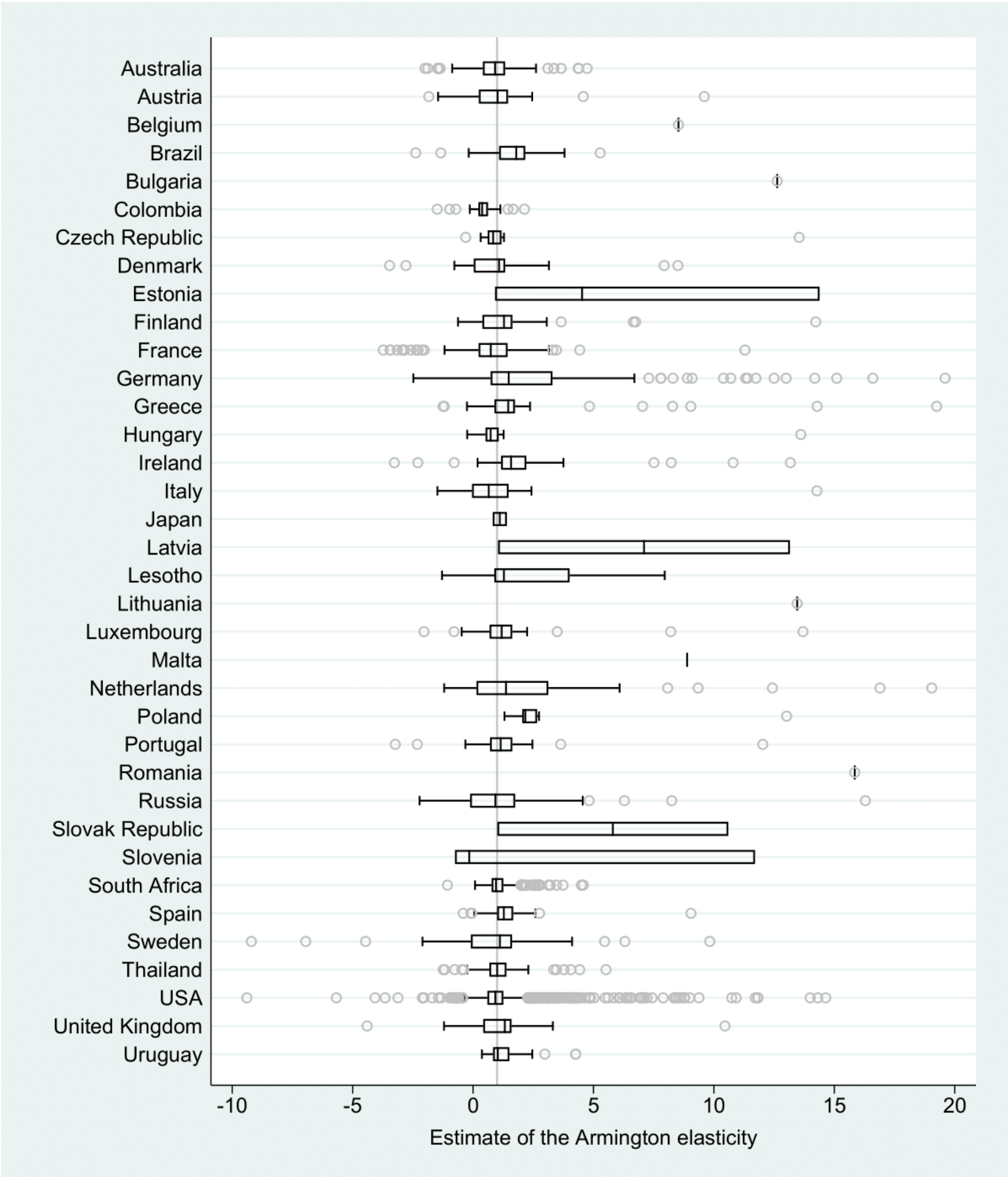
Figure 2.3: Estimates vary both within and across studies



Notes: The figure shows a box plot of the estimates of the Armington elasticity reported in individual studies. The length of each box represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. The dots show the outlying estimates with extreme values stacked at the values denoted as 'outliers.' The solid vertical line denotes unity.

The oldest study in our sample was published in 1977 and the most recent one in 2018, thereby covering more than 40 years of research. The mean reported elasticity is 1.5. Given that there are a few dramatic outliers in our data (their values climb to approximately 50 in absolute value), we winsorize

Figure 2.4: Estimates vary both within and across countries



Notes: The figure shows a box plot of the estimates of the Armington elasticity reported for individual countries. The length of each box represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. The dots show the outlying estimates with extreme values stacked at the values denoted as 'outliers.' The solid vertical line denotes unity.

the estimates at the 2.5% level; the mean is not affected by winsorization, and our results hold with alternative winsorizations at the 1% and 5% levels. Approximately 10% of the estimates are negative and commonly believed to occur due to misspecifications in the demand function and problems with import prices (Shiells *et al.* 1986). More than half of the estimates are larger than unity, which suggests that domestic and foreign goods can often be expected to form gross substitutes. Nevertheless, estimates differ greatly both within and between individual studies and home countries, as Figure 2.3 and Figure 2.4 demonstrate. To assign a pattern to this variance, for each estimate, we collect 43 explanatory variables describing various characteristics of data, home countries, methods, models, and quality; these sources of heterogeneity are examined in detail in Section 2.4.

Table 2.2: Armington elasticities for different subsets of data

		Unweighted			Weighted		
	No. of obs.	Mean	95% conf. int.		Mean	95% conf. int.	
Temporal dynamics							
Short-run effect	556	0.88	0.83	0.93	0.91	0.85	0.98
Long-run effect	2,968	1.56	1.49	1.63	1.74	1.65	1.82
Data characteristics							
Monthly data	488	1.04	0.97	1.11	1.18	1.12	1.24
Quarterly data	745	1.22	1.09	1.34	2.64	2.41	2.87
Annual data	2,291	1.62	1.54	1.70	1.32	1.25	1.40
Structural variation							
Primary sector	366	0.83	0.70	0.95	0.73	0.61	0.85
Agriculture, forestry, and fish.	*260*	*0.92*	*0.77*	*1.06*	*0.77*	*0.63*	*0.91*
Mining and quarrying	*103*	*0.58*	*0.33*	*0.84*	*0.38*	*0.14*	*0.62*
Secondary sector	3,044	1.46	1.40	1.52	1.40	1.34	1.46
Manufacturing	*2,963*	*1.46*	*1.40*	*1.52*	*1.40*	*1.34*	*1.46*
Utilities	*54*	*1.85*	*1.29*	*2.40*	*1.84*	*1.39*	*2.28*
Construction	*24*	*0.60*	*0.10*	*1.10*	*0.67*	*0.15*	*1.19*
Tertiary sector	75	1.42	1.13	1.71	1.25	0.90	1.61
Trade, cater., and accom.m.	*23*	*0.97*	*0.65*	*1.28*	*0.84*	*0.53*	*1.16*
Transport, stor., and comm.	*16*	*1.92*	*0.75*	*3.09*	*2.10*	*0.71*	*3.50*
Finance, ins., real est., bus.	*8*	*1.07*	*0.43*	*1.72*	*0.57*	*0.03*	*1.10*
Services	*21*	*1.63*	*1.35*	*1.92*	*1.47*	*1.19*	*1.76*
Developing countries	856	1.83	1.69	1.96	1.54	1.43	1.66
Developed countries	738	1.24	1.16	1.32	1.24	1.15	1.34
Publication status							
Published papers	1,385	1.23	1.13	1.32	1.65	1.52	1.78
Unpublished papers	2,139	1.60	1.53	1.68	1.61	1.53	1.68
All estimates	3,524	1.45	1.40	1.51	1.64	1.56	1.71

Notes: The definitions of subsets are available in Table 2.4. Weighted = estimates weighted by the inverse of the number of estimates reported per study. Several elasticities in our dataset are estimated for all industries or across more sectors; these observations are excluded from the table.

Table 2.2 provides a first indication of the potential causes of heterogeneity. We compute the mean values of the Armington elasticity estimates for different groups of data based on temporal dynamics (short- or long-run), data frequency, structural variation, and publication characteristics. To account for the unbalancedness of our dataset, we also compute mean estimates weighted by the inverse of the number of estimates reported per study so that each study gets the same weight. The table shows that the long-run elasticities are approximately twice as large as the short-run elasticities, which corroborates the arguments of Gallaway *et al.* (2003) and the common notion that short-run elasticities are smaller. In fact, Cassoni & Flores (2008) argue that smaller short-run estimates are given by the estimation design itself, unless overshooting occurs. Quarterly and annual data are typically used to capture the long-run effects (Gallaway *et al.* 2003) and thus can be expected to produce larger elasticities than monthly data, which is supported by the statistics shown in the table.

The smaller elasticities reported for the primary sector (with respect to other sectors) suggest that the products of agriculture, forestry, fishing, mining, and quarrying are more difficult to substitute with their foreign alternatives. Concerning agriculture, this finding can be explained, as Kuiper & van Tongeren (2006) point out, by a common, explicit or implicit, support of domestic (or even local) produce. In contrast, the largest elasticities are typically found for utilities (approximately 1.85) and transport, storage, and communication (1.92). The elasticity also tends to be 50% larger for developing countries than for developed countries. Finally, although the means suggest a difference between the typical results of published and unpublished papers, the weighted means, in which each study has the same weight, suggest that the publication process is not associated with the magnitude of the estimates of the Armington elasticity. This simple analysis suggests there is potential for systematic differences among the reported elasticities, but any particular conclusion can be misleading without accounting for the correlation between individual aspects of data and methodology, which we address in Section 2.4. It can also be misleading without correcting for publication bias, and we turn to this problem in the following section.

2.3 Testing for Publication Bias

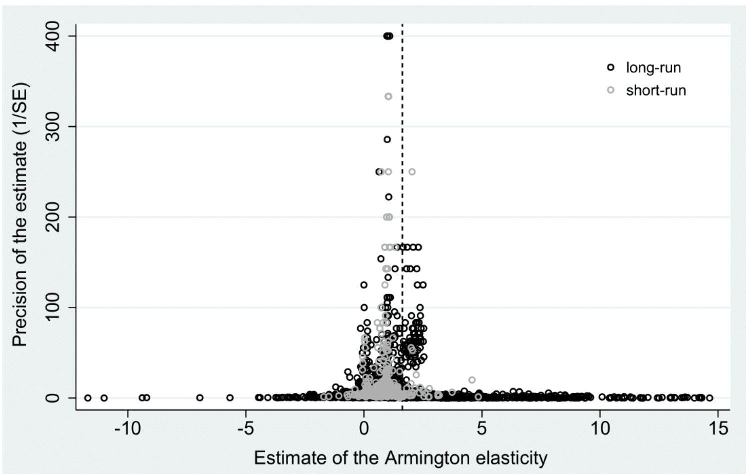
Publication bias is widespread in science, and economics is no exception: Ioannidis *et al.* (2017b) document that the typical estimate reported in economics is exaggerated twofold because of publication selection. Publication selection arises because of the general preference of authors, editors, and referees for estimates that have the "right" sign and are statistically significant. Of course, this is not to say that publication selection equals cheating: in contrast, it makes sense for (and improves the value of) an individual study not to focus on estimates that are evidently wrong. But when most authors follow the strategy of ignoring estimates that have the "wrong" sign or are statistically insignificant, our inference from the literature as a whole (and also from many individual studies) becomes distorted. Given the degrees of freedom available to researchers in economics, estimates with the "right" sign and statistical significance at the 5% level are almost always possible to obtain after a sufficiently large number of specifications have been tried. A useful analogy provided by McCloskey & Ziliak (2019) is the Lombard effect, in which speakers increase their vocal effort in the presence of noise: given noisy data or estimation techniques, the researcher has more incentives to search through more specifications for a significant effect. When statistical significance becomes the implicit requirement for publication, significance will be produced but will no longer reflect what the statistical theory expects of it.

A conspicuous feature of the Armington elasticity is that it must be positive if both domestic and foreign goods are useful to the consumer. Therefore, from the very beginning, the literature has shunned negative and zero estimates as clear artifacts of data or method problems. One of the first studies, Alaouze (1977, p. 8), notes, "we shall concentrate on the ...[*industries*]... for which the elasticity of substitution has the correct [*positive*] sign." Among the latest studies, Feenstra *et al.* (2018, p. 144) find that the estimated elasticity is negative for some varieties and isolate them from the dataset: "these data are faulty or incompatible with our model." As we have noted, this approach can improve the inference drawn from an individual study but generally creates a bias. Given the inherent noise in trade data, estimated elasticities for some industries or specifications will always be insignificant, negative, or both. For other industries or specifications, the same noise produces estimates that are much larger than the true effect. However, no upper bound exists that would immediately deem elasticities implausible; some domestic and foreign goods can be perfectly substitutable in theory. Therefore, the large estimates will be kept in the paper and interpreted. This psychological asymmetry between zero and infinity coupled with inevitable imprecision in data and estimation creates publication bias. One apparent solution is symmetrical trimming: when the authors ignore 10 negative or insignificant estimates, they should also ignore the 10 largest positive estimates. Winsorizing would be better still, but it is rarely employed in practice.

A common tool used to assess the extent of publication bias is the so-called funnel plot (Egger *et al.* 1997b). The funnel plot shows the magnitude of the estimated effect on the horizontal axis and the precision of the estimate (the inverse of the standard error) on the vertical axis. There should be no relation between these two quantities because virtually all techniques used by the researchers to estimate the Armington elasticity guarantee that the ratio of the estimate to its standard error has a symmetrical distribution (typically a t-distribution). Therefore, regardless of their magnitude and precision, the estimates should be symmetrically distributed around the true mean effect. With decreasing precision, the estimates become more dispersed around the true effect and thus form a symmetrical inverted funnel. In the presence of publication bias, the funnel becomes either hollow (because insignificant estimates are omitted), asymmetrical (because estimates of a certain sign or size are excluded), or both.

The funnel plot in Figure 2.5 gives us a mixed message, as we show short- and long-run estimates of the Armington elasticity separately. The short-run elasticities are symmetrically distributed around their most precise estimates, which are slightly less than 1. The long-run elasticities, in contrast, form an asymmetrical funnel: the most precise estimates are also close to 1, but among

Figure 2.5: Funnel plot suggests publication bias among long-run elasticities



Notes: In the absence of publication bias, the funnel should be symmetrical around the most precise estimates of the elasticity. The dashed vertical line denotes the simple mean of the full sample of elasticities. Outliers are excluded from the figure for ease of exposition but included in all statistical tests.

imprecise estimates, there are many more that are much larger than 1 compared to those that are smaller than 1. This finding is consistent with no publication selection among short-run elasticities and publication selection against negative and insignificant elasticities among long-run elasticities. Nevertheless, the funnel plot is only a simple visual test, and the dispersion of the long-run estimates could suggest heterogeneity in data and methods, the other systematic factor driving the estimated coefficients. Regression-based funnel asymmetry tests provide a more concrete way to test for publication bias. As we have noted, if publication selection is present, the reported estimates and standard errors are correlated (Stanley 2005; Stanley & Doucouliagos 2010; Havranek 2015):

$$\sigma_{ij} = \sigma_0 + \delta \cdot SE(\sigma_{ij}) + \mu_{ij}, \tag{2.4}$$

where σ_{ij} denotes i-th estimate of the Armington elasticity with the standard error $SE(\sigma_{ij})$ estimated in the j-th study; μ_{ij} is the error term. σ_0 is the mean underlying effect beyond publication bias (that is, conditional on maximum precision), and the coefficient δ of the standard error $SE(\sigma_{ij})$ represents the strength of publication bias. If $\delta = 0$, no publication bias is present. If $\delta \neq 0$, σ's and their standard errors are correlated, the correlation can arise either because researchers discard negative estimates of the elasticity (in which case the correlation occurs due to the apparent heteroskedasticity) or because researchers compensate for large standard errors with large estimates of the elasticity (the Lombard effect).

Table 2.3 presents the results of (2.4) using various estimation techniques run for three samples: the pooled set of elasticities, short-run elasticities, and long-run elasticities. Panel A uses unweighted data. In the baseline OLS model, the coefficient δ from (2.4) is not statistically significant for the short-run sample, and the estimated corrected mean is the same as the simple mean of 0.9. In the sample of long-run elasticities, in contrast, we find strong publication bias that decreases the underlying mean from 1.56 (the uncorrected mean) to 0.9 (the mean corrected for publication bias). The result for a pooled sample of short- and long-run elasticities is close to that of long-run elasticities because long-run elasticities dominate the dataset.

In the next model, we add study-level fixed effects to the baseline specification, which slightly deepens the difference between the mean short- and long-run effects beyond bias. Finally, for Panel A, we use a multilevel estimation technique that implements partial pooling at the study level and uses the data to influence the pooling weights. Given that the estimated elasticities

Table 2.3: All tests indicate publication bias among long-run Armington elasticities

	All	Short-run	Long-run
PANEL A: Unweighted estimations			
OLS			
SE (publication bias)	0.808***	0.0791	0.805***
	(0.0652)	(0.0826)	(0.0630)
Constant (effect beyond bias)	0.873***	0.867***	0.901***
	(0.133)	(0.0249)	(0.168)
Fixed effects			
SE (publication bias)	0.621***	-0.00578	0.627***
	(0.0588)	(0.104)	(0.0580)
Constant (effect beyond bias)	1.007***	0.883***	1.047***
	(0.0423)	(0.0192)	(0.0476)
Hierarchical Bayes			
SE (publication bias)	0.500**	-0.0810	0.630***
	(0.190)	(0.480)	(0.190)
Constant (effect beyond bias)	1.200***	0.887**	1.250***
	(0.240)	(0.310)	(0.0476)
PANEL B: Weighted OLS estimations			
Weighted by the inverse of the number of estimates reported per study			
SE (publication bias)	1.017***	0.0975*	1.033***
	(0.249)	(0.0514)	(0.251)
Constant (effect beyond bias)	1.011***	0.893***	1.046***
	(0.254)	(0.0694)	(0.303)
Weighted by the the inverse of the standard error			
SE (publication bias)	1.559	2.698	0.906**
	(0.969)	(2.213)	(0.431)
Constant (effect beyond bias)	0.761***	0.510	0.922***
	(0.217)	(0.325)	(0.205)
PANEL C: Non-linear estimations			
Weighted average of adequately powered (Ioannidis *et al.* 2017b)			
Effect beyond bias	1.049***	0.872***	1.101***
	(0.017)	(0.024)	(0.021)
Selection model (Andrews & Kasy 2019)			
Effect beyond bias	0.911***	0.863***	0.943***
	(0.015)	(0.018)	(0.021)
Stem-based method (Furukawa 2019)			
Effect beyond bias	0.992***	1.031***	0.994***
	(0.024)	(0.070)	(0.042)
Observations	3,524	556	2,968

Notes: The uncorrected mean of the estimates of the long-run Armington elasticity is 1.56. Panels A and B report the results of regression $\sigma_{ij} = \sigma_0 + \delta \cdot SE(\sigma_{ij}) + \mu_{ij}$, where σ_{ij} denotes i-th Armington elasticity estimated in the j-th study and $SE(\sigma_{ij})$ denotes the corresponding standard error. All = the entire dataset, Short-run = short-run Armington elasticities, Long-run = long-run Armington elasticities, SE = standard error. Standard errors, clustered at the study and country level, are reported in parentheses (except Hierarchical Bayes, which has posterior standard deviation in parentheses). The available number of observations is reduced for Ioannidis *et al.* (2017b)'s estimation (all 3,440; short-run 555; long-run 2,885) and Furukawa (2019)'s estimation (all 1,850; short-run 105; long-run 965). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Stars for hierarchical Bayes are presented only as an indication of the parameter's statistical importance to keep visual consistency with the rest of the table.

are nested within each study, hierarchical modeling is a convenient choice to analyze the variance in the elasticities: one can expect that the stochastic term of (2.4) depends on the design of each individual study and therefore does not have the same dispersion across individual studies. It follows that the regression coefficients δ are probably not the same across studies. Nevertheless, δ's should be related, and the hierarchical modeling treats them as random variables of yet another linear regression at the study level. We apply a hierarchical Bayes model and implement the Gibbs sampler for hierarchical linear models with a standard prior, following Rossi *et al.* (2005). The hierarchical model corroborates the evidence presented earlier but finds slightly weaker publication bias among the estimates of the long-run elasticity.

Panel B of Table 2.3 presents weighted alternatives to the baseline OLS model of Panel A. First, the regression is weighted by the inverse of the number of estimates reported by each study, so that both small and large studies are all assigned the same importance. Second, the regression is weighted by the inverse of the standard error so that more precise estimates are assigned greater importance. Panel B shows results that support the conclusions from Panel A. Finally, Panel C shows the latest alternatives to linear meta-analysis models. The problem with the linear regression that we have used so far is the implicit assumption that publication bias is a linear function of the standard error. If the assumption does not hold, our conclusion concerning publication bias can be misleading. Here, we apply three non-linear techniques that relax this assumption. The corrected means of both the short- and long-run Armington elasticity remain close to unity in all three alternative approaches: the weighted average of adequately powered estimates by Ioannidis *et al.* (2017b), the stem-based method by Furukawa (2019), and the selection model by Andrews & Kasy (2019).

Based on a survey involving more than 60,000 estimates, Ioannidis *et al.* (2017b) document that the median statistical power among the published results in economics is 18%. They show how low power is associated with publication bias and then propose a simple correction procedure that focuses on the estimates with power above 80%. Monte Carlo simulations presented in Ioannidis *et al.* (2017b) suggest that this simple technique outperforms the commonly used meta-regression estimators. The intuition of the model presented by Furukawa (2019) rests on the fact that the most precise estimates suffer from little bias: with very small standard errors, the authors can easily produce estimates that are statistically significant. While previous authors have recommended meta-analysts to focus on a fraction of the most precise estimates in meta-analysis (for example, Stanley & Doucouliagos 2010), Furukawa (2019) finds a clever way to estimate this fraction based on exploiting the trade-off between bias and variance (omitting studies increases variance). Andrews & Kasy (2019) use the observation reported by many researchers (for instance, Havranek 2015; Brodeur *et al.* 2016) that standard cut-offs for the p-value (0.01, 0.05, 0.1) are associated with jumps in the distribution of reported estimates. Andrews & Kasy (2019) build on Hedges (1992) and construct a selection model that estimates publication probability for each estimate in the literature given its p-value. They show that, in several areas, the technique gives results similar to those of a large-scale replication.

Several important findings can be distilled from the estimations reported in Table 2.3. First, we find publication bias among long-run elasticities but not among short-run elasticities. One explanation consistent with this result is that short-run elasticities are typically deemed less important than long-run elasticities, especially for policy purposes. They are often reported only as complements to the central findings of the paper. It can take time before consumers shift their demand between domestic and foreign goods; consequently, insignificant estimates of the short-run elasticity are more likely to survive the publication process than insignificant estimates of the long-run elasticity. Second, publication bias inflates the mean estimate of the long-run Armington elasticity by at least 50%, which can have a strong impact on the results of a model informed by the empirical literature in terms of the calibration of the elasticity. Third, the large difference between the short- and long-run elasticities reported in Table 2.2 (and observed in many studies, see Gallaway *et al.* 2003) is all but erased once publication bias is taken into account. In sum, we find robust evidence of publication bias in this literature. However, some of the apparent correlations between the estimated elasticities and their standard errors can be due to data and method heterogeneity. We turn to this issue in the next section.

2.4 Why Elasticities Vary

2.4.1 Potential Factors Explaining Heterogeneity

Three reasons for the systematic differences in the estimates of the Armington elasticity have been frequently discussed in the literature. First, studies using disaggregated data are often observed to yield larger estimates than studies using aggregate data (Imbs & Mejean 2015). Second, cross-sectional studies tend to yield larger estimates than time-series studies (Hillberry & Hummels 2013). Third, multi-equation estimation techniques typically give larger estimates than single-equation techniques (Goldstein & Khan 1985). Many literature reviews (including Cassoni & Flores 2008; Marquez 2002; McDaniel & Balistreri 2003), moreover, stress other characteristics of estimates and studies that can significantly influence the results. We present the first attempt to shed light on the sources of heterogeneity in Table 2.2. To investigate the heterogeneity among the estimates of the Armington elasticity more systematically, we codify 43 characteristics of the study design and augment equation (2.4) by adding these characteristics as explanatory variables. Given that publication bias affects only the long-run elasticity, we replace the standard error in the equation by an interaction term between the standard error and a dummy variable that equals one if the estimate corresponds to a long-run elasticity.

Table 2.4 lists all the codified variables, their definitions and summary statistics, including the simple mean, standard deviation, and mean weighted by the inverse of the number of observations reported in a study. For ease of exposition, we divide the variables into groups reflecting data characteristics (11 aspects), structural variations (11 aspects), estimation techniques (14 aspects), and publication characteristics potentially related to quality that are not captured by data and estimation characteristics (3 aspects). The distinction between short- and long-run elasticities is among the most important factors stressed in the literature (Gallaway *et al.* 2003). Nevertheless, in the previous section, we find that publication bias plagues the estimates of long-run elasticities and that beyond publication bias, short- and long-run elasticities have comparable magnitudes. In this section, we will examine whether the claim still holds when other possible systematic influences on the estimates of the Armington elasticity are taken into account.

Table 2.4: Description and summary statistics of the regression variables

Variable	Description	Mean	SD	WM
Armington elasticity	The reported estimate of the Armington elasticity.	1.45	1.78	1.64
Standard error (SE)	The reported standard error of the Armington elasticity estimate.	0.72	1.18	0.61

Continued on next page

Table 2.4: Description and summary statistics of the regression variables (continued)

Variable	Description	Mean	SD	WM
SE * Long-run effect	The interaction between the standard error and the estimated long-run Armington elasticity.	0.69	1.19	0.59
Temporal dynamics				
Short-run effect	=1 if the estimated Armington elasticity is short-term (reference category for the group of dummy variables describing temporal dynamics).	0.16	0.36	0.12
Long-run effect	=1 if the estimated Armington elasticity is long-term.	0.84	0.36	0.88
Data characteristics				
Data disaggregation	The level of data aggregation according to SIC classification (min = 1 if fully aggregated, max = 8 if disaggregated).	6.49	1.58	6.20
Results disaggregation	The level of results aggregation according to SIC classification (min = 1 if fully aggregated, max = 8 if disaggregated).	5.06	1.21	5.34
Monthly data	=1 if the data are in monthly frequency.	0.14	0.35	0.08
Quarterly data	=1 if the data are in quarterly frequency (reference category for the group of dummy variables describing data frequency).	0.21	0.41	0.25
Annual data	=1 if the data are in yearly frequency.	0.65	0.48	0.67
Panel data	=1 if panel data are used (reference category for the group of dummy variables describing time and cross-sectional dimension of data).	0.34	0.47	0.27
Time series	=1 if time-series data are used.	0.58	0.49	0.65
Cross-section	=1 if cross-sectional data are used.	0.08	0.27	0.08
Data period	The length of the time period in years.	14.24	9.76	17.08
Data size	The logarithm of the total number of observations used to estimate the elasticity.	4.64	1.93	4.55
Midyear	The median year of the time period of the data used to estimate the elasticity.	23.45	11.54	22.48
Structural variation				
Primary sector	=1 if the estimate is for the primary sector (agriculture and raw materials; reference category for the group of dummy variables describing sectors).	0.10	0.31	0.31
Secondary sector	=1 if the estimate is for the secondary sector (manufacturing).	0.86	0.34	0.58
Tertiary sector	=1 if the estimate is for tertiary sector (services).	0.02	0.14	0.03
Developing countries	=1 if the estimate is for a developing country (reference category for the group of dummy variables describing the level of development).	0.24	0.43	0.28
Developed countries	=1 if the estimate is for a developed country.	0.79	0.41	0.74

Continued on next page

Table 2.4: Description and summary statistics of the regression variables (continued)

Variable	Description	Mean	SD	WM
Market size	The logarithm of the market size of the home country (GDP in billions of USD, 2015 prices).	6.45	1.86	5.94
Tariffs	The tariff rate of the home country (weighted mean, all products, %).	6.78	7.15	6.07
Non-tariff barriers	Additional cost to import of the home country (USD per container).	0.94	0.26	0.97
FX volatility	The volatility of the exchange rate using the DEC alternative conversion factor (home country currency unit per USD).	0.58	0.55	0.69
National pride	Home bias captured by the percentage of "I am very proud of my country" answers from the World Values Survey.	0.53	0.22	0.51
Internet usage	The number of fixed broadband subscriptions of the home country (per 100 people).	2.91	5.02	1.23
Estimation technique				
Static model	=1 if a static model is used for estimation.	0.23	0.42	0.30
Distributed lag and trend model	=1 if a distributed lag or trend model is used.	0.10	0.30	0.27
Partial adjustment model	=1 if a partial adjustment model is used for estimation.	0.15	0.35	0.11
First-difference model	=1 if a first-difference model is used.	0.09	0.29	0.05
Error-correction model	=1 if an error-correction model is used.	0.04	0.20	0.04
Nonlinear model	=1 if a nonlinear model is used.	0.28	0.45	0.13
Other models	=1 if another model is used (reference category for the group of dummy variables describing models used).	0.11	0.31	0.10
OLS	=1 if the OLS or GLS estimation method is used.	0.48	0.50	0.67
CORC	=1 if the Cochrane-Orcutt or FGLS estimation method is used.	0.16	0.37	0.13
TSLS	=1 if the instrumental method is used.	0.09	0.28	0.06
GMM	=1 if the GMM estimation method is used.	0.24	0.43	0.10
Other methods	=1 if other types of estimation are used (reference category for the group of dummy variables describing the estimation method used).	0.03	0.17	0.05
Import constraint	=1 if the study includes some measure of import restriction.	0.03	0.18	0.06
Seasonality	=1 if the study controls for seasonality.	0.20	0.40	0.12
Publication characteristics				
Impact factor	The recursive discounted impact factor from RePEc.	0.12	0.24	0.17

Continued on next page

Table 2.4: Description and summary statistics of the regression variables (continued)

Variable	Description	Mean	SD	WM
Citations	The logarithm of the number of Google Scholar citations normalized by the number of years since the first draft of the paper appeared in Google Scholar.	1.26	1.01	1.00
Published	=1 if a study is published in a peer-reviewed journal.	0.39	0.49	0.65

Notes: SD = standard deviation, WM = mean weighted by the inverse of the number of estimates reported per study, SIC = Standard Industrial Classification system for classifying industries by a four-digit code. Market size, tariff and non-tariff barriers, FX volatility, and internet usage have been collected from the World Bank database (WB 2018), data on national pride from the World Values Survey (Inglehart *et al.* 2014). The impact factor is downloaded from RePEc and the number of citations from Google Scholar. The rest of the variables are collected from studies estimating the Armington elasticity.

Data characteristics. Many studies (Feenstra *et al.* 2018; McDaniel & Balistreri 2003; Welsch 2008, among others) argue that because intra-industry diversity decreases with an increasing level of sectoral aggregation, more aggregated data should yield smaller elasticities. Feenstra *et al.* (2018) note that some recent macro-studies (Bergin 2006; Heathcote & Perri 2002) estimate the aggregate elasticities around unity, while studies focusing on individual product groups (Broda & Weinstein 2006; Imbs & Mejean 2015) imply much stronger responses. McDaniel & Balistreri (2003) compare two articles on US data that use 3-digit SIC level (Reinert & Roland-Holst 1992) and 4-digit SIC level (Gallaway *et al.* 2003) aggregations and come to the same conclusion: higher disaggregation brings higher substitutability. We codify the *data disaggregation* variable according to the SIC classification. Fully aggregated, whole-economy data acquire the value of 1; in contrast, fully disaggregated product-level data acquire the value of 8. Given the consensus in the literature, we expect the variable to show a positive association with the reported elasticities. Furthermore, in some papers (such as Aspalter 2016; Mohler & Seitz 2012), the level of aggregation of the input data differs from the level of aggregation of the reported results. Imbs & Mejean (2015) argue that a pooled estimate that ignores heterogeneity across sectors tends to be biased downwards. To reflect the problem of aggregating the results, we create an additional variable based on the same principles as the variable for data aggregation.

Another commonly discussed issue is data frequency. It is related to the short- or long-run nature of the elasticity, but we control for this feature separately. Cassoni & Flores (2008) show that aggregation from monthly to quarterly data removes short-term adjustment patterns, such as overshooting (Cassoni & Flores 2010) or J-curve effects (Backus *et al.* 1994). They also note that *monthly data* often contain atypical observations that could misrepresent the underlying trade data. Gallaway *et al.* (2003), on the other hand, estimate long-run elasticities based on monthly and quarterly data and find no systematic difference in the estimates. Given that quantity measures are notoriously noisy, Hillberry & Hummels (2013) state that the measurement error often becomes exacerbated with monthly or quarterly data and high product disaggregation. The use of *quarterly* instead of *yearly data* may be necessary to gain a sufficiently large dataset, but Hertel *et al.* (1997) argue that problems associated with quarterly data could lead to overly inelastic estimates. A number of studies, including Aspalter (2016), Olekseyuk & Schurenberg-Frosch (2016), and Feenstra *et al.* (2018), use annual data, especially when the authors want to identify both micro and macro elasticities. Aspalter (2016) also suggests that the annual frequency of data often leads to a more consistent cross-country dataset.

We further distinguish among *time series*, *cross-section*, and *panel data*, using panel data as the reference category. The survey by McDaniel & Balistreri (2003) reports that cross-sectional data are associated with larger reported elasticities because cross-sectional estimates also consider supply conditions. Cassoni & Flores (2008), however, argue that the conclusion of McDaniel & Balistreri (2003) stems from comparing results based on heterogeneous analyses and data and point out that the impact of data cross-sectionality depends on the correct specification of the model and the estimation technique employed. The variable *data period* reflects how estimates differ when obtained over longer time periods, while the variable *data size* captures the potential effects of small-sample bias. We also control for the age of the data by including a variable that reflects the midpoint year of the sample (variable *midyear*) with which the Armington elasticity is estimated. Figure 2.2 suggests that the elasticity is increasing in time (and some studies, for example Schurenberg-Frosch 2015; Welsch 2008, observe a similar pattern). In this vein, Hubler & Pothen (2017) argue that globalization might have increased the Armington elasticity by decreasing the heterogeneity of products and reducing the market power of individual countries.

Structural variation. The elasticity of substitution might depend systematically on the characteristics of the product, industry, or country in question.

Blonigen & Wilson (1999) suggest that with greater physical differences, the elasticity of substitution between products decreases. Shiells *et al.* (1986) and more recent papers such as Faria & Haddad (2014), Nemeth *et al.* (2011), and Saikkonen (2015) provide evidence of how the Armington elasticity differs across industries. Moreover, Saito (2004) shows that heterogeneous goods (e.g., final products such as automobiles or medical equipment) are more difficult to substitute across countries than more homogenous goods (e.g., intermediate products such as glass or metals). Because we do not have enough variation in our dataset to control for the many individual product categories or industries (if all these controls were included, collinearity would skyrocket), we control for sectoral differences by dividing the sample into three groups: the *primary sector* with industries related to raw materials, the *secondary sector* with manufacturing industries, and the *tertiary sector* of services. Nevertheless, the data that we provide in the online appendix include more details and researchers can use these data and codes to construct implied elasticities for the individual industries in which they are interested.

We also control for the characteristics of the country for which the elasticity is estimated (the home country). Developing countries can be expected to face a larger pool of substitutable products abroad because the rest of the world encompasses the production of all levels of technology. In contrast, for developed countries with better production technologies, it might be more difficult to find adequate substitutes abroad. Moreover, Kapuscinski & Warr (1999) note that developing countries often provide poor data, and the resulting biases could lead to larger elasticities. We divide the countries into two categories: a group of *developed* countries, which includes Central and Western Europe, North America, Australia, New Zealand, and Japan; and a group of *developing countries*, which covers the rest of Asia, Latin America, and Africa.

It has been shown in the literature that even physically identical goods can be differentiated by aspects such as availability, customer service, and perception of quality. Linder (1961) suggests that countries with similar income per capita should trade more because their consumers have similar tastes, as reflected in the production of goods in each country (more details are provided in Francois & Kaplan 1996). Ideally, to capture these features of consumers' preferences, we follow the study on the border effect by Havranek & Irsova (2017) and create a variable representing the income dissimilarity of the home country and the corresponding foreign country. Because this bilateral approach is not feasible for the Armington elasticity literature, we use another representation of consumer preferences: we include a proxy variable *national pride* to capture consumer bias for home goods over foreign ones (Trefler 1995; Kehoe *et al.* 2017). The variable is constructed as the percentage of 'very proud' answers to the question 'How proud are you of your country?' from the World Values Survey (Inglehart *et al.* 2014). Wolf (2000), for example, shows that the home bias could go beyond the influence of typical quantifiable trade barriers and also exist on a sub-national level.

Several potential country-level determinants of the Armington elasticity have a strong connection to the border effect first presented by McCallum (1995). One of the common border effect determinants is *market size*: any border barrier in a small economy increases the ratio of within-country trade more than in a large economy. We thus expect this variable to have a positive association with the reported elasticity. To proxy for market size, we use GDP for the midpoint of the data period used in the study. Moreover, trade barriers and other extra transaction costs associated with crossing the border have also been considered an important determinant of the Armington macro elasticity (Lopez & Pagoulatos 2002). These trade frictions are captured by variables *tariff* (representing the tariff rate) and *non-tariff barriers* (representing the cost to import); all these data are obtained from WB (2018).

According to Parsley & Wei (2001), contracting costs and insecurity represent other potential determinants that affect cross-country trade and possibly the Armington elasticity. We approximate these additional trade frictions by the volatility of the exchange rate in the home country versus the US dollar (variable *FX volatility*). Parsley & Wei (2001) suggest that the exchange rate volatility may not only contribute to cross-border market insecurities but also explain the price dispersion of similar goods across the border. Finally, we account for information barriers and use the number of broadband subscriptions per 100 people as a measure of *internet usage*. The expansion of internet use creates new types of tradable services and is believed to have increased cross-border trade (IBRD 2009).

Estimation technique. A large variety of models and methods exist to estimate the Armington elasticity. To simplify, denoting the expression $\log(Q_M/Q_D)$ in (2.3) as y , $\log(P_D/P_M)$ as x, and $\log(\beta/1-\beta)$ as σ_0, we obtain the *static model* $y_t = \sigma_0 + \sigma_1 x_t + e_t$, where σ_1 is the Armington elasticity and e is the error term. Static models constitute approximately 23% of our dataset. Another category labeled *distributed lag and trend model* includes elasticities estimated using distributed lag models (Tourinho *et al.* 2003) and models with a time trend variable added to achieve data stationarity (Lundmark & Shahrammehr 2012): $y_t = \sigma_0 + \sum_{l=0}^{\tau} \sigma_{l+1} x_{t-l} + \sigma_{\tau+1} t + e_t$, $\tau \geq 0$. The *partial adjustment model*, on the other hand, allows for a non-instantaneous adjustment of the demand structure to the changes of the relative prices (for example Ogundeji *et al.* 2010) by adding the lagged dependent variable y_{t-1} among the explanatory variables and reads $y_t = \sigma_0 + \sigma_1 x_t + \sigma_2 y_{t-1} + e_t$ (Alaouze 1977, shows that the omission of the lagged dependent variable in cases where it is significant biases the estimates downwards).

If the corresponding levels of time series are not stationary or cointegrated, the authors take *first differences* (see Gibson 2003, for example). In some cases, the lagged value of the level of the explanatory variable is also included, and the authors end up with $\Delta y_t = \sigma_0 + \sigma_1 \Delta x_t + \sigma_2 x_{t-1} + e_t$. When the time series are cointegrated, authors also use an *error-correction model* to estimate the elasticity (such as Gan 2006, does); then, the model reads $\Delta y_t = \sigma_0 + \sigma_1 \Delta x_t + \sigma_2 y_{t-1} + \sigma_3 x_{t-1} + e_t$. Several studies, including Corado & de Melo (1983), Feenstra *et al.* (2018), and Saikkonen (2015), employ different forms of *non-linear models*. The non-linear model category constitutes 28% of our dataset. There is no unifying specification presentable in this case, as the individual approaches differ. The reference category for the group of dummy variables describing the models used to estimate the Armington elasticity is the variable *other models*, which covers the rest of the used approaches that do not fall under any of the above-mentioned categories.

Shiells & Reinert (1993) use the GLS technique, ML estimation, and simultaneous equation estimator that employs a distributed lag model to estimate the elasticities. They find the estimates to be relatively insensitive to the three alternative estimation procedures. Not all studies, however, come to the same conclusion of methodological indifference. To account for the potential effect of estimation techniques, we group the most frequently used methods of estimation into five categories: OLS estimation together with the GLS estimator (variable *OLS*), Cochrane-Orcutt estimation together with the FGLS (variable *CORC*), two-stage least squares and related techniques (variable *TSLS*), a separate group of *GMM* estimates, and all *other methods*, which represent the reference category for this group of dummies. We also include a control that equals one if the specification includes some measures of *import constraints*. Alaouze (1977) stresses that quantitative and tariff quota restrictions could bias the estimates of the elasticity because importers cannot fully utilize the advantages of price changes or must pay a fee when exceeding a certain amount of imported goods. Another coded aspect of the data is whether the authors control for *seasonality* in the demand function (Tourinho *et al.* 2010), which is a particularly important characteristic of agricultural products. Seasonality is commonly captured by quarterly dummies (see, for example, Ogundeji *et al.* 2010).

Publication characteristics. Despite the large number of variables we collect, the list of aspects potentially related to quality is unlimited. Therefore, we also employ several publication characteristics that can be expected to be correlated with the unobserved features of the quality of the paper. To see if published studies yield systematically different results, we include a dummy variable that equals one if the study is *published* in a peer-reviewed journal. To take into account the differences in the quality of publication outlets, we include the discounted recursive RePEc *impact factor* of the respective study (this impact factor is available for both journals and working paper series). Finally, for each study, we create a variable reflecting the logarithm of the number of Google Scholar *citations* normalized by the number of years since the first draft of the study appeared in Google Scholar.

2.4.2 Estimation

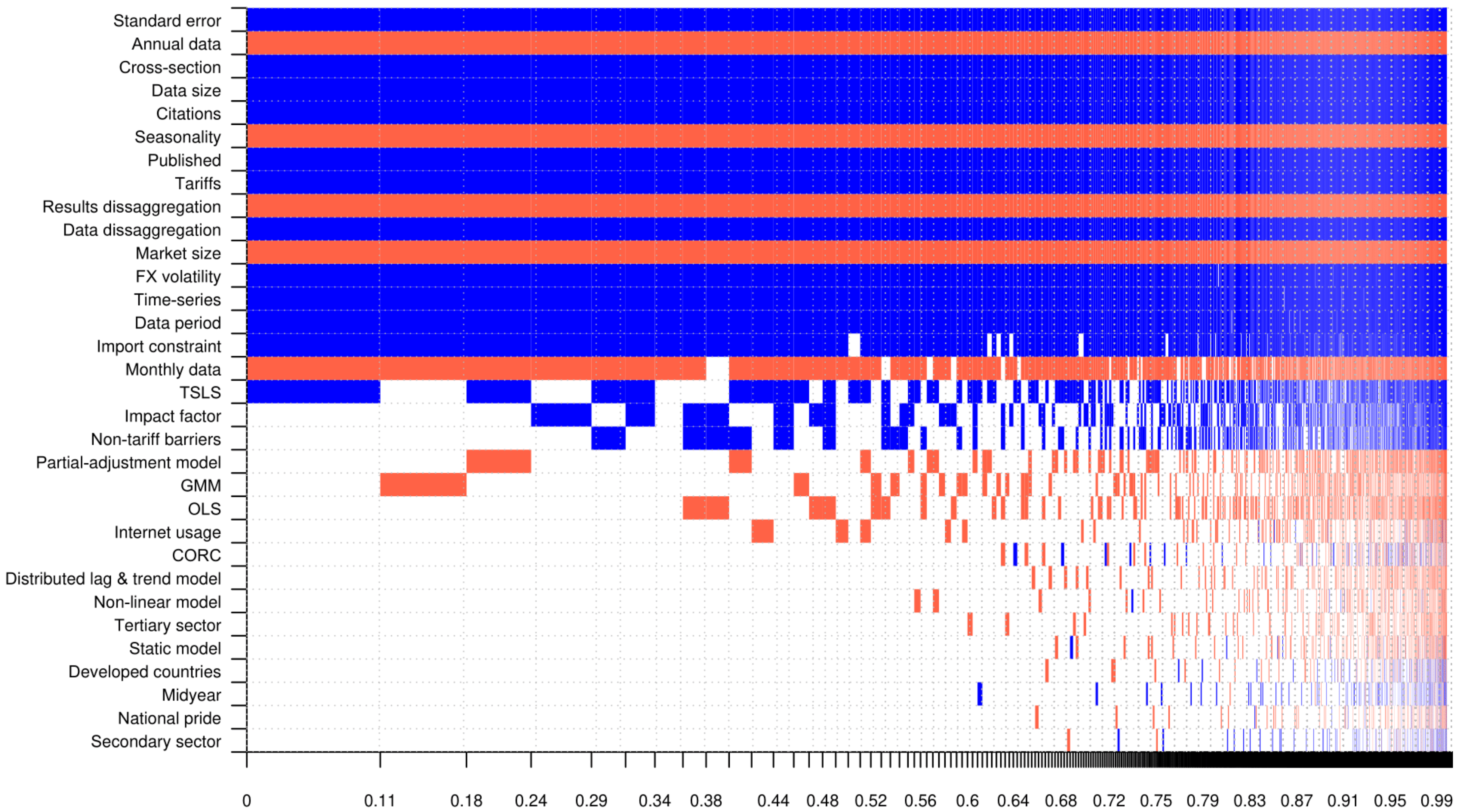
To relate the variables introduced above to the magnitude of the estimated Armington elasticities, one could run a standard regression with all the variables. But such an estimation would ignore the model uncertainty inherent in meta-analysis: while we have a strong rationale to include some of the variables, others are considered mainly as controls for which there is no theory on how they could affect the results of studies estimating the Armington elasticity. To address model uncertainty, we employ Bayesian model averaging (BMA). BMA runs many regressions with different subsets of the 2^{34} possible combinations of explanatory variables. We do not estimate all possible combinations but employ Monte Carlo Markov Chain (specifically, the Metropolis-Hastings algorithm of the `bms` package for R by Zeugner & Feldkircher 2015), which walks through the most likely models. In the Bayesian setting, the likelihood of each model is represented by the posterior model probability. The estimated BMA coefficients for each variable are represented by posterior means and are weighted across all models by their posterior probability. Each coefficient is then assigned a posterior inclusion probability that reflects the probability of the variable being included in the underlying model and is calculated as the sum of posterior model probabilities across all the models in which the variable is included. Further details on BMA can be found in, for example, Raftery *et al.* (1997) or Eicher *et al.* (2011). BMA has been used in meta-analysis, for example, by Havranek *et al.* (2015).

In the baseline specification, we employ the priors suggested by Eicher *et al.* (2011), who recommend using the uniform model prior (giving each model the same prior probability) and the unit information g-prior prior (giving the prior the same weight as one observation of the data). These priors reflect the lack of prior knowledge regarding the probability of individual specifications, model size, and parameter values. We use unweighted data to estimate the baseline but later provide weighted alternatives to evaluate the robustness of our results. Furthermore, as a robustness check, we follow Ley & Steel (2009) and apply the beta-binomial random model prior, which gives the same weight to each model size, as well as Fernandez *et al.* (2001), who advocate for the so-called BRIC g-prior. In addition, to avoid using priors entirely, we also apply frequentist model averaging (FMA). Following Hansen (2007), we use Mallow's criterion for model averaging and the approach of Amini & Parmeter (2012) towards the orthogonalization of the covariate space. Amini & Parmeter (2012) provide a comprehensive comparison of different averaging techniques, including Mallow's weights and other frequentist alternatives.

2.4.3 Results

Figure 2.6 visualizes the results of Bayesian model averaging. The columns of the figure denote the individual regression models, and the column widths indicate the posterior model probability. The columns are sorted by posterior model probability from left to right. The rows of the figure denote individual variables included in each model. The variables are ordered by their posterior inclusion probability from top to bottom in descending order. If a variable is excluded from the model, the corresponding cell is left blank. Otherwise, the blue color (darker in grayscale) indicates a positive sign of the variable's coefficient in the particular model; the red color (lighter in grayscale) indicates a negative sign. Figure 2.6 shows that approximately half of our variables are included in the best models, and the signs of these variables are robust across specifications.

Figure 2.6: Model inclusion in Bayesian model averaging



Notes: The figure depicts the results of Bayesian model averaging with a uniform model prior and unit information prior (Eicher *et al.* 2011). On the vertical axis, the explanatory variables are ranked according to their posterior inclusion probabilities from the highest at the top to the lowest at the bottom. The horizontal axis shows the values of cumulative posterior model probability for each model ranked from the highest on the left to the lowest on the right. All variables are described in Table 2.4. Numerical results are reported in Table 2.5. The blue color (darker in grayscale) means that the estimated parameter of the corresponding explanatory variable is positive. The red color (lighter in grayscale) indicates that the estimated parameter is negative. No color denotes that the corresponding explanatory variable is not included in the model. The results are based on the unweighted specification. The robustness checks in which the specification is weighted by the number of estimates reported per study and by the standard error of the estimate are provided in Table B2 in the Appendix A. Detailed diagnostics are provided in Table B1 and Figure B1.

Table 2.5: Why elasticities vary

Response variable:	Bayesian model averaging			Frequentist check (OLS)			Frequentist model averaging		
Estimate of the Armington elasticity	Post. mean	Post. SD	PIP	Coef.	SE	p-value	Coef.	SE	p-value
Constant	-1.38	NA	1.00	-1.53	0.70	0.03	-1.65	0.53	0.00
SE * Long-run effect	0.73	0.02	1.00	0.73	0.05	0.00	0.74	0.03	0.00
Long-run effect	0.00	0.02	0.03				0.17	0.12	0.13
Data characteristics									
Data disaggregation	0.19	0.04	1.00	0.18	0.10	0.07	0.20	0.05	0.00
Results disaggregation	-0.23	0.04	1.00	-0.22	0.10	0.03	-0.25	0.04	0.00
Monthly data	-0.59	0.13	1.00	-0.50	0.27	0.06	-0.48	0.16	0.00
Annual data	-1.16	0.12	1.00	-1.15	0.39	0.00	-0.89	0.16	0.00
Time series	0.54	0.13	1.00	0.59	0.46	0.20	0.83	0.17	0.00
Cross-section	2.21	0.19	1.00	2.34	0.40	0.00	2.09	0.22	0.00
Data period	0.03	0.00	1.00	0.03	0.01	0.00	0.03	0.01	0.00
Data size	0.32	0.02	1.00	0.32	0.09	0.00	0.33	0.02	0.00
Midyear	0.00	0.00	0.05				0.01	0.01	0.13
Structural variation									
Secondary sector	0.00	0.01	0.02				0.02	0.08	0.82
Tertiary sector	0.00	0.03	0.02				-0.07	0.17	0.68
Developed countries	0.00	0.02	0.02				0.22	0.15	0.15
Market size	-0.11	0.02	1.00	-0.12	0.06	0.05	-0.14	0.03	0.00
Tariffs	0.03	0.01	1.00	0.03	0.01	0.02	0.04	0.01	0.00
Non-tariff barriers	0.21	0.20	0.58	0.40	0.21	0.05	0.38	0.15	0.01
FX volatility	0.31	0.07	1.00	0.28	0.15	0.07	0.18	0.07	0.01
National pride	0.00	0.02	0.01				-0.17	0.19	0.37
Internet usage	0.00	0.00	0.07				0.01	0.01	0.46
Estimation technique									
Static model	0.00	0.01	0.02				-0.26	0.11	0.02
Distributed lag and trend model	0.00	0.01	0.01				-0.34	0.15	0.02
Partial adjustment model	-0.01	0.04	0.06				-0.29	0.12	0.02
First-difference model	0.00	0.02	0.03				-0.17	0.13	0.17
Nonlinear model	-0.01	0.07	0.05				-0.65	0.29	0.03
OLS	-0.14	0.18	0.47	-0.23	0.14	0.09	-0.39	0.18	0.03
CORC	-0.04	0.13	0.13				-0.25	0.18	0.15
TSLS	0.42	0.19	0.89	0.40	0.16	0.01	0.29	0.19	0.13
GMM	-0.03	0.11	0.09				-0.10	0.21	0.63
Import constraint	0.48	0.18	0.94	0.58	0.24	0.02	0.51	0.17	0.00
Seasonality	-0.57	0.11	1.00	-0.52	0.34	0.13	-0.37	0.13	0.00
Publication characteristics									
Impact factor	0.23	0.22	0.59	0.43	0.12	0.00	0.52	0.18	0.00
Citations	0.55	0.05	1.00	0.53	0.17	0.00	0.46	0.08	0.00
Published	0.57	0.09	1.00	0.61	0.30	0.04	0.58	0.15	0.00
Studies	42			42			42		
Observations	3,524			3,524			3,524		

Notes: SD = standard deviation. SE = standard error. PIP = posterior inclusion probability. Bayesian model averaging (BMA) employs the priors suggested by Eicher *et al.* (2011). The frequentist check (OLS) includes the variables recognized by BMA to comprise the best model and is estimated using standard errors clustered at the study and country levels. Frequentist model averaging (FMA) employs Mallow's weights (Hansen 2007) using the orthogonalization of the covariate space suggested by Amini & Parmeter (2012). All variables are described in Table 2.4. Additional details on the BMA exercise can be found in the Appendix A.

The numerical results of the BMA exercise with priors according to Eicher *et al.* (2011) are reported in Table 2.5. Additionally, we show two alternative estimations. First, we estimate simple OLS, which excludes the 13 variables that were deemed unimportant by the BMA exercise (according to Eicher *et al.* 2011, the effect of a variable is considered *decisive* if the posterior inclusion probability is between 0.99 and 1, *strong* between 0.95 and 0.99, *substantial* between 0.75 and 0.95, and *weak* between 0.5 and 0.75). OLS results mostly correspond with the results of BMA: the coefficients display the same signs and similar magnitudes, and their p-values typically correspond to the information extracted from the respective posterior inclusion probabilities (with the exception of country-level variables, which will be discussed below). Second, we estimate frequentist model averaging, which includes all variables used in the BMA model. FMA conclusions are mostly in line with the baseline; except that, unlike BMA, it considers estimation techniques to be important factors driving the magnitude of the Armington elasticity.

The complete set of robustness checks, including BMA exercises with alternative priors and weights, can be found in Table B2. When using alternative priors (according to Fernandez *et al.* 2001; Ley & Steel 2009), we obtain evidence that supports the conclusions of our baseline model. BMA weighted by the inverse of the number of estimates reported per study confronts our baseline model on estimation techniques, and we will discuss the differences later on. We also report BMA with precision weights, although such an estimation is problematic in our case because weighting by precision introduces artificial variation to the study-level variables. BMA results from Table 2.5 testify to the decisive importance of the effects caused by *data* and *results disaggregation*, the usage of *monthly* and *annual data*, *time-series* and *cross-section* type of input data, *data period* and *data size* of a study, the country's *market size*, imposed *tariffs*, *FX volatility*, a control for *seasonality*, the number of *citations*, and *published* studies. The results further point to substantial evidence of effects caused by imposed *import constraints* and weak evidence of effects caused by imposed *non-tariff barriers* and *nonlinear model* choice. We will concentrate on the variables for which we have the most robust evidence.

The presence of publication bias in the estimates of the long-run Armington elasticity is supported by evidence across all the models that we run. The reported long-run elasticities, therefore, are found to be systematically exaggerated due to publication bias even if we control for various data and method characteristics of the individual studies. The inclusion of these controls lowers the estimated magnitude of publication bias reported in Table 2.3, but only slightly (the coefficient decreases from 0.8 to approximately 0.7). At the same time, our results suggest that, after controlling for publication bias and other aspects of study design, the difference between the estimated short-run and *long-run elasticity* is, on average, close to zero.

Data characteristics. The evidence on the effect of *data disaggregation* is consistent with the prevalent opinion in the literature following mostly Hummels (1999): higher disaggregation of data leads to more homogenous products and brings higher international substitutability. The bias is thus believed to originate in the heterogeneity of goods included in aggregated categories. Our results suggest that the effect is statistically important; still, the economic importance of the effect seems relatively low (the coefficient equals 0.2 in Table 2.5) in comparison to other sources of heterogeneity. In the majority of the studies in our dataset, data disaggregation and *results disaggregation* have the same value, but some of the studies use disaggregated data while reporting aggregated elasticities. Imbs & Mejean (2015) show that if elasticities are heterogeneous, the aggregate elasticity of substitution is given by an adequately weighted average of good-specific elasticities. We find that, contrary to Imbs & Mejean (2015), the output data granularity (disaggregation of resulting elasticities) is negatively associated with the reported elasticities.

Data frequency is another systematic factor that influences the estimates of the Armington elasticity. Table 2.2 shows that elasticities estimated using datasets with annual and quarterly frequencies tend to be larger than when monthly data are employed for estimation. Hertel *et al.* (1997) states that, in general, with lower data frequencies, more inelastic estimates are to be expected, as adjustment patterns become lost in aggregation. When we control for publication bias and other aspects of study design, the elasticities estimated with *quarterly data* appear to be robustly higher—by approximately 1.5—than what any other data frequencies produce.

Our results also corroborate the importance of using *cross-sectional* data versus time-series data. When the time dimension of the data is accounted for, the estimated elasticities tend to be smaller by at least 1.5 (observation-weighted BMA in Table B2 puts the cross-sectional coefficient at 1.6 and the unweighted baseline at 2.2), although the length of the time series does not seem to play a substantial additional role. Studies with a small number of observations produce small estimates of the elasticity, which might reflect small-sample bias. Although some commentators in the literature note that the estimates of the Armington elasticity are increasing in time (Schurenberg-Frosch 2015; Welsch 2008; Hubler & Pothen 2017), we argue that once the study design is controlled for, no such pattern remains.

Structural variation. Given that the majority of studies deal with either the United States or Europe (and the economies of the United States, Germany, and France alone account for approximately 1,500 observations in our sample), our data sample suffers from a lack of cross-country variation, and the conclusions concerning the country-level variables should be taken with a grain of salt. Indeed, most of the country-level variables lose statistical significance in the frequentist check, where standard errors are clustered at the country level. With that disclaimer in mind, we briefly describe the results. Zhang & Verikios (2006) argues that small countries feature relatively low Armington elasticities because they are rather import-dependent and tend to boast highly specialized industries. The negative coefficient of variable *market size* across all models, albeit small, is not in line with this argument. Our results suggest that larger markets tend to have rather smaller Armington elasticities; some evidence from our weighted specification suggests that developed countries also feature smaller elasticities. Zhang & Verikios (2006), on the other hand, argue that developing countries have underdeveloped domestic industries that are often unable to compete with imports, which should contribute to smaller Armington elasticities.

Blonigen & Wilson (1999) find evidence that barriers to entry lower the elasticity of substitution between domestic and foreign goods. Our results indicate that barriers to trade, tariff or non-tariff related, have either economically unimportant or statistically insignificant effects on the reported Armington elasticity. This evidence is, however, not entirely conclusive because the baseline and alternative prior specification (first panel of Table B2) offer an unintuitive sign for the coefficient of non-tariff barriers, even though the evidence for this coefficient is rather weak. Volatility in the exchange rate, moreover, shows a statistically and economically important positive effect. Finally, we do not find our proxy for home bias or the spread of Internet use important for the magnitude of the elasticity of substitution.

Estimation techniques. The evidence on the systematic importance of model and estimation techniques is rather mixed. The baseline unweighted specification does not offer a strong case for any of the model or method choices to have a systematic impact on the estimated elasticity. The baseline specification suggests that lower reported elasticities are associated with OLS and that larger elasticities are associated with a control for endogeneity. In the study-weighted specification, the usage of the *static model*, *nonlinear model*, and *GMM* seem to have not only statistically but also economically important effects. The static model is often used to capture the long-run effect using OLS. Non-linear models also typically apply GMM to capture the long-run Armington elasticity. Goldstein & Khan (1985) argue that single-equation estimation techniques commonly generate price elasticities biased downward because they constitute a weighted average of the actual demand and supply elasticity. GMM is also commonly applied to help with endogeneity issues in the estimation procedures (Aspalter 2016). The non-linear estimation technique is applied differently in different studies, but many follow Feenstra *et al.* (2018), which is currently considered the best practice in the literature. Next, Huchet-Bourdon & Pishbahar (2009) show that estimation ignoring *import tariffs*, for example, may produce biased results. Our results suggest that if a control for tariffs is not included in the estimation, elasticities indeed tend to be systematically larger. The control for *seasonality* in the estimation model, on the other hand, seems to diminish the estimated elasticities.

Publication characteristics. Our results indicate a remarkably strong association between publication characteristics (publication in a peer-reviewed journal, the impact factor of the outlet, and the number of citations) and the reported results of a study. We interpret this association as the effect of quality on the results: higher-quality studies tend to report substantially larger Armington elasticities. However, a qualification is in order. Publication bias can influence this association, for example, if peer-reviewed journal and generally better outlets prefer larger elasticities. Moreover, if researchers calibrating their models also prefer large elasticities, they may preferentially cite studies that deliver such estimates. We experimented with adding additional interactions of the standard error and the publication characteristics, but none proved important. Therefore, we find no evidence that the association between quality and the size of the reported Armington elasticity is driven by publication bias.

The results presented so far suggest that publication bias exaggerates the mean Armington elasticity but that many questionable data and method choices may result in a downward bias (and some may also do so in an upward bias). Finally, studies of higher quality tend to report larger elasticities. In the remainder of this section, we attempt to put all of this information together and derive the mean Armington elasticity implied by the literature and corrected for all biases related to publication selection, data and methods, and quality. To this end, we construct a synthetic study that uses all 3,524 estimates but gives each estimate a weight based on our baseline BMA results and our definition of a "best practice" study. Such best practices are inherently subjective, depending on our decision about the best choices for data, method, and publication choices. We execute several robustness checks to ensure that the important results hold in different but plausible settings.

The best-practices estimate is a result of a linear combination of the BMA coefficients from the baseline specification in Table 2.5 and our chosen values for the respective variables. We prefer the most precise estimates (and, as a consequence, no publication bias), so we plug in zero for the standard error. We focus on the long-run elasticity because the long-run effect is the area in which most policy makers are interested. We prefer the full disaggregation of data and results. We also prefer panel data, the maximum size of the dataset, and the maximum length of the data period. We plug in the maximum for the midyear of data used in individual studies because we want to give more weight to recent information. We also prefer studies published in peer-reviewed journals with a large impact factor and those with a high number of citations.[2]

To estimate the best-practice mean elasticity for our entire sample, we evaluate all the structural variables, including the country-specific variables at the sample mean. The estimates for individual countries in Table 2.6, on the other hand, are estimated using country-specific values for the cross-country variables (these include *developed* economies, *market size*, *tariffs*, *non-tariff barriers*, *FX volatility*, *national pride*, and *internet usage*). We prefer estimates obtained using nonlinear models and the GMM estimator and estimations controlling for import constraints and seasonality. We also prefer annual data because they abstract from short-term fluctuations that might obscure the estimates of the elasticity; in meta-analysis, there is no lack of power that would force us to move to a higher (and noisier) frequency.

Table 2.6 reports the results of our best-practice exercise. The elasticities

[2]Three variables display large outliers: the number of *citations*, *data size*, and *impact factor*. To ensure that our estimates are not driven by the outliers, we take the 95$^{\text{th}}$ percentile of the value of these variables in our dataset. If we took the maximum, our resulting estimate of the elasticity would be larger.

Table 2.6: Armington elasticities implied for individual countries

	Mean	95% conf. int.	
Australia	3.2	1.8	4.6
Austria	2.8	1.5	4.0
Belgium	2.8	1.4	4.1
Brazil	3.2	1.4	5.0
Bulgaria	3.1	1.6	4.6
Colombia	3.4	1.4	5.4
Cyprus	3.0	1.5	4.6
Czech Republic	3.6	2.3	4.9
Denmark	2.7	1.3	4.0
Estonia	3.0	1.6	4.5
Finland	2.7	1.4	4.0
France	2.7	1.4	4.0
Germany	2.7	1.5	4.0
Greece	2.8	1.5	4.0
Hungary	3.1	1.8	4.4
Ireland	2.8	1.5	4.1
Italy	2.7	1.4	4.0
Japan	3.2	2.1	4.3
Latvia	3.0	1.6	4.4
Lesotho	3.9	1.8	5.9
Lithuania	3.0	1.6	4.3
Luxembourg	3.0	1.5	4.6
Malta	3.1	1.5	4.8
Netherlands	2.6	1.4	3.8
Poland	3.0	1.7	4.3
Portugal	2.9	1.5	4.4
Romania	3.1	1.7	4.5
Russia	3.4	1.6	5.1
Slovak Republic	3.1	1.6	4.5
Slovenia	2.9	1.5	4.4
South Africa	3.4	1.6	5.3
Spain	2.7	1.4	4.0
Sweden	2.7	1.5	3.9
Thailand	3.1	1.2	5.0
United Kingdom	2.9	1.8	4.1
United States	2.4	1.1	3.7
Uruguay	3.4	1.6	5.2
Euro area	2.6	1.5	3.7
All countries	2.9	1.3	4.4

Notes: The table presents the mean estimates of the Armington elasticity implied by the Bayesian model averaging exercise and our definition of best practices. The confidence intervals are approximate and constructed using the standard errors estimated by OLS.

implied for different countries after correction for publication and other biases range from 2.7 to 3.4; the mean estimate for the entire world is 2.9. The mean elasticities would be even larger if we preferred quarterly data instead of annual data, pushing the corrected mean to 4 for the overall sample (3.7 for the European Union and 3.6 for the United States). The elasticity would also be larger if we took the maxima instead of the 95% percentiles for data size, the impact factor of the outlet, and the number of citations, and if we preferred TSLS instead of GMM and cross-sectional data instead of panel data. The 95% confidence intervals, although quite wide, imply that the aggregate Armington elasticity of substitution is above 1.3 with a 95% probability. This finding resonates with Imbs & Mejean (2015) and their call for elasticity optimism.

2.5 Concluding Remarks

We present the first quantitative synthesis of the vast empirical literature on the elasticity of substitution between domestic and foreign goods, also known as the macro-level Armington elasticity (Feenstra *et al.* 2018). The elasticity is a key parameter for both international trade and international macroeconomics. In computable general equilibrium models commonly used to evaluate trade policy, the elasticity of substitution governs the effects of newly introduced tariffs, among other things. In open-economy dynamic stochastic general equilibrium models used by many central banks to evaluate and plan monetary policy, the elasticity of substitution governs the strength and speed of the exchange rate pass-through.

Consider, for example, two European central banks that, in the wake of the Great Recession, introduced exchange rate floors to limit their currencies' appreciation against the euro: the Swiss National Bank and the Czech National Bank. Currency depreciation (relative to the counterfactual without the currency floor) produces two effects relevant to the aggregate price level. First, imported goods become more expensive, which directly increases inflation. With a large elasticity of substitution between domestic and foreign goods, however, this effect becomes muted and delayed because consumers shift toward relatively cheaper domestic goods. Second, currency depreciation stimulates the economy by encouraging exports and discouraging imports, which raises inflation in the medium term. With a larger elasticity of substitution, this effect strengthens. Because both the Swiss National Bank and the Czech National Bank use open-economy dynamic stochastic general equilibrium models for policy analysis, the assumed size of the Armington elasticity played an important (if implicit) role in the decision on when and how to implement the exchange rate floor.

We collect 3,524 previously reported estimates of the Armington elasticity, which makes our paper one of the largest meta-analyses conducted in economics so far. Ioannidis *et al.* (2017b) survey 159 economics meta-analyses and report that the mean analysis uses 400 estimates. We also construct 34 variables that reflect the context in which researchers obtain their estimates. Several characteristics of the studies and individual estimates might affect the results systematically, as was claimed by previous studies: for example, the level of data aggregation (Hummels 1999), data frequency (Hertel *et al.* 1997), the distinction between short- and long-run effects (Gallaway *et al.* 2003), and estimation strategy (Cassoni & Flores 2008). Other studies stress the potential importance of structural determinants of the Armington elasticity at the industry or country level (Blonigen & Wilson 1999; Lopez & Pagoulatos 2002; McDaniel & Balistreri 2003). Our aim in this paper is to assign a pattern to the great variation observed among the reported estimates of the elasticity.

Our results, based primarily on the Bayesian and frequentist model averaging that address the model uncertainty inherent to meta-analysis, suggest that the single most important variable for the explanation of the variation in the reported elasticities is the standard error. Large standard errors are associated with large estimates, which is inconsistent with the property of almost all techniques used to estimate the elasticity: the ratio of the estimate to its standard error has a t-distribution (or other symmetrical distribution). The property implies that estimates and standard errors should be statistically independent quantities. The violation of independence suggests a preference for large estimates that compensate for large standard errors, which we further corroborate by employing the new non-linear techniques by Ioannidis *et al.* (2017b), Andrews & Kasy (2019), and Furukawa (2019). This publication selection results in an exaggeration of long-run estimates by more than 50% on average. After correcting for publication bias, we observe no systematic difference between the reported sizes of short- and long-run elasticities.

We find that a large part of the variation in reported elasticities can be explained by data characteristics. In particular, data aggregation, low frequency, and small samples are typically associated with smaller estimates. In contrast, the use of cross-sectional data tends to result in large estimates: on average, greater by 1.7 than when time-series data are used. After controlling for these data characteristics, we find no association between data age and the size of the reported elasticity. Thus, the larger elasticities reported by more recent studies are typically given by the move from time-series to cross-sectional data analysis. Our results also suggest that study quality (roughly approximated by publication status, the RePEc impact factor of the outlet, and the number of citations) is robustly associated with study results: higher-quality studies tend to report larger elasticities. We use all of this information to construct a synthetic study that draws on all 3,524 estimates but gives more weight to better estimates and controls for publication bias. While defining "better" estimates is inevitably subjective, we argue that, given plausible definitions of best practice, the best possible guess concerning the aggregate Armington elasticity is close to 3—at least based on the empirical research of the last 50 years since Armington (1969).

Three qualifications of our results are in order. First, the 3,524 estimates that we collect are not independent but likely correlated within studies and countries. We try to account for this problem by using Bayesian hierarchical analysis and clustering the standard errors (where possible) at the level of both studies and countries. Second, while we control for 34 aspects of studies and estimates, one could still add more variables, as the pool of potential controls is unlimited. We omit industry-level variables, for example, because their inclusion would cause serious collinearity. But the entire dataset together with the code is provided in the online appendix and allows interested researchers to focus on different subsets of variables. Third, while we do our best to include all studies reporting an estimate of the macro-level Armington elasticity, we might have missed some. This potential omission does not create a bias in meta-analysis as long as it is not conditional on study results.

Chapter 3

Bank Capital, Lending and Regulation: A Meta-analysis

Simona Malovana, Martin Hodula, Josef Bajzik, Zuzana Gric

Abstract We collected over 1,600 estimates on the relationship between bank capital and lending and construct 40 variables to capture the context in which these estimates are obtained. Accounting for potential publication bias, we find that a 1 percentage point (pp) increase in capital (regulatory) ratio results in around 0.3 pp increase in annual credit growth, while changes to capital requirements cause a decrease of around 0.7 pp. Using Bayesian and frequentist model averaging, we show that the relationship between bank capital and lending changes over time, reflecting the post-crisis period of increasingly demanding bank capital regulation and subdued profitability. We also find that the reported estimates of semi-elasticities are significantly influenced by the empirical approach chosen by researchers. Our findings suggest that the literature fails to provide policymakers with reliable estimates of the effects of capital regulation on bank lending, and our study offers insights that could help guide future research.

This paper was published in *Journal of Economic Surveys[2023, pp. 1-29, DOI: 10.1111/joes.12560]*. The views expressed are those of the authors and not necessarily those of the Czech National Bank.

3.1 Introduction

There is extant empirical literature assessing the effect of changes to bank capital on the extension of bank credit. The importance of quantifying this relationship has been one of the most pivotal research questions for almost two decades. The topic was given particular attention following the onset of the 2007–2009 Global Financial Crisis (GFC), when the likelihood of a credit crunch was under debate and again when the first quantitative easing programs were gradually implemented. The question has reemerged more recently with the gradual implementation of Basel III and an increasing use of macroprudential policy instruments. Following the implementation of Basel III, the observed minimum capital requirements effectively rose from 8% to 10.5%. However, due to all the additional prudential buffers, the capital requirements were able to reach as much as 20% (BCBS 2010). It is thus not surprising that the current research concerning the relationship between bank capital and lending has shifted towards assessing the effects of capital regulation on bank lending capacity.

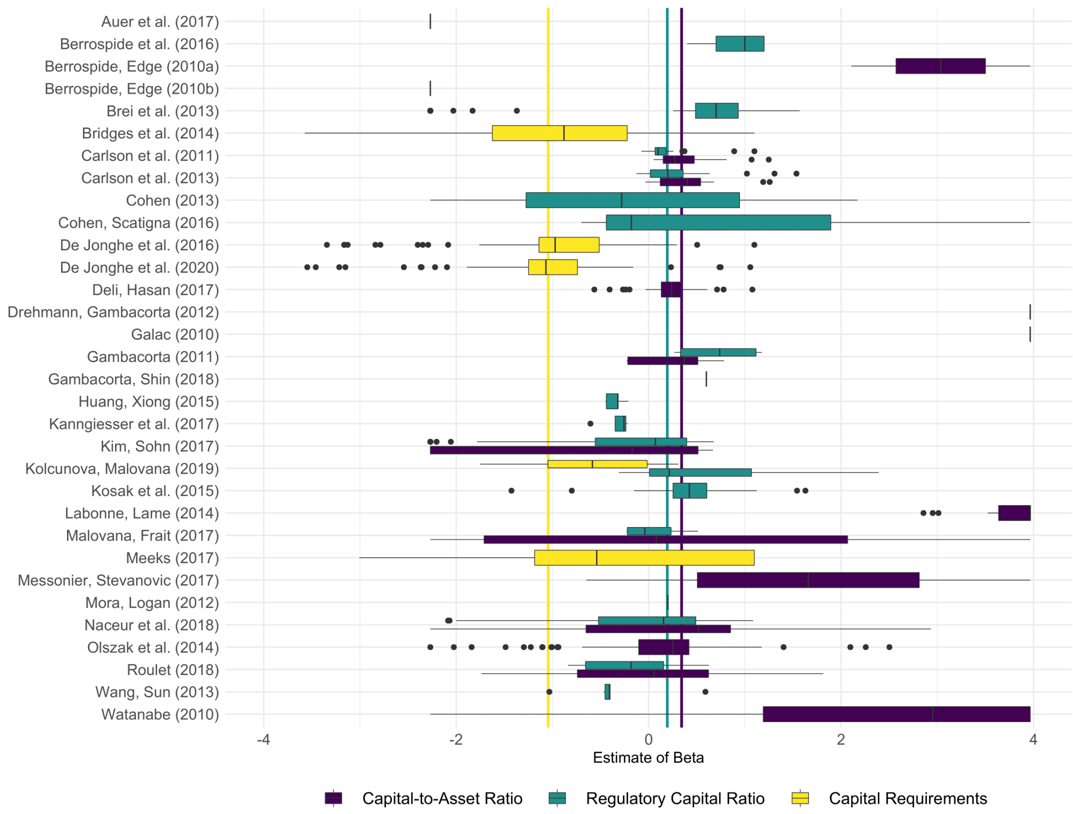
A conspicuous feature of the bank capital-lending literature is that the bank capital ratio may change for various reasons, ranging from regulatory (Peek & Rosengren 1997; De Jonghe *et al.* 2020) to economic and managerial (Houston *et al.* 1997; Berrospide & Edge 2010b; Gambacorta & Marques-Ibanez 2011a). As a result, there is a wide range of possible outcomes when quantifying the impact of changes in bank capital on bank lending. On the one hand, an increase in the bank capital ratio due to the introduction of a new capital regulation may dampen bank lending activities as a bank would try to avoid the higher costs of financing loans by capital (De Jonghe *et al.* 2020). On the other hand, a general increase in the bank capital (equity) ratio due to, for example, bank profit accumulation should be reflected in an increase in lending, suggesting a positive effect (Berrospide & Edge 2010b).

In this paper, we conduct a thorough review of the empirical literature on how changes in bank capital affect credit dynamics. Our approach involves gathering over 1,600 estimates from 46 papers that examine the relationship between bank capital and lending. Throughout the literature, there are three expressions of the capital ratio: a simple capital to asset ratio, a regulatory capital ratio that includes Common Equity Tier 1, Tier 1, and Tier 2 capital over risk-weighted exposures, and a capital requirements ratio that is defined as capital requirements over risk-weighted exposures. We note from Figure 3.1

that the literature exhibits significant fragmentation regarding the estimated coefficients that goes beyond the different expression of capital ratios.

To explain the differences, we collect an additional 40 variables that reflect the context in which the estimates were produced. The newly created database allows us not only to derive an "average" effect but also to explain why the estimates vary across different studies and to describe what the most commonly employed empirical strategy is. We use state-of-the-art meta-analytic techniques to estimate the true effect of changes to bank capital on bank lending, as well as the model averaging methods used to identify the significant drivers of the heterogeneity of the observed estimates.

Figure 3.1: Reported Estimates Vary Both Within and Across Studies



Note: The figure depicts a boxplot of the collected estimates on the effect of a 1 pp increase in capital ratio on bank annual credit growth (see equation 3.1). The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures. The length of each box represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. Short black vertical line denotes unitary semi-elasticity. Long colored vertical lines indicate the median semi-elasticity for each category of capital ratio. For ease of exposition, extreme outliers are excluded from the figure but included in all the statistical tests.

The collected estimates imply that a 1 pp increase in the simple capital-to-asset ratio is associated with a 0.3 pp increase in annual credit growth. Considering the regulatory capital ratio, the average estimated semi-elasticity decreases to about 0.2 pp. Neither of these two exhibits signs of publication bias and thus the estimated true effect is very close to a simple average across the collected semi-elasticities. In stark contrast, correcting for publication bias shrinks the mean semi-elasticity on capital requirements from -1.7 pp to around -0.7 pp.

Next, our findings indicate that various study characteristics are systematically associated with the reported results. Among the 40 variables we construct, the most important for model averaging are those related to data, the estimation technique and cross-country or regional differences. Specifically, we find that single-country studies with larger sample sizes exhibit a positive correlation with the collected semi-elasticities, while studies shielded from omitted variable bias with more favorable publication characteristics are generally negatively correlated with the reported estimates. Apart from data characteristics, estimates of the effect of changes to the simple capital-to-asset ratio are also found to be dependent on the variables reflecting the macro-financial characteristics of the countries analyzed. The heterogeneity in the estimates based on the regulatory capital ratio can thus mostly be explained by model specification. In the case of the literature on capital requirements, the standard error is the most important variable in terms of explaining the variation in the reported estimates. Large standard errors are associated with more negative estimates, supporting the existence of publication bias in this category.

Taken together, results of our meta-analysis suggest that the existing literature fall short on providing the policymakers with clear and reliable estimates of the effects of capital regulation on bank lending. In fact, we find that only a handful of studies capture the “pure” effects of changes in bank capital due to a newly endorsed capital regulation. Those few rely mostly on (semi)natural experiments, while the vast majority of studies rely on more or less precise identification strategies.

Therefore, to aid policymakers, we use the information obtained from the collected studies and the heterogeneity analyses to compute the mean semi-elasticity of the relationship between bank capital and lending based on the design of the most reliable studies. Using this, we attempt to show what the mean semi-elasticity would be if all studies used the same strategy as our preferred approach. Given its high policy relevance, we favor such primary study characteristics that represent a consistent and unbiased estimator and could better capture changes in bank capital due to capital regulation rather than other factors. We find that semi-elasticities implied by significant heterogeneity drivers are distinctly positive for the simple capital-to-asset ratio (0.4 pp) and negative for the regulatory capital ratio (-0.6 pp). Interestingly, both semi-elasticities turn negative when a prolonged period of low interest rates is considered. This shows that, as the post-GFC period of record-low interest rates was accompanied by strong macroprudential policy activity, capital regulation become the main driver of shifts in bank capital ratios.

Our paper relates to two strands of literature. First, this paper contributes to the broader empirical literature on the effects of changes to bank capital on the financial sector. The literature has focused predominantly on testing the implications of the bank capital level on the probability of crises (Demirguc-Kunt *et al.* 2013; Jordà *et al.* 2021) and finding the optimal capital level (Miles *et al.* 2013; Thakor 2014; Schwert 2018). A meta-analysis of Kočenda & Iwasaki (2022) explore the impact of the so-called CAMELS indicators (including capital adequacy) on bank survival, suggesting an economically negligible impact. Our paper provides a novel (and, to the authors' knowledge, the first) comprehensive synthesis of the empirical literature on the relationship between bank capital and lending. Previous related studies include VanHoose (2007) and Kashyap *et al.* (2010) who provide a narrative review of the theoretical and empirical literature. Second, our paper also contributes to the emerging literature on the effects of bank capital regulation on the real economy and the financial sector. Two recent meta-analytic studies include Araujo *et al.* (2020) who estimate the average effects of macroprudential policy on bank credit, house prices and the real economy, and Fidrmuc & Lind (2020) who present a meta-analysis of the impact of higher capital requirements on macroeconomic activity. In both cases, the meta-analyses are performed on a set of studies that rely on various dummy-coded indices capturing changes to bank capital regulation. A typical shortcoming associated with the dummy approach is the inability to actually quantify the effects of regulatory policies which is generally a key issue for policymakers (Alam *et al.* 2019). In our paper, we opt for the time-series approach where we select only those papers capturing continuous changes to bank capital.

We perceive the contribution of this paper to be threefold. First, quantifying the effect of changes to bank capital on the supply of credit is of utmost importance to policymakers. Obtaining a comprehensive overview of the findings of the literature goes well beyond the scope of individual studies that are, by nature, very selective. Second, we show the caveats associated with modelling the relationship between bank capital and lending as well as inform about the most commonly employed practices. Third, we present some indications that the relationship is changing over time and discuss the implications this would have for correctly estimating and assessing the impact of capital regulation.

The remainder of this paper is structured as follows: Section 3.2 introduces the interplay between bank capital, capital requirement and bank lending. Section 3.3 describes how we collect data from primary studies. Section 3.4 tests for publication bias and estimates the effect beyond bias. Section 3.5 explores the heterogeneity of the estimated semi-elasticities and Section 3.6 concludes.

3.2 Bank Capital, Capital Requirements and Lending

The level of capital is central to bank lending decisions, at least under the conditions of an imperfect market for bank equity and the existence of minimum capital requirements. Consider a well-capitalized bank with access to additional sources of capital. Such a bank will be able to accommodate any capital losses without having to reduce its assets (and hence its lending). Now consider the polar case where a bank actively manages its portfolio in order to maintain a constant capital ratio. For such a bank, with an observed capital ratio of 8%, a dollar reduction in capital would lead to a $12.5 reduction in its assets, including loans. Raising the bank capital ratio to 10.5% – the minimum level under Basel III – should therefore lead to a $9.5 reduction in assets during, let's say, times of crises, i.e. a reduction of almost 24%. This may well be the longer-term effect of raising the bank capital level. How much we would deviate from this idealized scenario in real terms and what the intermediate effect of increasing the bank capital ratio would be is, of course, an empirical question, and not an easy one, as the reason behind the increase in the bank capital ratio is often not directly observable.

The fact that the reasons behind changes in the bank capital ratio may vary largely complicates the empirical efforts to measure the effects of changes to bank capital on bank lending. With a reasonable degree of simplification, bank capital may change due to a regulatory change (e.g. a change in the minimum capital requirements) or for any other managerial or economic reason. A large body of research is focused on the former, i.e. bank behavior under capital regulation which cannot be removed from the relationship between bank capital and bank lending. In fact, early studies date way back to the 1990s when Basel I was introduced. Back then, many observers debated whether the newly introduced capital regulations were inhibiting lending (Bernanke *et al.* 1991; Hancock & Wilcox 1994; Berger & Udell 1994). This issue was reinstated following the most recent Basel Accord (Basel III), with the debate shifting to the costs associated with stricter capital requirements as compared to the benefits arising from greater financial and macroeconomic stability (Beltratti & Stulz 2012; Berger & Bouwman 2013; Thakor 2014).

With the ever-increasing use of both micro- and macro-prudential policy instruments, it is no wonder that researchers are urged to incorporate the regulatory constraints faced by banks in their estimates of the relationship between bank capital and lending. In fact, macroprudential policy, which is aimed at ensuring financial sector stability and resilience, has slowly gained prominence as a third economic policy. To achieve its goal, (macro)prudential policy has several tools at its disposal. For banks, to which such policy most commonly applies, these tools include capital- and borrower-based measures. Capital-based measures have been frequently used in both advanced and emerging market economies. Their importance and frequency of use increased significantly following the emergence of the GFC (Cerutti *et al.* 2017a; Alam *et al.* 2019). Capital-based measures encompass capital requirements aimed at increasing the loss-absorbing capacity of banks and the overall financial sector resilience to shocks of a different nature. By altering banks' funding costs, capital-based measures may also affect credit intermediation by banks. While there is no official macroprudential policy target, the immediate focus on credit dynamics is justified by the well-documented fact that credit booms typically precede crises (Jordà *et al.* 2011; Schularick & Taylor 2012).

A regulatory shock to the bank capital ratio is expected to be decisive for bank lending decisions if two conditions hold true: First, a bank – in response to heightened capital requirements – changes its funding structure in favor of equity. A contrasting case would be a bank holding capital well in excess of the minimum capital requirement, i.e. one which maintains a "capital surplus". Under such circumstances, increasing additional capital requirements may have a limited effect on the capital adequacy ratio of a bank simply because it would use the extra capital, thus shrinking the surplus.[1] Second, the observed change

[1]In fact, we can assume than even a bank holding a capital surplus would change its lending behavior following a regulatory tightening. A bank faces internal or implicit costs of

in the structure of bank funding increases bank funding costs. This builds on the presumption that equity is more costly than debt which is a generally accepted condition (Kashyap *et al.* 2010).

There is widespread agreement in the theoretical academic literature that the immediate effects of constraining capital standards are likely to be a reduction in total lending (VanHoose 2007). Empirical evidence is far more inconclusive. This stems from the several perils associated with the modelling of the relationship between capital and lending. For one, there are the identification issues. Since many capital-based regulatory measures are taken in response to developments in the financial sector, researchers are always running on edge with the risk of reverse causality that might bias their estimates. Regulatory actions may well coincide with faster growth in lending due to the focus of the policy itself. And even in the literature exploring the relationship between bank capital and lending beyond the regulatory cap, it is hard to tell if a change in capital is causing a change in lending, rather than reflecting it. During an economic recession, banks generally record larger losses on their existing loan portfolios, and this ultimately reduces their capital stock. Needless to say, during a recession, there are worse lending opportunities too. Thus, even in this strand of literature, it is the goal of researchers to make sure that their estimates are purged of these bias-inducing effects. The other broad challenge is to separate changes in bank capital due to a change in capital regulation from changes stemming from economic conditions or management decisions. As we outlined in this section, the source of the change in bank capital can be decisive for bank lending, with the two likely going in the opposite direction.

In the current strand of literature, the most valued studies use "natural experiments" where a shock to bank capital is perceived as exogenous and uncorrelated with any lending opportunities. For instance, Peek & Rosengren (1997) exploit a regulatory change concerning the US branches of Japanese banks to identify how shocks to capital impact loan supply. Nowadays, these natural experiments are often backed-up by detailed credit registry data which allow us not only to identify exogenous shocks to capital but also to dismantle credit supply/demand movements (see, for example, De Jonghe *et al.* 2020). Another approach to dealing with endogeneity issues is to separate banks according to different characteristics and compare the sub-samples. In this respect, Bernanke *et al.* (1991) is representative of the many studies that compare

funds, which are set on a consolidated basis. Further, a bank often sets up internal capital ratio targets above the minimum level dictated (Berrospide & Edge 2010b).

different groups of banks to assess the importance of capital shocks on lending. Other studies typically divide banks according to their capital level or capital surplus and hypothesize that banks with a low level of capitalization will base their lending decisions more on changes in capital. Hancock & Wilcox (1993) and Hancock & Wilcox (1994) have conducted prominent studies in which they estimate bank capital target functions and find significant correlations between capital relative to its target and subsequently to bank lending.

3.3 Collection Process and Dataset Formation

The main purpose of this paper is to explore the effect of bank capital and capital regulation on lending. As such, we do not limit our analysis to the relationship between capital requirements and loan supply, but also explore the impact of overall bank capitalization, both risk-sensitive and insensitive. Understanding the role of bank capitalization is integral to correctly assessing and anticipating the transmission of additional capital requirements. Our general knowledge of the existing literature on this topic and the first bird's-eye view of some prominent studies gave us the impression that we would potentially face significant heterogeneity in the variable definition, their transformation and the identification strategy. Therefore, we decided to provide a comprehensive overview which would inform the reader not only about the true effect but also about the predominant model specification, including the role of the variable definition, data characteristics and the researchers' preferred estimation approach.

In our selection procedure, we considered all the empirical studies involving some form of bank capital or capital requirements on the right-hand side of the relationship and lending on the left-hand side, regardless of the variable transformation. We collected 1,639 estimates from 46 studies. Our data collection procedure is described thoroughly in the online appendix. We limited our search to studies published in 2010 and later to account for the fact that capital regulation has been used more extensively since the GFC (Alam *et al.* 2019).

Although we collected estimates on all possible variable transformations, we found that the majority of them (85%) are based on the same transformation – credit growth and the level of capital ratio. The remaining estimates are scattered between different transformations, making direct comparison impossible. Even within the category of the growth-ratio transformation we encounter some heterogeneity, especially with respect to the definition of the capital ratio, which we explore further in our paper. To summarize the results, we calculate partial correlation coefficients, which show that the definition of capital ratio is more important than the preferred variable transformation.[2] Thus, we focus only on estimates from the growth-ratio transformation, which allows us to directly quantify the effect of changes to bank capital on lending and preserve the economic interpretation of the estimated effect. This approach provides a significant benefit compared to similar meta-analytic studies relying on dummy-coded macroprudential indices. Such findings will enable us to draw more convincing conclusions, not only on the true direction of the analyzed effect but also on its true size. This transformation is used in 32 studies and comprises a total of 1,395 estimates.

As a result, the semi-elasticities β entering the analysis in sections 3.4 and 3.5 refer to the following equation:

$$\%\Delta L_{it} = \beta CR_{it} + \gamma X_{it} + \epsilon_{it} \tag{3.1}$$

where $\%\Delta L_{it}$ is annual credit growth, CR_{it} is bank capital ratio and X_{it} is a vector of control variables for time t and unit i (country or bank).

In the absence of a "unified policy function", the literature includes several variations of equation (3.1), involving a different type and definition of credit variable or capital ratio, a different set of control variables or a different estimation approach. For instance, our sample studies consider three different types of capital ratios: a simple capital to asset ratio, a regulatory capital ratio which includes Common Equity Tier 1, Tier 1 and Tier 2 capital over risk-weighted exposures, and capital requirements which are defined as capital requirements (minimum, Pillar 2 add-ons and capital buffers) over risk-weighted exposures. We use state of the art meta-analytic techniques to construct summaries of the estimated semi-elasticities, aiming to verify the presence of publication bias, as well as to explain why the estimates may vary.

3.3.1 Early View of the Fragmentation

A bird's-eye view of the collected semi-elasticities suggest four stylized facts. First, estimates on the relationship between bank capital and lending vary sub-

[2]The partial correlation coefficient is a standardized measure of effect size that enables the comparison of estimates across different units of measurement. For further details, please consult the online appendix.

stantially, ranging from positive to negative values with a mean semi-elasticity of -0.09 but with median semi-elasticity of 0.15 (Table 3.1). For example, a central banker, wishing to incorporate the bank capital-lending semi-elasticity into a stress-testing framework, would have a difficult job finding the "correct" semi-elasticity value. The increased variance in the estimated semi-elasticities provides solid ground for a systematic evaluation of the published results which is shown in the next two sections.

Table 3.1: Breakdown into Categories of Different Capital Ratios

	Obs.	Articles	Mean	Median	5%	95%	Skewness
Total	*1,395*	*32*	*-0.093*	*0.149*	*-2.269*	*1.538*	*-2.72*
Capital-to-asset ratio	514	17	0.345	0.342	-2.221	3.794	0.60
Regulatory capital ratio	652	18	0.138	0.194	-1.383	1.114	0.10
Capital requirements	229	5	-1.737	-1.044	-8.926	0.302	-2.58

Notes: The table presents summary statistics of the collected estimates on the effect of a 1 pp increase in capital ratio on bank annual credit growth (see equation 3.1) winsorized at the 2.5% level from each side. Some articles include multiple different capital ratios; therefore, the sum of the articles across the different categories reported in the third column (the sum of 17, 18 and 5) exceeds the total number of primary studies included (32). The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures.

Second, differences in the semi-elasticity values can seemingly be well explained by the researcher's initial choice on how to express the bank capital ratio (CR_{it}). Table 3.1 makes it apparent that the use of a capital-to-asset ratio generates semi-elasticity estimates skewed towards the positive spectrum of the semi-elasticity distribution. The mean semi-elasticity value comes in at 0.35. A simple capital-to-asset ratio is generally used to capture banks' capital position (capitalization) and therefore, a positive effect on bank lending is expected. In stark comparison, using capital requirements generates negative semi-elasticity estimates centered around a mean value of -1.74. We collected estimates from five studies on capital requirements which mostly examine the impact of Pillar 2 capital requirements but also the changes to overall capital requirements, including various capital buffers (more details are provided in the online appendix). Considering regulatory capital ratios (e.g. the Tier 1 Basel regulatory ratio) then generates mean semi-elasticity estimates of 0.14 which are slightly more skewed to negative values than a simple capital-to-asset ratio. Apparently, studies where the capital ratio takes into account the riskiness of banks' operations are more likely to capture the effects of changes to bank capital under capital regulation where a negative effect can be expected.

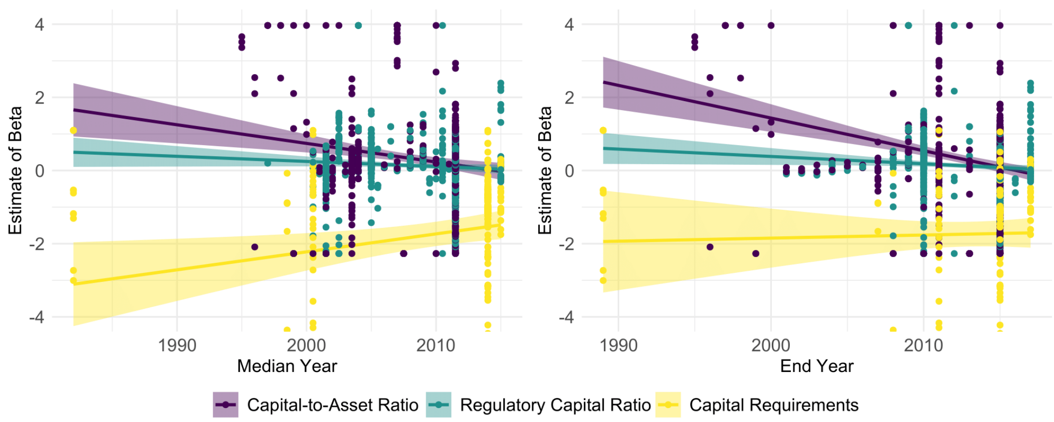
Third, it might be surprising that even when considering changes to Basel regulatory ratios, the literature does not paint a clear picture of the effects on bank lending. This could be due to several factors that might be in play. First, banks can change their actual capital ratio quite frequently and for various reasons that are not necessarily linked to changes in the capital regulation (see, for example, Guidara *et al.* 2013; Almazan *et al.* 2015; Bahaj *et al.* 2016). Thus, estimates based on observed capital ratios are noisier indicators of what regulatory changes may imply than estimates based on regulatory requirements. Second, banks typically hold capital in excess of what is required by the regulator (capital surplus), the level of which has been shown to be decisive for the response of bank lending to a shock to capital (Berrospide & Edge 2010b). It is therefore less likely for a shock to the bank capital ratio to be binding for a bank with a high capital surplus, simply because it would use the extra capital and allow the capital surplus to shrink (Kolcunová & Malovaná 2019).

Fourth, the reported semi-elasticity follows an interesting pattern in time (Figure 3.2). Estimates shifted towards more recent period tend to report semi-elasticity estimates at closer to zero. This is found to be true for both groups of studies, i.e. those using a simple capital-to-asset ratio and those with a regulatory capital ratio. A similar trend has been observed in other meta-analyses, and this phenomenon is commonly referred to as "time-lag bias" (Ioannidis 1998; Astakhov *et al.* 2019; Xue *et al.* 2022). This bias may arise because larger and more significant estimates are more likely to be published earlier. In our line of work, the explanation could be that shocks to bank capital are getting smaller through time. In another words, a one unit change to a bank capital ratio (simple or regulatory) demands a lower response in terms of bank lending. In case of a regulatory shock, this would mean that it was historically more costly for banks to raise external equity. Studies performed on earlier datasets would thus represent the upper bound of the possible effects of regulatory shocks on bank capital.

Another explanation is that we see the empirical manifestation of a "model risk", stemming from the pro-cyclicality of the Basel accords (Le Leslé & Avramova 2012). One of the features of banking regulation under the Basel III is that banks can determine risk-weights themselves, using internal models (IRB) rather than pre-set values for a given asset class. Model risk posits that under the IRB, the risk-weights are smaller and more pro-cyclical. Linked to the reported semi-elasticity, the existence of a model risk may weaken the effects of capital regulation. Yet another explanation would be that bank regulators have exploited the phasing-in of the new capital requirements more often, giving banks enough time to take the newly required capital out of retained earnings instead of cutting back on lending. In case of a shock to bank capitalization, the Basel II and III accords have led to a general increase in bank capital levels and surpluses. A negative shock to bank capital is thus expected to demand a lower response rate in terms of lending dynamics.

In the analyses to come, we focus on each of the capital ratio separately given their directly observable differences.

Figure 3.2: Reported Estimates Change Over Time



Note: The figure depicts a scatter plot of the collected estimates on the effect of a 1 pp increase in capital ratio on bank annual credit growth (see equation 3.1) relative to the median (left panel) and end (right panel) year. The median and end year are calculated for each primary study and capital ratio based on the time period of the data sample used in the estimation. The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures. For ease of exposition, extreme outliers are excluded from the figure but included in all the statistical tests.

3.4 Publication Bias

Publication bias occurs when there is a systematic difference between the distribution of the results produced and those reported by the researchers. Even the best-published study in our meta-analytic data set[3], Gambacorta & Marques-Ibanez (2011a), admits that publication bias may be an issue in the literature on bank capital and lending:

> *"The coefficient on the standard capital-to-asset ratio often has an incorrect negative sign, which casts some doubt on the role of this indicator in capturing the effect of a bank's capital position on bank lending."*

[3]We identify the "best-published" study based on the recursive impact factor and number of citations in Google Scholar.

On the same note, the theoretical literature seems to have set up the "correct" relationship between more stringent capital regulation and lending (VanHoose 2007):

> *"There is widespread agreement in the theoretical academic literature that the immediate effects of constraining capital standards are likely to be a reduction in total lending."*

The findings that do not fit the generally accepted narrative, i.e. where the researcher estimates the effect to be positive or not statistically significant, may thus be sensitive to publication bias. This behavior is described as the Lombard effect (McCloskey & Ziliak 2019), originally defined in biology, which refers to a situation where researchers (intentionally or not) need to try harder to achieve the estimates consistent with their intuition if the data are imprecise or noisy. To investigate publication bias, we use established techniques, such as graphical and econometric analysis, which assume that the estimated effect β and its standard error should be independent. Econometric methods used to estimate β typically produce a symmetrical distribution, resulting in zero correlation between estimates and standard errors. However, intentional or unintentional reporting of statistically significant results by researchers can lead to a correlation between estimates and standard errors.

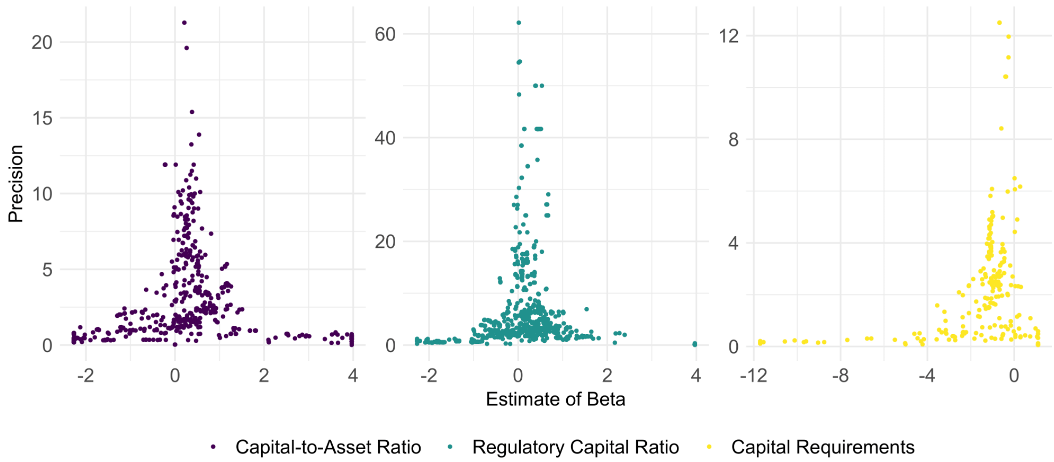
Graphically, the publication bias can be examined using a funnel plot that relates the estimate to its precision, measured by the inverse of the estimated standard error. In the absence of publication bias, precise estimates should cluster near the underlying mean value of the parameter, while less precise estimates should be more dispersed, creating a funnel shape. Therefore, if the estimated semi-elasticities are symmetrically distributed around the mean effect, there is likely limited publication bias.[4] However, if there is dominant and systematic heterogeneity, the funnel plot may be asymmetric. We will explore this heterogeneity and its underlying causes in the following section.

Figure 3.3 displays our funnel plots, categorized by capital ratios. The semi-elasticities linked to capital requirements appear to be heavily skewed towards negative values and form an asymmetrical funnel, indicating the presence of publication bias (the right panel). However, the funnel gives a mixed message if we look at the simple capital-to-asset ratio or regulatory capital ratio (the left and middle panel). In both cases, the most precise estimates are centered around slightly positive values. For simple capital-to-asset ratios, the

[4] A funnel plot, as a simple graphical measure of publication bias, was first proposed by Egger *et al.* (1997b).

right portion of the funnel might be a little heavier than the left one while for the regulatory capital ratio, the distribution is skewed more towards negative values. Nevertheless, the funnel plot is only a simple visual test, and the dispersion of the estimates could suggest heterogeneity in data and methods, the other systematic factor driving the estimated coefficients. To support our initial impression, we complement the visual analysis with a battery of formal regression-based tests.

Figure 3.3: Funnel Plots



Note: The figure depicts a funnel plot of the collected estimates on the effect of a 1 pp increase in capital ratio on bank annual credit growth (see equation 3.1). Precision is calculated as the inverse of standard error. The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures. Collected estimates are winsorized at the 2.5% level from each side.

Publication bias can be econometrically tested using a wide range of methods. As stated earlier, if there is no publication bias, estimates and their standard errors should not be correlated, and therefore, we can test the following specification (Stanley 2005; 2013):

$$\beta_{i,j} = \alpha + \gamma SE_{i,j} + \epsilon_{i,j} \tag{3.2}$$

where $\beta_{i,j}$ is the ith estimated semi-elasticity and $SE_{i,j}$ its standard error for each study j; α is the effect beyond bias and γ is the intensity of publication bias. If the γ coefficient is statistically significant, publication bias is present. The underlying mean effect corrected for publication bias is then captured in the coefficient α. However, this specification exhibits heteroskedasticity because the right-hand-side variable (SE) accounts for the variance of the left-hand-side variable (β estimate). To address this issue, we divide equation (3.2) by the standard error of the estimate, as suggested by, for example, Zigraiova *et al.* (2021); Stanley (2013; 2005):[5]

$$t_{i,j} = \gamma + \alpha \frac{1}{SE_{i,j}} + \omega_{i,j} \tag{3.3}$$

where $t_{i,j}$ represents the t-statistics for the ith estimated semi-elasticity from study j. The coefficients α and γ continue to capture the effect beyond bias and the intensity of publication bias, respectively.

The procedure outlined is known in meta-analytical literature as the FAT-PET, which combines the funnel plot asymmetry test (FAT) and the precision-effect test (PET). We supplement this approach by using a quadratic model of publication bias called PEESE (precision-effect estimate with standard error), which was developed by Stanley & Doucouliagos (2014):

$$t_{i,j} = \gamma SE_{i,j} + \alpha \frac{1}{SE_{i,j}} + \xi_{i,j} \tag{3.4}$$

where the coefficients α and γ again represent the effect beyond bias and the intensity of publication bias, respectively. This model relaxes the assumption of linearity between the estimated effect β and its standard error. The rationale behind squaring the standard error is that small studies (i.e., studies with high standard errors) are more susceptible to reporting exaggerated effect sizes. This problem is believed to be less prevalent in studies with high statistical power. While the PET method is most effective when the true effect is zero, the PEESE method outperforms it when the true effect is not zero. To take advantage of the strengths of both methods, Stanley & Doucouliagos (2014) propose combining them.

In accordance with Zigraiova *et al.* (2021), we estimate equations (3.3) and (3.4) using three different methods. Firstly, we include dummy variables for each study as fixed effects to account for unobserved study-level characteristics, allowing us to examine within-study variation while ignoring other sources of variation. Secondly, we introduce weights that give equal weight to each study, correcting for the asymmetry that results from the varying number of estimates per study. These weights are based on the inverse of the number of estimates produced by each study. Finally, to control for potential endogeneity, we use an

[5]Please note that this procedure is equivalent to using the weighted least squares method, where the weights are the inverse of the estimate's variance.

instrumental variable approach. Stanley (2005) recommends using the inverse of the square root of the number of observations as an instrument for the standard error, as have many other researchers in the field (Zigraiova *et al.* 2021; Gechert *et al.* 2022; Cazachevici *et al.* 2020).

A more advanced group of non-linear techniques for detecting publication bias exists beyond PEESE. To test the robustness of our results, we employ a battery of these advanced tests, which includes five methods: the top 10 method by Stanley *et al.* (2010), the weighted average of adequately powered (WAAP) by Ioannidis *et al.* (2017b), the selection model by Andrews & Kasy (2019), the stem-based method by Furukawa (2019), and the kinked method by Bom & Rachinger (2019). The top 10 method and the WAAP method focus on estimates with sufficient statistical power. The former removes 90% of the least precise estimates, while the latter employs only those estimates whose statistical power exceeds 80%. The stem-based method builds on a trade-off between bias (squared) and variance, with the most precise estimates being the least biased and omitting estimates leading to an increase in variance. Thus, optimizing this trade-off may result in the desired "true" effect. The kinked method searches for the precision threshold above which publication bias is unlikely.[6]

Table 3.2 presents the estimation results, separately for each capital ratio and divided into three panels based on the estimation approach. Empirical tests support our intuition gained from visually inspecting the funnel plot. The two sub-samples of estimates on capital-to-asset ratio and regulatory capital ratio show a somewhat limited evidence of significant publication bias. The estimated effect beyond bias across different methods is very close to bias-uncorrected mean as reported in Table 3.1. One explanation consistent with this result is that there are generally no strong *a priori* views on the direction of the relationship in the simple capital-to-asset ratio and regulatory capital ratio categories. This only highlights the need to further examine the sources of heterogeneity in the estimated semi-elasticities, as we can rule out the problem of publication selection.

We find significant evidence of publication bias in the sample of semi-elasticities linked to capital requirements. This is not that surprising due to strong prior intuition on the potential effects of changes to capital regulation on bank lending. However, the bias estimated using FAT-PET-PEESE, albeit

[6]For further information on each of the non-linear techniques, please refer to sources such as Cazachevici *et al.* (2020) or Gechert *et al.* (2022).

economically meaningful, is not always statistically significant. This may be due to a relatively small sample size and the potentially non-linear relationship between the collected estimates and their standard errors. The latter might speak in favor of relying on advanced non-linear techniques which estimate the effect beyond bias, i.e. the mean effect corrected for publication bias, at between -0.5 pp and -0.8 pp. Estimates are statistically significant at 1% and are robust across the different methods employed. The uncorrected mean linked to studies using capital requirements is -1.7 which is significantly above the corrected mean, and it suggests that the effect of changing the capital requirements on bank lending might be systemically exaggerated. Nevertheless, given that this particular literature is relatively scarce due to data limitations, it is important to interpret our findings cautiously. It can be viewed as a potential area for future investigation when a greater number of studies concentrating on changes to capital requirements are available.

In meta-analysis, two other commonly used methods are the random effects model and the three-level model. The random effects model assumes that effect sizes vary both within and between studies, and uses estimates of within-study and between-study variance to obtain an overall effect size estimate. The three-level model extends the random effects model by adding a third level of analysis to account for variation in effect sizes due to differences in primary study characteristics, such as differences in data or methods. However, these methods assume that only one effect is reported per study, which is not our case. As mentioned above, we prefer to use other methods. However, it may be useful to report regression results based on the random effects and three-level models to explore sources of heterogeneity within and between studies.

In the online appendix, we present results from both methods, which indicate that the majority of heterogeneity in our data arises from between-study variability. This is particularly evident for the simple capital-to-asset ratio, although we also observe some within-study heterogeneity for the regulatory capital ratio and capital requirements. Overall, the level of heterogeneity in our data ranges from moderate to substantial (Higgins & Thompson 2002),[7] emphasizing the need for further investigation to identify the underlying drivers of heterogeneity.

[7]Higgins & Thompson (2002) provide a rule of thumb on how to interpret a popular statistics I^2 used in meta-analysis to measure heterogeneity. This statistic reports the percentage of variability in the effect sizes that is not caused by sampling error. The values around 50% are reported as moderate heterogeneity while values around 75% as substantial heterogeneity.

Table 3.2: Estimation of Publication Bias

	Capital-to-Asset Ratio	Regulatory Capital Ratio	Capital Req.
Panel A: FAT-PET			
Study-level fixed effects			
Effect beyond bias (1/SE)	0.365***	0.265*	-0.442***
	(0.141)	(0.159)	(0.089)
Publication bias (constant)	-0.049	-0.305	-1.112***
	(0.198)	(0.334)	(0.261)
Weighted least squares			
Effect beyond bias (1/SE)	0.202*	0.182*	-0.406*
	(0.099)	(0.1)	(0.152)
Publication bias (constant)	0.274	-0.205	-0.866***
	(0.286)	(0.591)	(0.133)
Instrumental variable approach			
Effect beyond bias (1/SE)	0.399***	0.424***	-0.895***
	(0.046)	(0.046)	(0.134)
Publication bias (constant)	-0.153	-1.350***	-0.033
	(0.157)	(0.328)	(0.337)
Panel B: PEESE			
Study-level fixed effects			
Effect beyond bias (1/SE)	0.353**	0.261	-0.408***
	(0.141)	(0.16)	(0.096)
Publication bias (SE)	-0.054**	-0.242	0.077
	(0.026)	(0.215)	(0.067)
Weighted least squares			
Effect beyond bias (1/SE)	0.240***	0.170**	-0.616***
	(0.081)	(0.077)	(0.127)
Publication bias (SE)	0.001	0.016	-0.080
	(0.001)	(0.009)	(0.068)
Instrumental variable approach			
Effect beyond bias (1/SE)	0.317***	0.237***	-0.683***
	(0.016)	(0.01)	(0.033)
Publication bias (SE)	0.014	-0.521***	-0.303***
	(0.041)	(0.194)	(0.055)
Panel C: Advanced methods			
Top 10 method (Stanley *et al.* 2010)			
Effect beyond bias	0.252***	0.221***	-0.608***
	(0.026)	(0.028)	(0.094)
WAAP (Ioannidis *et al.* 2017b)			
Effect beyond bias	0.263***	0.181***	-0.750***
	(0.037)	(0.023)	(0.076)
Stem-based method (Furukawa 2019)			
Effect beyond bias	0.196*	0.021	-0.651***
	(0.107)	(0.187)	(0.082)
Kinked method (Bom & Rachinger 2019)			
Effect beyond bias	0.287***	0.240***	-0.482***
	(0.023)	(0.013)	(0.043)
Observations	514	652	229
Studies	16	18	5
Observations per study (mean)	32	36	46

Note: Panel A & B: Standard errors, clustered at the study level, are reported in parentheses. For the weighted least squares, the inverse of the number of estimates reported per study is used as a weight. For the instrumental variable approach, the inverse of the square root of the number of observations is used as an instrument of the standard error. Panel C: the top 10 method by Stanley *et al.* (2010), the weighted average of adequately powered (WAAP) by Ioannidis *et al.* (2017b), the selection model by Andrews & Kasy (2019), the stem-based method by Furukawa (2019), and the kinked method by Bom & Rachinger (2019). The number of observations under the Top 10 and WAAP methods was reduced to 52 and 33 for the capital-to-asset ratio sub-sample, 66 and 101 for the regulatory capital ratio sub-sample, and 23 and 36 for the capital requirements sub-sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In the next section, we will conduct an analysis to identify the sources of heterogeneity and examine the presence of publication bias while controlling for all possible factors that contribute to heterogeneity.

3.5 Drivers of Heterogeneity

The empirical literature on the relationship between bank capital and lending shows a high degree of heterogeneity that goes beyond the different expressions of the bank capital ratio. While the impact of raising capital requirements on bank lending proves to be sizeable and negative, the effect of changes to the simple capital-to-asset ratio and the regulatory capital ratio remains unclear in terms of its direction and size, as we have shown earlier in Figure 3.1 and Figure 3.3. Publication bias was proved not to explain this large variation.

Meta-analytic studies usually have to deal with some heterogeneity of the collected estimates. In principle, the heterogeneity should be low if the subject of the meta-analysis is a deep (structural) parameter obtained from a model that correctly describes the data generating process. In such a case, some heterogeneity can be driven by an econometric approach to estimating the model. However, if the subject of the meta-analysis is a reduced-form parameter or the model does not correctly follow the data generating process, the heterogeneity of the collected estimates is expected to be high and driven by data characteristics and model specification. The latter is a case for modelling bank behavior, suggesting that model specification and data characteristics will play an important role in the derived effect.

In this section, we control for 40 variables to better understand the differences between studies. About three quarters of the variables come from the primary studies while the rest are structural (external) variables capturing cross-country or regional differences. These are usually collected from first-rate databases such as those of the World Bank, the OECD or Eurostat. We provide the description and summary statistics of all variables entering the analysis in the online appendix.

3.5.1 Estimation

Given the large number of control variables, we decided to employ model averaging techniques, both Bayesian and frequentist. These techniques allow for model uncertainty, which is often present in complex models with many potential explanatory variables. In contrast to OLS, which requires a specific model specification with a fixed set of explanatory variables, model averaging techniques considers a range of models and assigns a weight to each model based on how well it fits the data. This allows for the possibility that different models may be appropriate for different subsets of the data, and that the results may be sensitive to the choice of model specification. The Bayesian model averaging (BMA) allows us to estimate the probability that an individual explanatory variable would be included in the underlying model. In addition, the BMA also provides a way to incorporate prior information into the analysis, which can be particularly useful when the sample size is small, or the data are noisy. By combining prior information with the observed data, this technique can produce more robust and reliable estimates of the model parameters. The frequentist model averaging (FMA) then serves as a useful robustness check.

The goal of BMA is to find the best possible approximation of the distribution of each regression parameter. Our data can provide 2^{40} variable combinations to run as regressions which would be potentially very time-consuming. To cut the estimation time, we use a Markov chain Monte Carlo process with the Metropolis-Hastings algorithm which only goes through the most likely models (Zeugner & Feldkircher 2015). The probability of each model is then turned into the respective weight and the weight of each model is captured by a measure called posterior model probability (PMP). The estimated coefficients for each variable are equal to the weighted sum of the variable coefficients through all the models based on the PMP of each model. This estimated coefficients have assigned posterior inclusion probability (PIP) representing the sum of the posterior model probabilities through all the models, where the variable is included.

BMA requires explicit priors concerning the model (model prior) and regression coefficients (g-prior). Following Eicher *et al.* (2011), we use a combination of unit information g-prior (UIP) and uniform model prior as a baseline. This setting expresses our lack of knowledge regarding the particular probabilities of individual parameter values as the prior assigns the same weight to the regression coefficient of zero and to the observation in the data. As a robustness check, we analyze the sensitivity of our results to different prior choices. For instance, we should account for the fact that we employ a relatively high number of explanatory variables which may succumb to collinearity even though we discarded several of them upfront. Therefore, we also employ the dilution model prior proposed by George (2010), adjusting the model probabilities by the determinant of the correlation matrix of the particular variables included in the suggested model.[8] Further, we also employ a combination of the Hannan-Quinn (HQ) g-prior and random model prior (Fernandez *et al.* 2001; Ley & Steel 2009) and a combination of the BRIC g-prior and random model prior. The HQ g-prior adjusts data quality and is recommended, for instance, by Feldkircher & Zeugner (2012) or Zigraiova *et al.* (2021). The BRIC g-prior, which is widely used in literature, minimizes the prior effect on the results (Zeugner & Feldkircher 2015). The use of random model priors thus means that equal prior probability is given to every model size (Gechert *et al.* 2022). This way we show a lack of prior knowledge about the model's distribution. All robustness checks that focus on the different priors are reported in the online appendix.

In the following sub-section, we interpret only the BMA means with a posterior inclusion probability (PIP) of above 0.5, following the approach proposed by Jeffreys (1961) and Havranek *et al.* (2021). They further divide the interpretation into the following groups: the effect is deemed 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 decisive if the PIP is greater than 0.99. We use the Bayesian approach as our baseline estimation technique, and the frequentist approach as a robustness check. For the "frequentist check", we run a simple OLS regression with clustered standard errors, including only variables with a PIP of above 0.5. While the latter approach may be controversial because it does not incorporate the uncertainty from estimating models in previous steps into the standard errors, recent papers such as (Gechert *et al.* 2022) and (Bajzik *et al.* 2020) also use this procedure. We believe that this simple exercise can add to the overall picture. The adjusted R-squared, which captures the portion of variance of the collected estimates explained by the chosen characteristics, ranges from 50% to 70%. The numerical results of both the FMA and frequentist check are presented in the online appendix.

3.5.2 Results

The numerical results of BMA for all three capital ratios are shown in Table 3.3. The corresponding graphical output is reported in the online appendix. We first consider two samples of studies split on whether they consider the simple capital-to-asset ratio or regulatory capital ratio. The third sample of stud-

[8]In case of high correlation, the determinant is close to one and the model receives little weight and vice versa. This prior was used in meta-analysis for instance by Bajzik *et al.* (2020).

ies, which uses capital requirements, is described as a special case bearing in mind the lower number of semi-elasticities available and the relatively uniform estimation approach employed in primary studies.

The BMA results confirm the presence of publication bias in our third sample for the relationship between capital requirements and lending, even after controlling for additional variables in a full meta-regression. The posterior inclusion probability of the standard error is 0.96, and the slope coefficient is around -0.25, consistent with the lower end of the range provided by univariate regressions in section 3.4. In contrast, there are no signs of publication bias for the other two capital ratios, as confirmed by the posterior inclusion probability of the standard error being below 0.1. This finding is in line with our univariate regression results as well. Additionally, we find that approximately half of the additional characteristics collected from primary studies or external databases are significant drivers of heterogeneity, with the signs of these variables being robust across specifications. We will discuss more each of these drivers in the following sub-sections.

Table 3.3: What Drives the Heterogeneity of Collected Estimates – Bayesian Model Averaging

	Capital-to-asset ratio			Regulatory capital ratio			Capital requirements		
	P. mean	P. SD	PIP	P. mean	P. SD	PIP	P. mean	P. SD	PIP
Constant	-3.555	-	1.000	-2.947	-	1.000	-1.936	-	1.000
St. error	0.002	0.008	0.087	0.001	0.011	0.068	-0.249	0.095	0.959
Data characteristics									
No. of observations	0.262	0.045	1.000	0.245	0.024	1.000	0.199	0.080	0.946
Confidential data	6.269	0.994	1.000	-2.047	0.689	0.942			
Other region	5.312	1.091	0.999	0.538	0.508	0.663			
US	-2.910	1.257	0.962	-1.323	0.702	0.923			
Single-country	3.811	1.298	0.951	2.133	0.779	0.962			
Midpoint	-3.561	1.967	0.917	0.219	0.591	0.194	-0.074	0.178	0.207
Corporate credit	0.514	0.316	0.751	-0.156	0.150	0.691			
Household credit	-0.182	0.306	0.298	-0.020	0.113	0.186	3.596	0.595	1.000
Quarterly frequency	-0.562	1.108	0.259	2.971	0.685	0.999			
Macro-level data	-0.001	0.045	0.047	0.000	0.031	0.054			
Total credit							2.649	0.864	0.976
Model specification and estimation									
Missing control variables	3.160	0.438	1.000	0.053	0.175	0.128			
Missing interest rate in eq.	-2.574	0.463	1.000	0.467	0.186	0.916			
Other method	-7.692	2.424	0.979	-1.158	0.401	0.994			
Fixed-effects method	-1.966	0.907	0.918	-0.747	0.378	0.852			
Lagged by 1Y or more	2.127	1.131	0.826	0.481	0.298	0.794	-0.024	0.118	0.092
Contemporaneous	-0.654	1.257	0.269	0.017	0.121	0.076			
Add. lag in eq.	-0.177	1.356	0.171	2.079	0.749	0.956	0.037	0.242	0.088
Time fixed effects incl.	0.024	0.078	0.129	0.006	0.031	0.069			
Add. capital in eq.	-0.021	0.098	0.080	0.960	0.276	0.992	-0.022	0.299	0.067
Dynamic model	0.019	0.112	0.068	-0.760	0.432	0.824			
Other interaction	0.003	0.032	0.049	-0.001	0.019	0.046			
Some interaction in eq.	0.002	0.044	0.045	-0.001	0.023	0.042	-0.352	0.412	0.499
GMM method	0.002	0.038	0.045	-0.011	0.049	0.082			
Crisis	0.002	0.037	0.044	0.006	0.033	0.071			
Publication characteristics									
Published	3.316	0.771	1.000	-0.006	0.091	0.095			
Impact factor	-1.454	0.235	1.000	-0.513	0.111	0.999			
Citations	-1.684	0.621	0.989	-0.051	0.176	0.126			
Central bank publication	2.574	0.785	0.972	-0.424	0.630	0.419			
Publication year	-0.963	0.823	0.691	-0.264	0.292	0.525	-0.304	0.634	0.247
External variables									
Ext: inflation deviation	0.336	0.044	1.000	0.007	0.020	0.219			
Ext: low for long	-0.460	0.066	1.000	-0.031	0.054	0.378	-0.169	0.036	0.998
Ext: fin. openness	6.499	1.253	1.000	0.265	0.339	0.534			
Ext: bank size	-0.004	0.015	0.179	-0.003	0.007	0.218			
Ext: 3M interest rate	-0.040	0.208	0.094	0.024	0.080	0.177			
Ext: spread	0.046	0.271	0.090	0.008	0.099	0.168			
Ext: unemployment	-0.004	0.075	0.080	0.062	0.085	0.442			
Ext: house price growth	0.001	0.009	0.073	-0.001	0.005	0.118			

Note: The table presents the estimation results of a collected estimate of the beta coefficient on the primary study characteristics and external (structural) variables, searching for potential sources of heterogeneity. Bayesian model averaging employs a combination of the uniform model prior and the unit information g-prior recommended by Eicher *et al.* (2011). P. mean – posterior mean, P. SD – posterior standard deviation, PIP – posterior inclusion probability. The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between the various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures. The set of characteristics for the subset of semi-elasticities on capital requirements is reduced due to little heterogeneity in the studies and the multicollinearity of some variables.

Capital-to-Asset Ratio

The sample of studies using a simple capital-to-asset ratio to explain changes in bank lending has a mean semi-elasticity of 0.35, with a significant variation between the top and bottom 5% reported (-2.22 and 3.79 respectively). Such studies generally aim to capture the effects of the bank capital position or the level of bank capitalization. As such, prior intuition and the true effect identified in the previous sections of our paper is that higher bank capitalization should inflate lending.

Data characteristics. We find sample size to be one of the most important factors influencing the estimates of the relationship between bank capital and lending. Semi-elasticities estimated using larger datasets tend to be higher which may be the result of higher in-sample variation captured by more observations or by a change in the relationship over time. The latter is supported by the fact that studies performed on more recent datasets, as captured by the midpoint of the estimation period, exhibit a less positive or indeed even a negative effect of bank capital on lending. The shift in the estimated relationship over time may, to some extent, reflect increasingly demanding (and binding) capital regulation, especially in Europe and the US. The decisively positive effect of studies conducted on countries outside Europe, together with the substantially negative effect of studies performed solely on US data, speak in favor of this interpretation. Some structural characteristics give additional credit to this hypothesis (see below).

More generally, single-country studies deliver higher positive semi-elasticity estimates, even more so when using confidential data sources.[9] Due to more detailed data, these kinds of studies can theoretically be more successful than others in correctly identifying shocks to bank capital that are plausibly unrelated to lending opportunities. As such, they are less likely to suffer from endogeneity bias.[10]

We further find substantial evidence that primary studies focusing on the impact of changes in the bank capital position on corporate credit produce more positive semi-elasticity estimates than those focusing on the impact on total credit (or household credit). This suggests that corporate credit is far more

[9] In our sample, studies employing confidential data are exclusively single-country studies and are generally performed by central bank researchers.

[10] For example, lending growth and capital can be endogenously determined through the performance of borrowers. Confidential loan-level data may allow for the inclusion of borrower characteristics that would significantly reduce potential bias.

sensitive to changes to a bank's capital position than credit extended to other economic sectors. This echoes back to the literature which shows that when faced with a change in capital requirements, banks will make trade-offs between different assets and between the different options of how to move to a higher capital ratio. Not surprisingly, banks will often choose to reduce their high-risk exposures (Akram 2014; Mendicino *et al.* 2018). Empirical studies find that banks tend to shrink their portfolios of corporate loans, which generally attract a higher risk weight, more than their portfolios of domestic loans (Bridges *et al.* 2014; Bahaj *et al.* 2016).

Model specification and estimation. Another important way in which estimates differ is the estimation technique used. Studies that try to shield the estimates from the omitted variable bias by using a fixed effect estimator tend to report lower semi-elasticity values than those using simple OLS. Similarly, studies that include both bank-specific and macro-economic variables in their specification report lower semi-elasticities than those missing some controls. The one exception is interest rates which, if missing in the model, translate into a generally weaker relationship between capital and lending. In our stream of literature, the omitted variable can take the form of various unobserved managerial or regulatory decisions that if negatively correlated with the bank capital ratio would downward-bias the estimated semi-elasticities. Overall, it seems that studies that try to address these issues report less positive or negative estimates which would mean the endogeneity bias would generally work upwards for this particular relationship between capital and lending. We also find that studies which consider a short-term relationship (up to one year) between bank capital and lending report lower semi-elasticity than studies considering a capital ratio lagged by one year or more. This may imply that changes to bank capital affect lending more negatively over a shorter horizon and more positively over a longer horizon. This links to the literature on the adjustment mechanism of banks' capital ratio to various shocks, e.g. a regulatory tightening (Memmel & Raupach 2010; Jokipii & Milne 2011) and the Covid-19 pandemic (Couaillier *et al.* 2022; Mathur *et al.* 2023). Studies show that following a regulatory tightening, low-capitalized banks respond dramatically to make sure they maintain their capital ratio above the regulatory minimum.

Publication characteristics. Our results indicate a strong association between four publication characteristics (year of publication, publication in a peer-reviewed journal, journal impact factor and the number of citations) and the reported results. We interpret this association as the potential effect of quality: while studies published in refereed journals tend to report generally more positive bank capital-lending semi-elasticities, more recent and higher-quality studies (those with a higher impact factor and more citations) tend to report decisively fewer positive or potentially even negative semi-elasticities. This means that the effect may be overestimated in the lower-quality series, while the negative association between the year of publication and the estimated semi-elasticity lends support to our hypothesis of the changing relationship over time. Interestingly, working papers published by central banks tend to report more positive semi-elasticities.

External variables. We identified three external variables which play a decisive role in the relationship between bank capital and lending. First, the positive relationship is stronger in countries with worsened economic conditions, which is captured by high or unstable inflation. A negative shock to bank capital during less tranquil times in countries with unfavorable macroeconomic conditions could significantly decrease credit. Second, a prolonged period of low interest rates, as captured by the “low for long” variable, weakens the positive effect or potentially reverses it. The variable refers mostly to the period after the GFC when bank profitability was subdued and capital regulation was tightened. For many banks, capital requirements became binding as their capital buffers (capital above minimum capital requirements) became depleted. Therefore, additional capital requirements may require banks to increase their capital position without extending additional credit to the economy or may even depress lending. The “low for long” variable could capture the change in the relationship between bank capital and lending after the GFC, which may have become weaker or even negative. Finally, we find that the degree of financial openness of a given country considered in the analysis is an important factor in explaining heterogeneity. Studies performed on a more financially open country tend to report more positive semi-elasticities. A more financially open banking sector is more prone to international spillovers of various shocks which foster the dependence of bank lending decisions on the level of capitalization.

Regulatory Capital Ratio

The sample of studies using a regulatory capital ratio to explain movements in bank lending has a mean semi-elasticity of 0.14 with the variation between the top and bottom 5% reported more in favour of the negative spectrum of the semi-elasticity distribution (-1.38 and 1.11 respectively). Studies considering the risk-sensitive capital ratio are generally interested in the estimation of the relationship between bank capital and lending under the umbrella of bank capital regulation. While the distribution of semi-elasticities on the simple capital-to-asset ratio is skewed to the right (i.e. more towards positive values), the distribution of semi-elasticities on the regulatory capital ratio is skewed to the left (i.e. more towards negative values).

Contrary to the risk insensitive capital-to-asset ratio, prior intuition linked to explaining a shock to regulatory capital ratio is rather fuzzy. Studies considered in our sample tend to understand such a shock in different ways ranging from a plain shock to bank capitalization (Brei *et al.* 2013) to the imposition of a regulatory tax (Naceur *et al.* 2018) to a direct proxy for macroprudential policy (Wang & Sun 2013). The absence of clear prior intuition is actually a good motivation for looking at the factors explaining the heterogeneity of reported semi-elasticities. It could be that different researchers have different priors on the analyzed effect and this may be reflected in their choice of the data, model and estimation technique.

Since the simple capital-to-assets ratio and regulatory capital ratio share some commonalities in terms of calculation, it is not surprising that we also find several common factors that explain the heterogeneity for both. Specifically, single-country studies with a larger sample size are positively correlated with the reported semi-elasticities in both sub-samples. Similarly, studies shielded from omitted variable bias with more favorable publication characteristics are negatively correlated with the reported semi-elasticities in both sub-samples. However, contrary to the sub-sample of the simple capital-to-asset ratio, we find a much smaller role for external factors.

Data characteristics. On top of the factors described above, we find that larger data frequency is positively correlated with the estimated semi-elasticities: semi-elasticities estimated using datasets with annual frequency tend to be smaller than those employing quarterly or monthly data for estimation. We further find that US-based studies tend to report substantially smaller semi-elasticities than other regions, even more so when using confidential data sources. Contrary to the previous subset of data, the effect of a higher regulatory capital ratio on lending is less positive or even negative if corporate loans are considered.

Model specification and estimation. We found that a number of characteristics were strongly associated with the model specification and estimation. First, the reliance on a dynamic model specification tends to shift the distribution of semi-elasticity estimates to negative territory. Not surprisingly, including the lagged dependent variable may explain a lot of the variance in the credit dynamics initially captured by the capital variable. Given the generally high persistence in the stock of credit, not including the persistence term may significantly overestimate the reported semi-elasticity linked to the capital ratio. In addition, a rich lag structure and an additional capital variable included in the same model shifts the semi-elasticity towards more positive values. This effect can be viewed from two angles. The inclusion of additional lags of the reported capital ratio or different capital variables may absorb the variation in credit dynamics, and potentially act in the opposite direction as the capital ratio primarily studied in this sub-section. In other words, the positive correlation may be a result of multicollinearity in the primary study. On the other hand, assuming that the researchers checked and corrected for potential variable multicollinearity, a richer lag structure and additional capital variable may allow the model to account for period-to-period persistence in the level of capital, e.g. adjustments banks may make in advance of planned changes in their balance sheets.

Finally, we find that the inclusion of interest rates in a model is detrimental to the reported semi-elasticity. Studies that are missing interest rates from their model specification report have more positive semi-elasticities. Interest rates are generally meant to capture changes to monetary policy which can directly affect bank lending and bank capital. Numerous studies show that monetary policy tightening depresses lending via the bank lending channel (Kishan & Opiela 2000; Disyatat 2011; Albrizio *et al.* 2020). At the same time, higher interest rates improve the profitability of a bank which can increase its capital via retained earnings (Borio *et al.* 2017; Altavilla *et al.* 2018). This would suggest that the omitted-variable bias associated with not including the interest rate into the model works upwards.

Capital requirements

Studies considering changes to bank capital requirements offer the most precise approximation of changes to bank capital due to a change in capital regulation. Unfortunately, owing to a lack of hindsight and detailed supervisory data, few such studies exist, giving us only a handful of semi-elasticity estimates (Table A2 in the online appendix). Moreover, the studies show little heterogeneity in terms of data characteristics and econometric approach, significantly reducing the set of control variables entering the analysis. BMA estimates thus need to be interpreted with care as they are meant to serve as a first attempt to perform a heterogeneity analysis in this field of literature. Changes to capital requirements, if binding, are expected to decrease bank lending due to the immediate effect on bank funding costs. If bank capital is costly, reliance on such funding can lead to a decrease in the supply of credit. Prior intuition thus strongly favors a negative effect of rising capital requirements on bank lending.

The presence of publication bias in the estimates of the effect of capital requirements is supported by evidence across all the models we run. Thus, the reported semi-elasticities are found to be systematically exaggerated due to publication bias even if we control for the additional characteristics of the individual studies. The bias is found to work downwards, therefore studies tend to favor more negative semi-elasticity estimates which are in line with economic intuition.

We further find that studies based on smaller samples report more negative semi-elasticities. This association may simply reflect that studies performed on more recent datasets, that are naturally shorter, will identify a stronger negative effect of capital regulation, since changes to bank capital requirements took place predominately after the GFC. A negative correlation with the "low for long" variable speaks in the favor of this interpretation. Similarly to the previous two capital to lending relationships, we found that the type of bank credit entering the analysis played a significant role. The recurring theme is a stronger reaction of corporate credit to changes in capital ratios, which may be either be more positive (changes in the simple capital-to-asset ratio) or more negative (changes in the regulatory capital ratio or capital requirements).

3.5.3 Economic Significance of Key Variables and Implied Semi-Elasticity

The model averaging analysis indicates that the reported semi-elasticities for the subsets related to the simple capital-to-asset ratio and the regulatory capital ratio may be biased due to the researchers' choice of empirical approach and the changing nature of the relationship over time. Therefore, we compute the mean semi-elasticity based on our preferred best practice methodology employed in the literature. With that, we attempt to show what the mean semi-elasticity would be if all the studies used the same strategy as the one that we prefer. We calculate the mean semi-elasticity as fitted values directly from the BMA output using variables that the BMA analysis deems important (PIP above 0.5).[11] As the characteristics used to derive the best-practice specification explain up to 70% of the variance in collected estimates, we deem this exercise robust and informative.[12]

Our preference is for studies that utilize consistent and unbiased estimators, which can effectively capture changes in bank capital arising from capital regulation. We find that more recent single-country studies, which employ confidential data samples with higher frequency and more observations, tend to align with our subjective definition of best practice. Additionally, we favor dynamic models that include both bank-level and macroeconomic control variables, and which are estimated using the fixed effect regression method. Furthermore, we prioritize more recent publications in refereed journals with higher impact factors and citation rates. We also include external variables selected by the BMA approach and differentiate between corporate and household credit under the best practice approach.

Since our best practice is subjective, we show how the mean semi-elasticity deviates if we alter some of the key variables. First, we calculate mean semi-elasticities using an inferior empirical approach that misses some key control variables, includes the additional capital variable or lag of respective capital ratio, and is estimated by simple OLS. Second, we change the value of the "low for long variable" to its 90th percentile to show the full manifestation of a prolonged period of low interest rates on the relationship between bank capital and lending. Table 3.4 presents results for all three subsets. Since the studies on capital requirements are methodologically homogeneous, the significant heterogeneity drivers are limited, and thus we cannot provide implicit semi-elasticity estimates for an inferior empirical approach.

In addition to the implied semi-elasticity, we also estimate the mean effect beyond bias by utilizing all available information and correcting for publication bias, as shown the first row of Table 3.4. The effect beyond bias is calculated using the complete meta-regression output and all control variables, except for the slope coefficient on the standard error. This means that the effect beyond

[11] We assign variables with a PIP below 0.5 and non-preferred variables a value of zero, and then we use the *predict* function provided by the *BMS* package in R to compute fitted values based on the MCMC frequencies of all models. Confidence intervals are obtained from the predictive densities of the fitted values using the *pred.density* function from the same package.

[12] In the online appendix, we present frequentist checks that control for the variables used to calculate implicit mean semi-elasticities. The adjusted R^2 is reported as the goodness-of-fit measure for these specification.

bias calculated from the multivariate regression is an extension of the effect beyond bias estimated in section 3.4 in a univariate context.[13] Comparing the multivariate and univariate estimates, we find that they are similar, especially for the subsets related to the simple capital-to-asset ratio and the regulatory capital ratio. However, the effect is somewhat stronger for the third sample on capital requirements, which reflects a lower estimated intensity of publication bias, as noted at the beginning of sub-section 3.5.2.

Table 3.4: Mean Semi-Elasticities Implied by Significant Heterogeneity Drivers

	(1) Capital-to-Asset Ratio		(2) Regulatory Capital Ratio		(3) Capital Requirements	
	Estim.	68% CI	Estim.	68% CI	Estim.	68% CI
Effect beyond bias	0.34	(-0.03, 0.72)	0.14	(-0.12, 0.40)	-1.36	(-2.27, -0.45)
Best practice	0.43	(-0.81, 1.08)	-0.62	(-0.75, 0.22)	-0.72	(-1.91, 0.27)
Corporate credit	0.61	(-0.60, 1.24)	-0.65	(-0.79, 0.19)	-0.97	(-2.16, 0.02)
Household credit	0.39	(-0.84, 1.01)	-0.62	(-0.76, 0.22)	-0.55	(-1.73, 0.44)
Inferior empirical app.	0.26	(-0.98, 0.91)	0.52	(0.51, 1.50)	-	-
Prolonged per. of low ir*	-2.42	(-3.68, -1.83)	-1.08	(-1.20, -0.24)	-3.55	(-4.68, -2.53)

Note: The table presents the mean estimate of the semi-elasticity of the relationship between bank capital and lending implied by the BMA, the collected estimates and our definition of best practice in the first row, as well as changes to key variables in the remaining rows. That is, the table attempts to show what the mean semi-elasticity would be if all studies used the same strategy as the one that we prefer (best practice). In addition, we attempt to show the economic significance of the key variables by calculating the mean semi-elasticity for different sub-groups of characteristics. The mean semi-elasticity is calculated as the fitted values based on the BMA output and a matrix of chosen study characteristics (only variables with a PIP of above 0.5 are considered). We report 68% confidence intervals in brackets which are retrieved from the predictive densities of the fitted values. The predictive density is a mixture density based on the best models identified by the BMA. *We replaced the values of the low for long variable with its 90th percentile to see the full manifestation of this effect. The low for long variable was deemed decisive (PIP > 0.99) in the subset of capital-to-asset ratios and capital requirements but unimportant (PIP = 0.38) in the subset of regulatory capital ratios; thus, estimates for the latter should be interpreted with caution.

The best practice estimate of the relationship between the capital-to-asset ratio and bank credit growth is 0.43 pp (column 1) while that of the relationship between the regulatory capital ratio and bank credit growth is -0.62 pp (column 2). The latter is in stark contrast to both the simple average and the mean effect corrected for publication bias calculated across collected semi-elasticities (between 0.1 and 0.3 pp). Generally, semi-elasticities implied by significant heterogeneity drivers are distinctly on the positive side for the simple capital-to-asset ratio and on the negative side for the regulatory capital ratio. This supports our view that the relationship between the regulatory capital ratio and bank lending reflects the impact of changes to bank capital

[13]To calculate the effect beyond bias from a multivariate regression output, we employ the identical method as for the calculation of implied semi-elasticity. This means, that we use the *predict* and *pred.density* functions, which are provided by the BMS package in R.

regulation whereas the simple capital-to-asset ratio reflects changes to bank capitalization. Part of the regulatory policy is contained in the denominator of the regulatory capital ratio. Unlike the simple capital-to-asset ratio, changes to the regulatory capital ratio can be caused by changes in regulation (i.e. the method of calculation of risk-weighted exposures) and the riskiness of the underlying exposures. A slightly stronger reaction, both positive and negative, was derived for corporate credit relative to household credit. On the contrary, the inferior empirical approach brings both semi-elasticities slightly closer to zero, suggesting that the correct model specification and estimation technique is key in identifying the true effect.

The best practice estimate of a 1 pp increase in capital requirements slows down bank annual credit growth, on average, by 0.72 pp (column 3). Similarly to the previous two ratios, the effect is stronger for corporate credit relative to household credit. Given the limited heterogeneity in the approach to estimating this effect, the implied semi-elasticity of the relationship between capital requirements and bank lending serves merely as a robustness check for our results on publication bias. Reassuringly, the best practice semi-elasticity stays close to the estimates of the effect beyond bias (Table 3.2), which is substantially above the uncorrected mean of -1.7 pp. Importantly, the effect is very close to that of the regulatory capital ratio, supporting our intuition presented above that the relationship between the regulatory capital ratio and bank lending reflects the impact of changes to bank capital regulation. This may be relevant for at least two reasons. First, it is in line with the literature highlighting the role of bank capitalization and the risk sensitivity of capital regulation in the transmission of capital requirements. Second, it allows for proxy changes in capital regulation using the regulatory capital ratio if the data on capital requirements are not available to the researcher. In such a case, the choice of a suitable model specification and estimation approach would be crucial.

Interestingly, implied semi-elasticities for all three capital ratios are brought down by a period of low interest rates. The relationship between bank capital and lending seems to be significantly affected by the post-crisis period of highly accommodative monetary policy and more demanding capital regulation. Even risk-insensitive bank capitalization is no longer positively associated with bank lending. Statistically, this can be driven by the fact that the period of low interest rates coincides with substantially increased macroprudential policy activity (Alam *et al.* 2019). Record low interest rates have depressed bank profitability and may have encouraged risk taking and hence increase the probability of default (for a literature review, see Malovaná *et al.* 2022). The associated risks were aimed to be contained by increased engagement of macroprudential policy. Thus, capital regulation could have become the main driver of changes in banks' capital ratios.

3.6 Concluding Remarks

We present the first quantitative synthesis of the vast empirical literature on the effects of changes to bank capital on lending. Bank capital is an important factor for both stakeholders and policymakers because it affects the stability of the financial system and the availability of credit. Yet, despite its significance, the existing literature displays considerable fragmentation, with estimates of the relationship ranging from positive to negative values. Our dataset, consisting of over 1,600 estimates from 46 studies, allows us to identify sources of fragmentation using state-of-the art meta-analytic techniques as well as Bayesian and frequentist model averaging methods.

We provide several key findings. First, the researcher's choice on how to express the bank capital ratio has an important impact on the estimated effect of changes to bank capital on lending. Semi-elasticities based on simple capital-to-asset ratio and regulatory capital ratio tend to be positive, while those based on capital requirements are strongly negative. Taking a simple average of the collected estimates, a one percentage point increase in the simple capital-to-asset ratio or regulatory capital ratio leads to about 0.3 pp and 0.1 pp increase in annual credit growth, respectively. On the contrary, a one percentage point increase in minimum capital requirements results in a 1.7 pp decrease in lending. Correcting for publication bias reduces the latter effect to around -0.7 pp, with the range of corrected estimates being between -0.4 pp and -0.9 pp. Notably, the first two groups of semi-elasticities exhibit little or no signs of publication bias.

Second, we investigate whether some characteristics of the primary studies or external factors drive the observed heterogeneity. Since the simple capital-to-assets ratio and regulatory capital ratio share some commonalities, it is not surprising that we also identify several common factors that explain the heterogeneity for both, especially those linked to model specification and estimation. On the contrary, external factors such as macro-financial and institutional characteristics only play a prominent role in the relationship with the capital-to-assets ratio.

Interestingly, the relationship between bank capital and lending appears to have changed after the GFC, as evidenced by the weakening and potential reversal of the positive effect during a prolonged period of low interest rates. This period was characterized by increasingly demanding bank capital regulation and subdued bank profitability, which may have impacted the relationship. Banks may have struggled to maintain voluntary capital buffers, and any additional requirements may become binding, limiting their ability to lend. Furthermore, studies conducted in Europe and the US show a more negative effect, likely due to more stringent capital regulation in these regions.

Due to the high variability of the estimated coefficients, the regulatory capital ratio is a rather noisy indicator of changes in capital regulation and can thus capture both changes in the bank's total capitalization and the capital requirements. However, considering key heterogeneity drivers identified by model averaging techniques allow us to put more emphasis on the regulation-induced changes to capital ratio. If we consider single-country studies that use confidential data with higher frequency and a superior empirical approach, the mean semi-elasticity turns negative. This highlights the strong influence of researchers' choice of empirical approach on the final estimate of semi-elasticity. Nevertheless, researchers aiming to estimate the effect of changes in capital regulation are still better off considering changes directly to capital requirements.

Therefore, our study does not yield typical policy recommendations but rather provides "instructions" on how to conduct future policy-relevant empirical research, specifically how to estimate the effect of capital regulation on bank lending accurately.

Chapter 4

Trading volume and stock returns: A meta-analysis

Josef Bajzik

Abstract I examine 468 estimates on the relationship between trading volume and stock returns reported in 44 studies. I study publication bias together with Bayesian and frequentist model averaging to explain the heterogeneity in the estimates. The results yield three key conclusions. First, publication bias distorts the findings of the primary studies. Second, the predictability of stock returns varies with different markets and stock types. Third, different data characteristics, structural variations and methodologies used drive the heterogeneity in the results of the primary articles. In particular, one should be cautious when using monthly data or VAR models.

This paper was published in *International Review of Financial Analysis [2021, 78, pp. 1-20, DOI: 10.1016/j.irfa.2021.101923]*. The study was supported by the Charles University, projects GA UK No. 297521 (writing, submission fees) and No. 1430220 (proofreading) and by Donatio Universitatis Carolinae (project start). The author notes that the paper represents his own views and not necessarily those of the Czech National Bank.

4.1 Introduction

Does the trading volume affects the stock returns or not? This question has attracted traders at least for decades. And since many new traders enter the market via on-line platforms at least during the COVID crisis this question is still in the forefront. For example, Chiah & Zhong (2020) find a large spike in trading volume in 37 international equity markets during the COVID.

Studying return-volume relationship is really interesting both theoretically and practically. Beaver (1968) said: "An important distinction between the price and volume tests is that the former reflects changes in the expectations of the market as a whole while the latter reflects changes in the expectations of individual investor." Morgan (1976) continues with suggestion that volume is connected with systematic risk and thus with stock returns. Hence the trading volume falls among the possible determinants of the stock returns. That the trading volume affects the stock returns suggest even Brennan *et al.* (1998). Thus, the trading volume comes to the discussion of the stock returns beside the well-known factors proposed by Fama & French (1992; 1993; 1996) and Jegadeesh & Titman (1993; 1995).

Moreover, Karpoff (1987) added several other reasons to study the trading-volume effect. First, this type of research provides insight into financial markets' structure. Second, it is seminal for event studies, that use a price and volume data to draw conclusions. Third, the return-volume relationship has significant implications for futures markets researches. These suggestions make the findings related to this topic even more valuable.

The first studies discussing the price-volume relationship originated in the US in the 1960s (Granger & Morgenstern 1963; Godfrey *et al.* 1964). The focus on the US continued for the next decades (e.g., Crouch 1970; Jain & Joh 1988). By the turn of the millennium, researchers from every continent had started to show interest in the topic. For example, two decades ago, Lo & Wang (2000) found almost two hundred articles related to trading volume. The articles were from various fields – economics, finance, and accounting. Furthermore, during the last two decades, many more articles have been published. Thus, although the Fama & French (1992) do not mention the trading volume as a factor for stock price determination, this conclusion have come into question again with the increased number of articles.

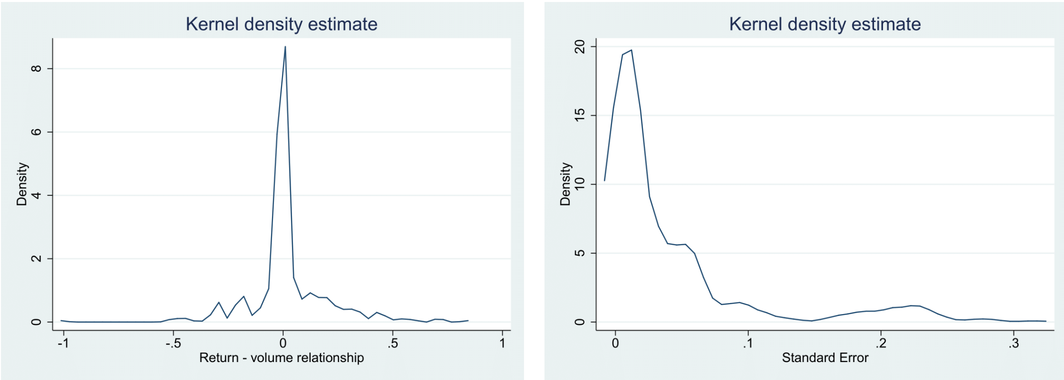
For instance, Lee & Rui (2002) and Gurgul *et al.* (2007) find a small or even negligible relationship between trading volume and stock returns. On the other hand, Brennan *et al.* (1998) and Chordia *et al.* (2001) contradict this conclusion. They were not alone in doing so: Mahajan & Singh (2009a) and Akpansung & Gidigbi (2015) provided overviews of the currently available literature related to the topic. While they pointed out reasons for the major differences in the existing literature, no consensus about the magnitude of the effect emerged. Thus it is important to fill the research gap and shed light on this area.

I decide to thoroughly investigate the trading volume-stock return relationship through a meta-analysis. Meta-analysis is quantitative review of all empirical results related to the given topic. It is an effective way to draw the real inferences from the various findings from the primary studies, which are often contradicting. Moreover, meta-analysis scrutinize systematically the differences in empirical findings. Since the meta-analysis works with a large number of independent variables, it addresses the endogeneity problems from which the primary study may suffer. Besides it compares the results also qualitatively. Next, it identifies study-invariant factors such a sample selection bias and measurement errors that may affected individual studies. It reveals the degree of influence of factors such as time variations and cross-country variations across the primary studies. Last, but not least, it explains the impact of economic fundamentals such as degree of market development on the market efficiency (Kim *et al.* 2019).

Altogether, I collect 468 estimates from 44 studies and 49 variables capturing the context in which the studies derived their findings. Figure 4.1 and Figure 4.2 provide an overview of the literature on trading volume and stock returns.

Three observations can be drawn from these figures. First, the median value of the estimated relationship approaches zero, and most of the estimates appear close to this value. Second, the reported values seem to decline over time. It may be caused simple by the so-called “Prometheus effect”, when the empirical effects decline over time after initial novel findings (Ioannidis 2008). On the other hand, the usage of the improved data and sophisticated techniques in the recent studies may induce it. The decline may spring by other essential changes as well. The third and last observation refers to the increasing variance of the estimates. Instead of converging toward some consensus, the estimates from the literature diverge over time. Again, it is unclear, whether it is caused by the more precise techniques used by the recent studies, or whether the new approaches bring the variance. The variance may be caused, for instance, by studying the relationship in the diverse countries and continents in the recent

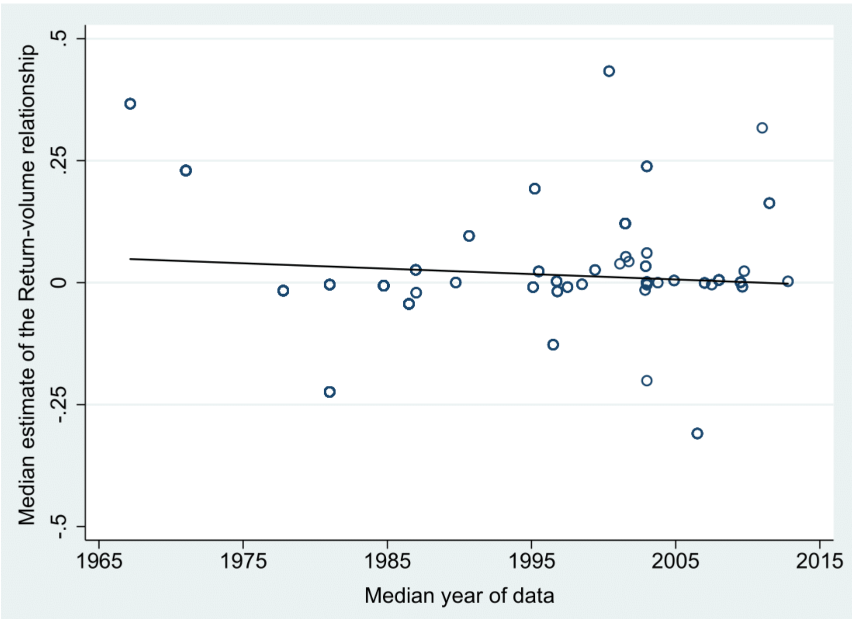
Figure 4.1: Kernel densities of the return-volume relationship and corresponding standard errors



(a) (b)

Notes: The Figure 4.1 depicts kernel densities for return-volume relationship (on the left) and corresponding Standard Errors (on the right). Since primary studies employ many different estimation approaches, partial correlation coefficients normalize the estimates and the winsorization handles with the outliers.

Figure 4.2: Decrease in mean and increase in variance of return-volume estimates over time



Notes: The Figure 4.2 captures median estimates per study of return-volume relationship at the vertical axis and measures the median year of the data used in particular studies at the horizontal axis.

years.

The last two observations about the declining trend and the increasing variance provide additional reasons for conducting a systematic assessment of all published results. The most suitable method for such an evaluation is a meta-analysis (Imai *et al.* 2021). The main contribution of such an analysis consists mainly in explaining the differences in the estimates over time, market data used and estimation techniques employed by the primary studies.

A meta-analysis addresses publication bias as well as model uncertainty issues. I follow seminal works such as Havranek & Irsova (2017) and employ the most modern techniques for correcting for publication bias together with Bayesian model averaging (Raftery *et al.* 1997) and frequentist model averaging (Amini & Parmeter 2012). To correct for publication bias, I start with the graphical visualization proposed by Egger *et al.* (1997a). Then, I add simple formal tests using ordinary least squares (OLS), the between effect and weighted least squares (WLS) (Stanley & Doucouliagos 2012). Furthermore, an extension of the formal tests is provided by means of the latest improvement suggested by Bom & Rachinger (2019). Moreover, newly developed stem-based method (Furukawa 2019) complement the investigation. Finally, the presence of publication bias at least in contemporaneous cases, is identified. The mean after correction for publication bias has a negligible value.

Moreover, other study-specific aspects affect the corrected mean. The results of both Bayesian model averaging (BMA) and frequentist model averaging (FMA) indicate that data characteristics, structural variation and different methodological approaches explain a large part of the inconsistency in the primary results. For example, usage of *Monthly* data or *VAR* models makes the effect of trading volume on returns substantially more negative.

Other causes of variation are the type of stocks and country of origin. For instance, this analysis reveals that the trading volume may predict the stocks of firms. On the other hand, the effects of trading volume on the stocks of the banks or overall indexes turn to be insignificant. This conclusions I find even across the continents. The same holds for markets in the North America, Europe, Australia and Asia. Only estimates for developing countries differ. Trading volume predicts the stock return on emerging markets better than on the developed ones. So on the developing markets one can partially predict via trading volume development of return of any kind of stocks. Thus, one should bear in mind the specifics of each stock when forming a portfolio, calibrating a model, preparing a trading strategy or conducting research.

The rest of the article has the following structure. Section 4.2 describes the procedure for collecting the primary studies. Section 4.3 investigates the presence of publication bias in the literature. Section 4.4 addresses heterogeneity in the primary studies and provides implied estimates. Section 4.5 summarizes the paper.

4.2 Data Collection

The data collection and estimation itself follow the guidelines for meta-analyses in economics proposed by Havranek & Sokolova (2020). In the first step, I search for all relevant studies. Based on related literature surveys conducted by Mahajan & Singh (2009a) and Akpansung & Gidigbi (2015) and the workhorse methods in this field (Brennan *et al.* 1998; Chordia *et al.* 2001), I design a search query for Google Scholar. The final query returns all relevant articles related to the volume-return relationship. The query is worded as follows: trade | trading and volume and "expected stock return" | "stock return" | "price changes". This search goes through the full text of the study regardless the precise formulation of the title, abstract, and keywords Gechert *et al.* (2020). Reading of the abstracts leads to the removal of three-fourths of the articles. I then read the full text of the rest of them. The latest study, from March 2019, was added, and then the literature search is terminated.

The articles deploy four comprehensive and distinctive strategies for studying the trading volume-return relationship. First, authors such as Lee & Rui (2000); Statman *et al.* (2006); Chuang & Lee (2006); Gurgul *et al.* (2007) focus on the effect of lagged returns on current trading volume. They follow an intuitive logic, supposing that people invest in stocks that displayed profits in the last season. The results of these studies support this intuition. The authors find that most of their estimates are significant. The second group of articles tests Granger causality. Granger (1969) proposes this methodology to test for "a correlation between the current value of one variable and past values of other variables" (Brandle 2010). VAR models serve as the baseline for these tests. In the context of return and volume, these models assess whether volume Granger-causes returns and vice versa (e.g., Mestel *et al.* 2003; Akpansung & Gidigbi 2015). The literature describes the relationship as weak or nonexistent.

The third group of studies, starting with Ying (1966), investigates stock markets by trading volume growth and stock returns. Ying (1966) finds a significant and positive relationship between volume growth and corresponding returns using the S&P 500 composite index. Similar findings are obtained, for example, by Gervais *et al.* (2001) when expanding Ying's analysis and by Watkins (2007) when using monthly NASDAQ data. The fourth and last group of articles studies the trading volume-return relationship itself. This group is represented by, for example, the aforementioned Brennan *et al.* (1998); Chordia *et al.* (2001), who find a negative and significant relationship using US stock data.

Since the results based on these four approaches are mutually incomparable, I study the fourth group only. There are several reasons for this decision: First, the aim of Fama-French factor models and Amihud & Mendelson's approach is to determine stock returns based on trading volume, *not vice versa.* I thus eliminate the first group. Second, the discussion is not focused on the question of simply *whether* there is a relationship but on *how much* trading volume affects returns. This excludes the Granger-causality testing studies.

Third, many articles studied in this meta-analysis observe the trading volume and return relationship. This is not the case for the previous groups, including the group studying growth in returns. Last but not least, even though some differences remain in the approaches to estimation in the fourth group, all of them can be accounted for with the meta-analytical tools described in the rest of this Section and Section 4.4.

A detailed description of the study selection path is provided in Figure B1. In addition to skipping the articles investigating the opposite relationship, Granger causality or growth in volume, I dropped the researches without measurements of the uncertainty of the estimates. The test for the presence of publication bias requires either standard errors or other metrics derived from standard errors. This condition stops the inclusion of some key contributions such as Chordia & Swaminathan (2000).

The final set of estimates meeting all conditions for the meta-analysis consists of 468 observations from 44 studies. The majority of the studies focus on US stock markets (e.g., Crouch 1970; Epps & Epps 1976). Nonetheless, numerous studies from recent years assess emerging markets (e.g., De Meiros & Van Doornik 2008; Tapa & Hussein 2016) and China (e.g., Shu *et al.* 2004). The data contains observation for both trading volume and stock returns and trading volume and expected stock returns.

The data include both published articles and working papers. While using only published studies offers reassurance of the quality of the estimates, the inclusion of unpublished papers does not negatively affect the results. Rusnak *et al.* (2013) discuss the usage of working papers and their effect on publication bias and suggest, "Authors who would preferably publish some estimates would do it rationally in early stage of publication." The same idea is raised by Doucouliagos & Stanley (2013) and their evidence on 87 meta-analyses. They conclude that there is "no difference in the magnitude of publications selection between unpublished and published studies". The same confirms even Astakhov *et al.* (2019) studying the relationship between firm size and stock returns. Moreover, the inclusion of both published and unpublished articles enables me to study the difference between these two subgroups and helps better reveal the drivers of heterogeneity. An overview of the articles used appears in Table 4.1. The oldest study in the collected sample is from 1970 and the newest from 2019; therefore, the data have almost 50 years of coverage. The data are available upon request.

Table 4.1: Studies included in the meta-analysis

Al-Jafari & Tliti (2013)	Long *et al.* (2018)
Assogbavi *et al.* (2007)	Louhichi (2012)
Brandle (2010)	Loukil *et al.* (2010)
Brennan *et al.* (1998)*	Mahajan & Singh (2008)
Ciner (2002)	Mahajan & Singh (2009a)
Ciner (2003)	Mahajan & Singh (2009b)
Chang & Wang (2019)*	Marshall & Young (2003)
Chen *et al.* (2001)	McGowan & Muhammad (2012)
Chordia *et al.* (2001)*	Narayan & Zheng (2010)
Crouch (1970)*	Ochere *et al.* (2018)
Datar *et al.* (1998)*	Pisedtasalasai & Gunasekarage (2007)
De Meiros & Van Doornik (2008)	Rotila *et al.* (2015)
Devanadhen *et al.* (2010)	Saatcioglu & Starks (1998)*
Epps & Epps (1976)*	Sheu *et al.* (1998)
Hafner (2005)	Shu *et al.* (2004)
Han *et al.* (2018)	Sana Hsieh (2014)
Hu (1997)	Tahir *et al.* (2016)
Le & Mehmed (2009)	Tapa & Hussein (2016)
Lee & Rui (2000)	Tripathy (2011)
Lee & Rui (2002)*	Yin & Liu (2018)
Lewellen (2015)	Yonis (2014)
Lin & Liu (2017)*	Zhong *et al.* (2018)

Notes: Table 4.1 provides overview of all the primary studies employed in this meta-analysis. The studies with the asterisk are those published in top-tier journals (specified in Table 4.6.

Despite the strict selection criteria for the articles, several inconsistencies remain. These relate to the measures of return and volume themselves. In the case of return measures, most of the studies employ returns or absolute returns. The definition of returns is as follows:

$$\Delta Ret = ln(P_t) - ln(P_{t-1}) = ln(P_t/P_{t-1}), \tag{4.1}$$

where P stands for price and t captures time horizon.

Some older papers, such as Crouch (1970), employ price changes instead of returns. Nevertheless, authors now prefer returns to price changes since returns can allow comparisons between different stocks, firms, or studies. Moreover, some authors use *Abnormal* returns instead of returns. Abnormal returns are above-average returns from the previous time frame (e.g., Yin & Liu 2018). Other authors prefer *Excess* returns, considering only returns above the risk-free rate (e.g., Chordia *et al.* 2001). In particular, the last method is widely applied in Fama-MacBeth types of models. Fama-MacBeth models adapt this measure from the capital asset pricing model (CAPM) and arbitrage pricing theory (APT) (Brennan *et al.* 1998).

Measurement of trading volume has also evolved over time. Early authors such as Crouch (1970) and Epps & Epps (1976) employ the number of shares traded as their volume measure. However, the turn-of-the-century study of (Datar *et al.* 1998) suggests, "The number of shares traded by itself is not a sufficient statistic for the liquidity of a stock since it does not take into account the differences in the number of shares outstanding or the shareholder base". These authors, together with Brennan *et al.* (1998), proposed two alternatives. First, the turnover rate is associated with the investor holding period. Second, the dollar trading volume is related to how long a dealer waits to turn around his position (Chordia *et al.* 2001). Finally, Lo & Wang (2000) compare all these approaches and recommend turnover as the most natural proxy for trading volume in the stock market. Thus, turnover is the preferred measure in most studies today (e.g., Long *et al.* 2018; Chang & Wang 2019; Zhong *et al.* 2018).

Last but not least, the authors differ even in their approach to return-volume relationship measurement. One group, represented by Brennan *et al.* (1998); Chordia *et al.* (2001), explores the effects of past volume on expected stock returns. These major authors in the field (Brennan *et al.* 1998; Chordia *et al.* 2001) suggest a negative effect of lagged volume on expected returns. The second, similarly sized group (e.g., Epps & Epps 1976; Datar *et al.* 1998) studies the relationship of volume and returns in the same time period. Unlike the dynamic relationship, the contemporaneous relationship between returns and volume clarifies information about trading volume asymmetry.

And this is the main difference one should bear in mind when thinking about the contemporaneous and the dynamic effects between the trading volume and stock returns. Karpoff (1987) and McMillan & Speight (2002) clarify that the contemporaneous relationship in general reveals the "degree of asymmetry of volume in bull and bear markets" (Karpoff 1987). The explanation why it is so may differ from author to author, some prefer mixture of distributions hypothesis (Clark 1973), others are in favor of sequential arrival of information hypothesis of (Copeland 1976). The first group (e. g. Epps & Epps 1976; Tauchen & Pitts 1983) looks at the trading volume as at the proxy for the unobservable directing process of returns, second group (e. g. Jennings *et al.* 1981) pinpoints that the new information is disseminated across the trader sequentially, thus they adjust their behaviour one by one. All in all both of these hypothesis comes to conclusion that the contemporaneous relationship deals with asymmetry.

On the contrary, the dynamic relationship between trading volume and stock returns does not work with information asymmetry, but shows the informational efficiency of the market McMillan & Speight (2002). Concretely, the question is, whether dynamic relationship is consistent or not with weak-form efficiency, since if it is it would imply that investors are able to make systemic profits based on the information about trading volume (Fama 1965). Besides, for instance Lee & Rui (2002) see another important information in studying the dynamic effect. They highlight the possibility of spillover effect from one market to another. This might be another reason, why the dynamic relationship is important. Both the dynamic and contemporaneous approaches yield inconclusive results across empirical evidence (Hu 1997; Akpansung & Gidigbi 2015; Poudel & Shrestha 2019). Thus, this study incorporates both relationships, distinguishes them with a dummy, and discuss their true value.

The variability in the measures of returns and volume obliges me to employ the approach of Valickova *et al.* (2015a). In their investigation of financial development and economic growth, these authors define four measures of economic growth (dependent variable) and ten variables for financial development (independent variable). The partial correlation coefficients (*PCCs*) enable comparability of the estimates at the cost of losing some information. *PCCs* come from the t-statistic of the estimate and the residual degrees of freedom (Greene 2008). The sign of a partial correlation coefficient is the same as the sign of the original coefficient:

$$PCC_{ij} = \frac{t_{ij}}{\sqrt{t_{ij}^2 + df_{ij}}}, \tag{4.2}$$

where r_{ij} stands for the partial correlation coefficient of the ith estimate of the jth study. t denotes the corresponding t-statistic, and df the degrees of freedom. The *PCCs* provide only the recalculation of the estimates stemming from different specifications into the comparable form. Since I am interested in the comparison between the different original specifications, I captured each different group of dependent and independent variable by dummy in the final Equation 4.6 for studying the drivers of the heterogeneity.

In addition to the estimates themselves, the corresponding standard errors need recalculation. I again follow the approach adopted by Valickova *et al.* (2015a), as suggested by Doucouliagos & Stanley (2013). They adapt the formula from Fisher (1954):

$$SEr_{ij} = \frac{PCC_{ij}}{t_{ij}}, \tag{4.3}$$

where SEr_{ij} denotes the standard error of the particular partial correlation coefficient PCC_{ij}. The t_{ij} expresses the t-statistic from the ith regression of the jth study. In regards to other authors employing the partial correlation coefficient in economic meta-analyses, I can mention, for instance, Doucouliagos (2005). An overview of the distribution of *PCCs* per study used in my research is provided in Figure 4.3.

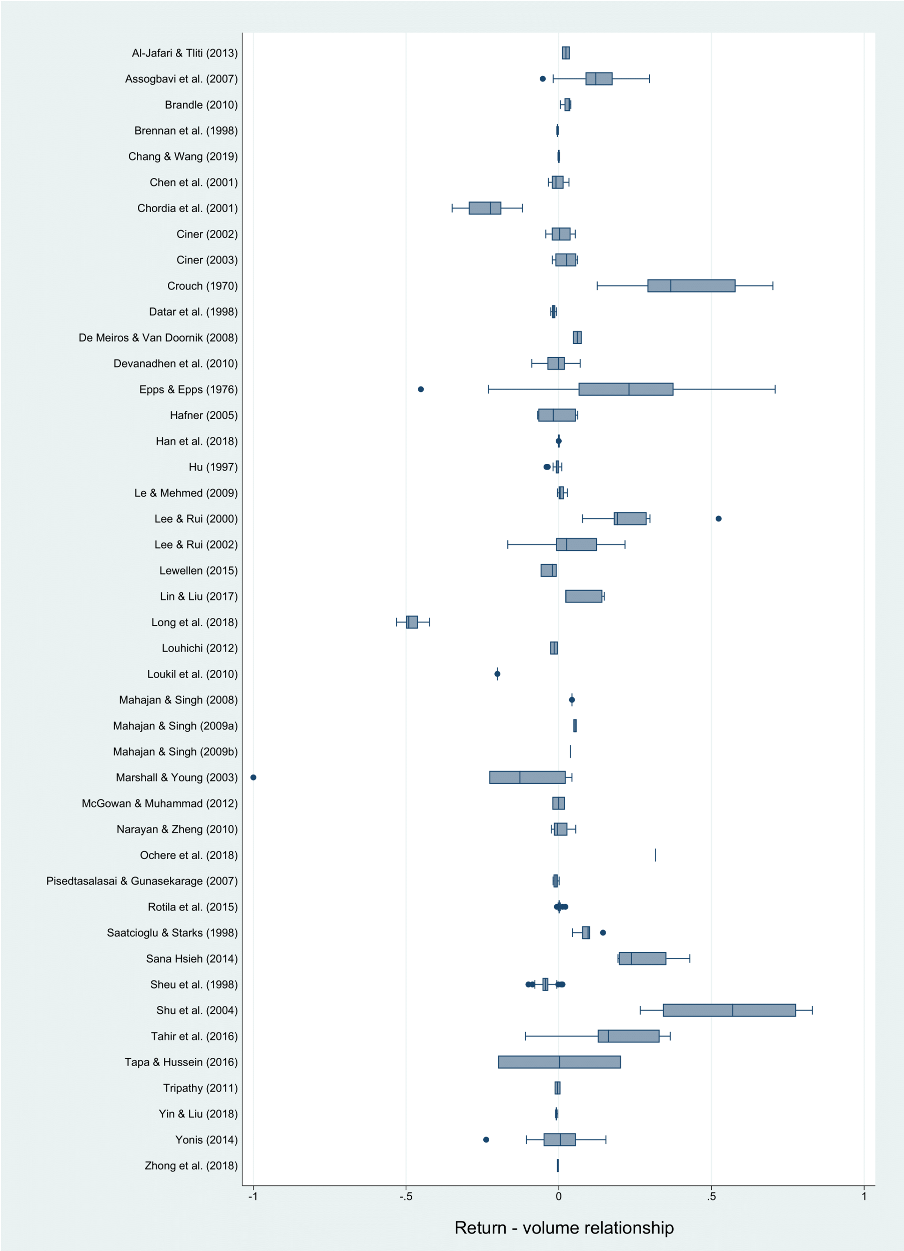
The mean reported estimate of the return-volume relationship is 0.071 for contemporaneous case and -0.026 for dynamic case. Winsorization at 2.5% helps to deal with some extreme outliers in the data, but the means remain the same after winsorization. Obviously, the estimates vary greatly both between and even within the primary studies. Thus, 49 explanatory variables collected for each observation address this variance. The task of these variables is to clarify diverse data characteristics such as data type, methodology used or market size.

These variables are very important in the meta-analysis. They capture the study-specific factors such as data frequency, estimation methodology or sample length. Besides these detect the outward surroundings where the study emerges. It is important to collect all variables that may possibly affect the results. These variables help me to understand the heterogeneity among the primary estimates and address the omitted variable bias during the further estimation.

That it is crucial to do so show, for instance, meta-analyses by Havranek *et al.* (2012); Ehrenbergerova *et al.* (2021). They indicate, for instance, that different type of data gives different estimates. And, for example, Valickova *et al.* (2015a) find out that the quality of the publication might affect the results.

That these variables really affects the estimates confirms even the authors of the primary studies themselves. For instance, Kim *et al.* (2019) suggest that the

Figure 4.3: Variation in the estimates both across and within studies



Notes: The box plot of the estimates of return-volume relationship published in primary studies highlights both median and interquartile range (P25 – P75). The coverage of whiskers reaches from (P25 – 1.5*interquartile range) to (P75 + 1.5*interquartile range). The dots capture remaining outlying estimates. The winsorization handles with overall outliers before computational tasks.

lower-frequency data are subjected more to market frictions than the higher-frequency data. Elsewhere, Chordia *et al.* (2001) or Datar *et al.* (1998) discuss different results for stocks from different markets. And, for instance, Lee & Rui (2002) and Saatcioglu & Starks (1998) pinpoint differences in estimates from developed and emerging countries.

Thus it is very important to capture these differences within and across the studies to evaluate the literature properly. The foretaste of these distinctions provides numerically the Table 4.1, but the more in detail explanation of the variable together with other statistics are to be found in Section 4.4.

Table 4.2: Mean return-volume estimates for different subsets of data

	No. of obs.	Studies	Mean	Stand. Dev.	95% conf. int.	
Temporal dynamics						
Contemporaneous	224	28	0.071	0.011	0.050	0.093
Dynamic	244	28	-0.026	0.008	-0.041	-0.010
Data characteristics						
Hourly data	52	4	0.182	0.030	0.121	0.242
Daily data	118	21	0.074	0.013	0.048	0.099
Weekly data	32	2	0.113	0.016	0.08	0.147
Monthly data	266	18	-0.045	0.006	-0.058	-0.033
Panel data	286	20	-0.038	0.006	-0.051	-0.025
Time series data	175	22	0.118	0.012	0.094	0.143
Cross-sectional data	7	2	-0.009	0.013	-0.041	0.023
Structural variation						
All stocks	220	18	-0.051	0.007	-0.029	-0.008
Indexed stocks	92	18	0.057	0.013	-0.050	0.010
NASDAQ stocks	9	2	-0.103	0.042	-0.200	-0.007
Banks stocks	18	3	0.042	0.023	-0.005	0.090
Firms stocks	129	9	0.123	0.015	0.092	0.154
Developing countries	136	25	0.058	0.013	0.031	0.084
OECD countries	332	24	0.006	0.008	-0.01	0.022
Publication status						
Published papers	367	38	0.024	0.009	0.007	0.042
Unpublished papers	101	6	0.008	0.007	-0.006	0.021
Top	200	9	0.016	0.013	-0.010	0.042
All estimates	468	44	0.021	0.007	0.007	0.035

Notes: Table 4.6 provides a complete description of the definitions of subsets. Winsorization at 2.5% and 97.5% levels deals with the outliers.

A glimpse of the heterogeneity provides Table 4.2. It summarizes the mean values of the return-volume relationship for different subgroups of data. These subgroups consider temporal dynamics, data frequency, type of data, type of stocks and publication characteristics.

The dynamic estimates show negative and significant effects, as in Brennan *et al.* (1998); Chordia *et al.* (2001). On the other hand, the contemporaneous estimates display slightly positive effects, as in Hiemstra & Jones (1994). The dynamic relationship usually connects the panel data with monthly frequency; thus, the means for the panel and monthly data subgroups are negative, like the mean of the dynamic relationship subgroup. In contrast, time series data at higher frequencies exhibit substantially positive results. Different types of stocks provide the following information. Firm stocks tend to have the highest mean estimate. On the other hand, NASDAQ stocks remain the lowest by a substantial margin. This finding appears even in Brennan *et al.* (1998); Chordia *et al.* (2001).

Besides the distinction between developing and OECD countries appears. Last but not least, top-tier publications, published and unpublished papers do not seem to differ significantly. In summary, this simple analysis proposes systematic differences among the reported estimates, but without correction for publication bias as in Section 4.3 and proper investigation of the sources of heterogeneity as in Section 4.4, any conclusions drawn will be misleading.

4.3 Publication Bias

The phenomenon of publication bias extensively affects economic literature. Ioannidis *et al.* (2017a) find that estimates reported in the economics literature are typically exaggerated twofold because of publication bias. It is understandable that authors naturally prefer a statistically significant estimate with the expected sign. This preference makes sense. One should not have to focus on evidently wrong estimates. On the other hand, substantial ignorance of statistically insignificant estimates with the "wrong" sign distorts the literature as a whole.

Addressing this subject, McCloskey & Ziliak (2019) discussed the Lombard effect. Like speakers who raise their voice in the presence of noise, researchers particularly augment their efforts to find a significant effect in the case of noisy data or poor estimation techniques. Statistically significant estimates at the 5% level with the "correct" sign are nearly always possible to reach in economics in the presence of the freedom to choose from among a large number of different specifications. On the other hand, statistically significant results gained in this manner no longer reflect the primary theoretical purpose of conducting statistical tests.

A classical perspective from which to study the return-volume relationship is in relation to the efficient market hypothesis. Basu (1977) observes, "While there is substantial empirical evidence supporting the efficient market hypothesis, many still question its validity". The same holds even today. For instance, Malkiel (2003) emphasizes that "pricing irregularities and even predictable patterns in stock returns can appear over time and even persist for short periods". Thus, some authors may find or try to find no significant effect, with null estimates, or make excessive efforts to find significant results (the Lombard effect). Therefore, studying how estimates of the trading-volume relationship are obtained is a compelling topic to scrutinize. It is important to determine whether estimates differ simply because of different economic and data backgrounds (Section 4.4) or because of selection by authors.

A common tool for detecting the extent of publication selection is the so-called funnel plot, first proposed by Egger *et al.* (1997a). A funnel plot depicts the magnitude of the estimated effect on the horizontal axis. The vertical axis then captures the precision, measured by the inverse of the estimated standard error. Since the studies on the return-volume relationship provide standard errors with a symmetrical distribution (usually a t-distribution), the estimates should have a symmetrical distribution around the true mean effect regardless of their magnitude and precision.

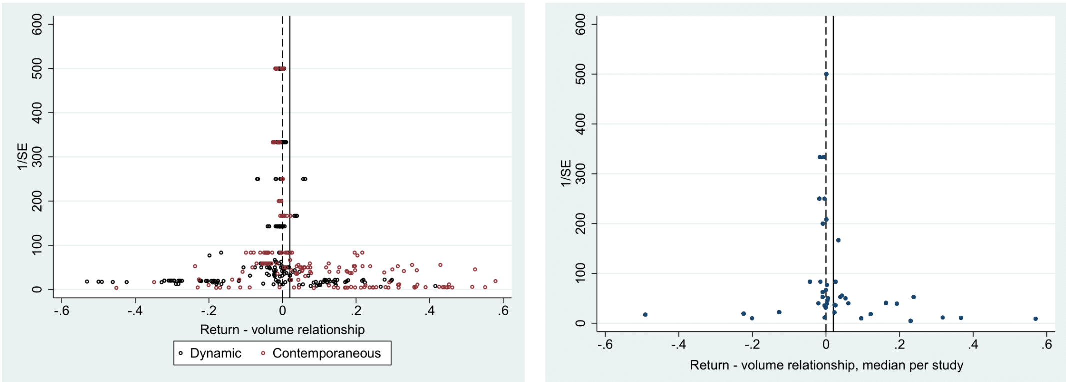
The estimates become further from the true effect as precision decreases. Thus, the estimates form a symmetrical inverted funnel. In the presence of publication bias, the funnel plot should be asymmetrical or hollow. The discarding of estimates of a particular sign or magnitude would cause this asymmetry, while the rejection of statistically insignificant estimates would cause the hollowness. The worst case arises when the funnel plot is both asymmetrical and hollow (Egger *et al.* 1997a).

Figure 4.4, which presents contemporaneous and dynamic estimates separately, gives a clear message. The depicted funnel plots show that the dynamic estimates are distributed more or less equally around zero. The same holds for the median estimates for each study. In contrast, the contemporaneous estimates are skewed to the left, which indicates the possible presence of publication bias in this case.

The funnel plot represents only a simple visual test. A more reliable way to check for publication selection offer regression-based funnel asymmetry tests. The following base regression (Stanley & Doucouliagos 2012) explores the correlation between the return-volume relationship and its standard error $SE(r_{ij})$:

$$r_{ij} = \beta_0 + \beta_1 SE(r_{ij}) + e_{ij}, e_{ij} \sim N(0, \sigma^2), \tag{4.4}$$

Figure 4.4: Funnel plot: Little evidence of publication bias in this field



(a) (b)

Notes: Without the publication bias the scatter plot seems like an inverted funnel symmetrical around the most precise estimates. The left panel depicts all estimates distinguished by the time dynamics. The right panel shows median estimates per study. The solid line stands for overall mean relationship. The dashed line is set at zero. The computational tasks includes even outliers in winsorized form, but for the ease of exposition the funnels excludes them.

where r_{ij} stands for the ith estimate of the partial correlation coefficient between expected stock returns and trading volume from study j. β_0 expresses the mean underlying effect beyond publication selection bias, and the coefficient β_1 reveals the strength of publication bias. The aforementioned Lombard effect or discarding of estimates with the "wrong" sign may cause the correlation. If $\beta_1 = 0$, publication bias is not present in the field. Otherwise, publication bias is present.

I estimate Equation 4.4 with four different estimation methodologies. First, I use simple OLS with standard errors clustered at the level of individual studies and countries. The two-way clustering follows the suggestion of Cameron *et al.* (2012). Second, I run a panel data regression employing between effects. Third, I follow Stanley & Doucouliagos (2012) and Astakhov *et al.* (2019) in multiplying Equation 4.4 by $1/SE(r_{ij})$. This assigns more weight to more precise studies and directly deals with heteroskedasticity. Therefore, the weight $1/SE(r_{ij})$ is called *Precision*. In the fourth specification, instead of *Precision*, I use the inverse number of estimates per study as a weight.

In addition to commonly used and widely known publication bias detection techniques, I employ two recently developed advanced techniques. Estimating β_0 from Equation 4.4 yields an unbiased estimate of the mean corrected for publication bias only if publication selection is proportional to the standard error. Nevertheless, in practice, I am dealing with an unknown functional form of the publication selection procedure.

Therefore, first, I employ the advanced estimator introduced by Furukawa (2019). This approach, known as the stem-based method, works only with the most precise estimates, optimizing the number of estimates for investigating publication bias by minimizing the mean squared error of the estimates. This conservative, fully data-dependent, nonparametric method robustly alleviates publication bias under various assumptions. The second advanced technique was specified by Bom & Rachinger (2019). Their method also known as kinked method. It improves on the precision effect estimate with standard error (PEESE) test for publication bias.

It should be also noted that these advanced techniques in comparison with the previous four do not directly address the magnitude of publication bias. These focus rather on the *true* effect beyond bias. It means that they specify, what the effect would be when no bias in literature would be present.

Table 4.3: Formal tests on the presence of publication bias

	All	Contemporaneous	Dynamic
PANEL A: Unweighted estimations			
OLS			
SE (publication bias)	0.867***	0.876***	-0.169
	(0.091)	(0.100)	(1.247)
	[-0.487; 2.349]	[-0.844; 4.775]	[-3.729; 2.714]
Constant (effect beyond bias)	-0.013	0.027	-0.021
	(0.014)	(0.018)	(0.013)
	[-0.063; 0.042]	[-0.021; 0.106]	[-0.094;-0.004]
Between effects			
SE (publication bias)	1.069**	1.229**	1.436**
	(0.483)	(0.601)	(0.621)
	-	-	-
Constant (effect beyond bias)	-0.002	0.038	-0.057*
	(0.026)	(0.033)	(0.029)
	-	-	-
PANEL B: Weighted OLS estimations			
Weighted by the inverse of the number of estimates reported per study			
SE (publication bias)	0.965**	0.960*	0.860
	(0.478)	(0.531)	(0.799)
	[-1.091; 2.995]	[-2.644; 4.613]	[-1.405; 2.454]
Constant (effect beyond bias)	0.001	0.044*	-0.041**
	(0.012)	(0.021)	(0.018)
	[-0.025; 0.026]	[0.000; 0.090]	[-0.083; -0.007]
Weighted by the inverse of the standard error			
SE (publication bias)	0.771**	1.672***	-0.807
	(0.376)	(0.444)	(1.255)
	[-2.330; 2.498]	[-6.747; 7.385]	[-3.585; 2.019]
Constant (effect beyond bias)	-0.009**	-0.014	-0.003
	(0.004)	(0.010)	(0.003)
	[-0.037; -0.002]	[-0.148; 0.019]	[-0.122; 0.070]
PANEL C: Non-linear estimations			
Stem-based method (Furukawa 2019)			
Effect beyond bias	-0.006**	0.001	-0.006***
	(0.003)	(0.006)	(0.001)
Kinked method (Bom & Rachinger 2019)			
Effect beyond bias	-0.011***	-0.023***	-0.003***
	(0.002)	(0.004)	(0.001)
Observations	468	224	244

Notes: The uncorrected mean of the contemporaneous estimates is 0.071 and for the dynamic one -0.026. Standard errors reported in parentheses are clustered at the study and country level (except between effects; the usage of two-way clustering follows Cameron *et al.* 2012). The square brackets report 95% confidence intervals from wild bootstrap clustering and Rademacher weights with 999 replications (except between effects; the implementation follows Roodman 2020). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4.3 presents the results of all six specifications. At the first glance, it seems through the first column that the publication bias is present across All estimates. But based on the observation from Figure 4.4 it is more interesting to focus on the publication bias for dynamic and contemporaneous estimates separately.
The first row shows the baseline result of the OLS regression of the partial correlation coefficient on its standard error. The β_1 coefficients indicating the possible presence of publication bias are both positive and significant. This suggests a strong selective reporting bias for contemporaneous estimates. Besides the results for contemporaneous estimates show negative but insignificant constant representing the underlying mean partial correlation coefficient corrected for reporting bias. On the other hand, the the dynamic estimates do not show any bias and the effect beyond bias is also negligible. Based on these observations alone it seems that authors make efforts to find an effect of trading volume on stock returns, but there is none in contemporaneous and even in dynamic case.
The second part of Panel A of Table 4.3 exhibits the results of the panel data regression with between effects. The between effects indicate an even stronger selective reporting bias than that found in the case of OLS. It holds for both cases contemporaneous and dynamic. The corrected partial correlation coefficient again appears insignificant for contemporaneous estimates On the contrary, among the dynamic estimates a negative and significant effect beyond bias emerges. However, even in this case, the effect is not substantial: Doucouliagos (2011), in his guidelines on partial correlation coefficients, considers such an effect not even "small". He defines a small effect as one ranging from 0.07 to 0.17 in absolute value, a medium effect as one ranging between 0.17 and 0.33, and a large effect as one above 0.33.
Moving on, Panel B reports the analysis of the WLS estimation with the precision and inverse number of estimates per study as weights. The findings derived from these two specifications simply accentuate the findings from Panel A. They confirm the presence of bias in the contemporaneous case and a negligible effect beyond bias in both cases. Finally, Panel C summarizes the results of the nonlinear techniques, which are focused on the true effect beyond bias. This tests show ambiguous findings on the significance and sign of the contemporaneous effect.

Table 4.4: Formal tests on the presence of publication bias - Top Journals

	All	Contemporaneous	Dynamic
PANEL A: Unweighted estimations			
OLS			
SE (publication bias)	0.946***	0.829***	-1.205
	(0.030)	(0.035)	(1.894)
	[-0.954; 2.974]	[-4.987; 5.420]	[-33.180; 9.564]
Constant (effect beyond bias)	-0.041***	0.032***	-0.052
	(0.007)	(0.008)	(0.042)
	[-0.253; 0.242]	[-0.241; 0.200]	[-0.608; 1.119]
Between effects			
SE (publication bias)	1.163	1.057	0.836
	(0.808)	(0.844)	(1.809)
	-	-	-
Constant (effect beyond bias)	-0.008	0.068	-0.068
	(0.070)	(0.086)	(0.105)
	-	-	-
PANEL B: Weighted OLS estimations			
Weighted by the inverse of the number of estimates reported per study			
SE (publication bias)	0.992***	0.874***	0.746
	(0.075)	(0.045)	(0.972)
	[0.089; 2.667]	[-1.357; 5.111]	[-14.980; 9.706]
Constant (effect beyond bias)	0.002	0.078***	-0.061**
	(0.008)	(0.009)	(0.027)
	[-0.116; 0.122]	[-0.090; 0.211]	[-0.572; 0.333]
Weighted by the inverse of the standard error			
SE (publication bias)	0.416***	1.341***	-2.543**
	(0.067)	(0.064)	(1.054)
	[-2.973; 2.001]	[-58.680; 27.120]	[-22.970; 1.092]
Constant (effect beyond bias)	-0.008***	-0.009***	-0.002
	(0.001)	(0.002)	(0.002)
	[-0.191; 0.173]	[-0.435; 0.190]	[-0.729; 0.980]
PANEL C: Non-linear estimations			
Stem-based method (Furukawa 2019)			
Effect beyond bias	-0.006***	-0.002	-0.006***
	(0.001)	(0.002)	(0.001)
Kinked method (Bom & Rachinger 2019)			
Effect beyond bias	-0.004***	-0.012***	-0.001
	(0.001)	(0.002)	(0.001)
Observations	200	118	82

Notes: The uncorrected mean of the estimates is 0.016. Standard errors reported in parentheses are clustered at the study and country level (except between effects; the usage of two-way clustering follows Cameron *et al.* 2012). The square brackets report 95% confidence intervals from wild bootstrap clustering and Rademacher weights with 999 replications (except between effects; the implementation follows Roodman 2020). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

On the other hand, they find the dynamic effect negative and significant after publication bias is corrected. These results line up with the findings of, for instance, Brennan *et al.* (1998). Moreover, the dynamic effect of the trading volume passes even the test for new expected stock returns determinant proposed by Harvey *et al.* (2016). Harvey *et al.* (2016) suggest that new expected stock return determinant should have the t-statistic greater than 3.0.

In Table 4.4 I restrict the analysis only to the estimates from the top-tier journals. I was thinking about inclusion of estimates only from Top 5 journals, but such a choice would leave me with only three articles and that would be insufficient for the analysis. Thus, I use Financial Analysts Journal and its article influence score from the Web of Science as a threshold for a peer-review quality and expand the selection a bit. This journal published several key finance articles but still allows enough better-ranked articles to be involved in the restricted analysis. In the choice I follow Bajzik *et al.* (2020) who in discussion of international trade use similarly-ranked Canadian Journal of Economics as a journal-quality threshold. Finally, in the restricted sample I have 200 observations from nine studies.

According to Rusnak *et al.* (2013) and Doucouliagos & Stanley (2013) the analysis of the top-tier articles should not differ significantly from the previous ones. This expectation turns to be almost true, since the results do not differ significantly in the sign and magnitude. But still one may observe the top journal estimates are more precise, thus the results about the presence of publication bias are a bit more convincing. On the other, the bootstrap confidence intervals turn more unstable.

Furthermore, I use the Caliper test proposed by Gerber *et al.* (2008); Gerber & Malhotra (2008); Bruns *et al.* (2019) to supplement the inspection of publication bias. This test focuses on different results selection stages called p-hacking and HARKing. In general, in cases of publication bias, authors simply do not publish results with insignificant estimates, but in the case of p-hacking, authors include only models with significant estimates in the study. In the case of HARKing, authors set their hypothesis after the results are already known Bruns *et al.* (2019). That the p-hacking (and HARKing) may be a problem in area of expected stock returns indicate, for instance, Harvey (2017).

The caliper test does not reveal anything about the corrected effect. It is based on the study of the break in reported t-statistics, where the break around the usual significance threshold indicates selective reporting. When authors do not report selectively, the distribution of t-statistics remains even around the usual significance thresholds of 1.96, 1.645, and 1.

Table 4.5: Caliper test

		Contemporaneous		Dynamic	
	Caliper size	0.1	0.2	0.1	0.2
90%	Above C	-	40.0%	22.2%	22.2%
	Below C	-	60.0%	77.8%	77.8%
	p-value	-	0.704	0.095	0.014
95%	Above C	66.7%	66.7%	85.7%	73.3%
	Below C	33.3%	33.3%	14.3%	26.7%
	p-value	0.347	0.207	0.047	0.068
99%	Above C	66.7%	50.0%	30.0%	55.6%
	Below C	33.3%	50.0%	70.0%	44.4%
	p-value	0.465	1.000	0.223	0.651

Notes: The table provides caliper tests following Bruns *et al.* (2019) for caliper sizes 0.1 and 0.2 and for the hypothesis of a 50:50 distribution. The numbers express the share of observations in a given interval around the significance threshold. The test parameter follows $C = \frac{n_{oc}}{n_{oc}+n_{uc}}$, where n_{oc} and n_{uc} stand for the number of observations with t-statistics in the interval above and below the threshold.

The results summarized in Table 4.5 suggest no selection around the 90% and 99% levels. On the other hand, breaks at the middle level – around the 95% interval – indicate bias. The results hold for different caliper sizes. This means that authors push their estimates above the 95% level but do not do so at the 10% and 1% levels. Moreover, contrary to the conclusions on publication bias, this type of bias distorts the dynamic estimates.

All in all, the formal tests reveal following. First, there is a publication bias present in the literature among the contemporaneous estimates. Second, the corrected mean for contemporaneous effect is negligible and insignificant. These two findings show that the authors tend to publish positive relationship, when there is none. Third, the dynamic estimates do not subject to the publication bias and are of significant value. But even in this case, their true value is closer to zero than the results suggested by the literature. Fourth, authors are particularly likely to provide biased estimates around the higher confidence intervals.

4.4 Drivers of the Relationship

On the other hand, the results from publication bias section may suffer endogeneity problems, since other aspects than publication bias may influence the value of the corrected mean. Thus it is important to collect all the factors that may possibly influence the return-volume relationship, both study-variant and study-invariant, and compare them qualitatively.

At least five different explanations of the difference in the estimates of the return-volume relationship have been repeatedly mentioned in the literature. The first three of them – the different measures of volume and returns and the difference between contemporaneous and dynamic estimates – I have already discussed in Section 4.2. As a fourth reason, the existing literature suggests an effect of data frequency. The lower-frequency data are subjected more to market frictions (Kim *et al.* 2019). On this problem point even Saatcioglu & Starks (1998) studying the results from weekly data (Granger & Morgenstern 1963) and monthly data (Rogalski 1978).

The fifth explanation is that the differences may spring from data aggregation. The less aggregated data are prone to exhibit frictions. For instance, studies such as Jain & Joh (1988) and Lee & Rui (2000) use aggregate data from stock markets. Their findings again indicate a positive relationship between return and volume. In contrast, Chordia *et al.* (2001) distinguish *NASDAQ* stocks, since the trading volume-stock return effect appears more negative for this exchange than for the NYSE or AMEX (which belong to the *All* group). Other groups include *Index* stocks Han *et al.* (2018), *Banks* stocks (Rotila *et al.* 2015; Al-Jafari & Tliti 2013) and *Firms* stocks (Tahir *et al.* 2016; Datar *et al.* 1998). The findings from the studies focused on these groups vary.

Table 4.6: Description and summary statistics of the regression variables

Label	*Description*	*Mean*	*SD*	*Median*
SE	Estimates of standard errors of return-volume relationship (winsorized at 1% level)	0.04 (0.04)	0.06 (0.06)	0.02 (0.02)
Data characteristics				
Contemporaneous	=1 if the return-volume relationship is contemporaneous	0.48	0.50	0.00
Dynamic	=1 if the return-volume relationship is dynamic	0.52	0.50	1.00
Returns	=1 if the returns in any form are estimated	0.96	0.20	1.00

Continued on next page

Table 4.6: Description and summary statistics of the regression variables (continued)

Label	*Description*	*Mean*	*SD*	*Median*
Price change	=1 if the price change is estimated instead of returns	0.04	0.20	0.00
Normal	=1 if the returns or price change itself are used	0.42	0.50	0.00
Absolute returns	=1 if the returns or price change are in absolute terms	0.12	0.32	0.00
Abnormal returns	=1 if the returns are defined as abnormal	0.03	0.16	0.00
Excess returns	=1 if the returns are defined as excess	0.43	0.50	0.00
Turnover	=1 if the volume is expressed as the number of shares traded during a time period divided by the number of shares outstanding at the end of the time period	0.55	0.50	1.00
Dollar volume	=1 if the volume is expressed in terms of dollar volume of the trade	0.13	0.34	0.00
Shares traded	=1 if the volume is expressed in terms of shares traded	0.32	0.47	0.00
Detrended series	=1 if the volume series was detrended	0.11	0.31	0.00
Data period	Length of time period	14.53	10.81	11.5
Data size	Total of observation (in logarithms)	8.61	3.27	8.19
Midyear	The logarithm of the mean year of the data used minus the earliest mean year in our data plus one	2.99	0.75	3.13
Hourly data	=1 if the data were collected hourly or more frequently	0.11	0.32	0.00
Daily data	=1 if the data were collected daily	0.25	0.44	0.00
Weekly data	=1 if the data were collected weekly	0.07	0.25	0.00
Monthly data	=1 if the data were collected monthly	0.57	0.50	1.00
Panel	=1 if the panel data were used	0.61	0.49	1.00
Time series	=1 if the time series data were used	0.37	0.48	0.00
Cross-section	=1 if the cross-sectional data were used	0.02	0.12	0.00
Structural variation				
All	=1 if the research relies on the data for the whole stock-exchange at least	0.47	0.50	0.00
Index	=1 if the cumulative returns value for stocks from particular index was used	0.20	0.40	0.00
NASDAQ	=1 if the cumulative returns value for NASDAQ stocks is used	0.02	0.14	0.00
Banks	=1 if the returns relate only to banking sector	0.04	0.19	0.00
Firms	=1 if the returns relate to firms stocks (i. e. do not relate to the banks)	0.28	0.45	0.00
Developing country	=1 if the estimate is for developing country	0.29	0.46	0.00
OECD	=1 if the estimate is for OECD country	0.71	0.46	1.00
Market size	Market size in terms of GDP (billions of dolars) in midyear of data (in logarithms)	7.90	1.39	7.11
North America	= 1 if the observation is linked to the North America	0.47	0.50	0.00

Continued on next page

Table 4.6: Description and summary statistics of the regression variables (continued)

Label	*Description*	*Mean*	*SD*	*Median*
Asia	= 1 if the observation is linked to Asia	0.36	0.48	0.00
Europe	= 1 if the observation is linked to Europe	0.13	0.34	0.00
Australia	= 1 if the observation is linked to Australia	0.02	0.15	0.00
Other Continents	= 1 if the observation is linked to Latin America or Africa	0.02	0.15	0.00
Estimation technique				
Fama-Macbeth	=1 if the Fama-Macbeth model is used	0.48	0.50	0.00
VAR	=1 if the VAR model is used	0.24	0.42	0.00
Simple model	=1 if the simple linear model is used	0.22	0.42	0.00
GARCH	=1 if the ARIMA with GARCH in error term is used	0.06	0.24	0.00
Monday	=1 if effect of Monday or January trading is considered	0.03	0.17	0.00
Trimmed	=1 if the primary dataset was trimmed	0.09	0.29	0.00
January excluded	=1 if all months but January are included in the primary dataset	0.08	0.27	0.00
OLS	= 1 if OLS estimation method is employed	0.43	0.50	0.00
MLE	= 1 if MLE estimation method is employed	0.04	0.20	0.00
GMM	=1 if GMM estimation method is employed	0.29	0.45	0.00
Other methods	=1 if other types of estimation is employed	0.24	0.43	0.00
Publication characteristics				
Impact factor	Discounted recursive impact factor from RePEc IDEAS	0.64	0.97	0.05
Citations	The logarithm of the number of Google Scholar citations normalized by the number of years since the publication year plus one	1.96	1.49	1.60
Published	=1 if the article was published	0.78	0.41	1.00
Top	=1 if the estimate was published in journal with ranking higher than *Financial Analysts Journal*	0.43	0.50	0.00

Notes: This table shows the mean, standard deviation and median for each variable used in the estimation. The variables in *italics* serves as benchmark variables in BMA and FMA estimations. None of the variables has correlation above 0.85. The effects in the brackets for standard error shows values for the unwinsorized estimates. The average partial effect method (Wooldridge 2015) is used for the means of all variables in logarithms. Market sizes are collected from the World Bank database. (The World Bank,2019). GDP is in billions of US dollars at the midyear point of the data. For Taiwan only, data are obtained from the National Statistics Republic of China (National Statistics Republic of China, 2019), since the World Bank does not provide information for Taiwan. These values are recalculated to US dollars based on the midyear NTD-USD exchange rate according to the Federal Reserve Bank. (Federal Reserve Bank: Real Effective Exchange Rates, 2019) For each midyear, I use the year-end exchange rate to recalculate the current value. For 1981 only, I use the exchange rate from 31 December 1983, since earlier data are not available. The impact factor is downloaded from RePEc, and the number of citations is downloaded from Google Scholar. In setting the variable *Top* I follow the methodology of Bajzik *et al.* (2020). As a proxy for a peer-review quality I use the Financial Analysts Journal and its article influence score from the Web of Science. The rest of the variables are collected from studies investigating the return-volume relationship.

These five explanations represent only a few reasons why the published estimates might differ among themselves. Besides these reasons I capture by variables other aspects of the primary studies. Among these variables belongs estimation methodology, sample length or publication characteristics. These variables consider both the study-specific factors and the outward surroundings of the study. It is important to collect all variables that may possibly affect the primary estimates that the results will not suffer from endogeneity. I present my capture of the origins of the heterogeneity in Table 4.6.
To inspect the heterogeneity between the estimates of the return-volume relationship, I capture 49 features of the individual study design and expand Equation 4.4 by adding these features as independent variables. All the classified variables, with their definitions, mean, standard deviation and median lists the Table 4.6. The variables are divided into subgroups for ease of exposition. Twenty-two aspects relate to data characteristics, 13 to structural variation, 11 to estimation techniques, and the last three to publication characteristics. As it is already visible from presented discussion, the Table 4.2 and the Table 4.6 more variables possibly drive the results from primary studies. Thus, I discuss them now more in detail and present the final results later in this section.

Data characteristics. In addition to the different proxies for return and volume measures mentioned in Section 4.2, other study-invariant characteristics arose during data collection. For example, the studies are divided by data type – that is, whether the data are *Cross-sectional*, *Time series* or *Panel*. In terms of the direction of the effect of any of these characteristics, there is no prior knowledge in this field (Akpansung & Gidigbi 2015). On the other hand, that the type of data may affect the estimates show, for instance, the meta-analyses of Cazachevici *et al.* (2020); Ehrenbergerova *et al.* (2021).
In addition to the previously mentioned distinctions, I discern the number of observations, *length* of the time period in years and *Midyear* of the data used. I expect both variables to have negative sign. In case of the *length* variable is it because, on shorter data there is higher probability that some inefficiencies appears (Schwert 2003). In the case of the *Midyear* the negative sign might be caused simply by so-called "Prometheus effect". It means that the empirical effects decline over time after initial novel findings (Ioannidis 2008). In case of the financial market it might be due to effort to present the rejection of the EMH (Kim *et al.* 2019).
Due to similar reasons I considered using the year of publication (*Pubyear*) to

capture differences in publishing data, but its correlation of above 85% with *Midyear* led me to discard this idea. Then, I considered adding the squares of *Midyear*, but again, the correlation with its linear term of above 97% did not allow me to do so. According to the findings of Schurenberg-Frosch (2015), linear terms suffice. Furthermore, one variable from each mentioned group of variables is dropped due to the dummy variable trap.

Structural variation. In addition to the different volume and return measures, I investigate and capture the research area in each article with dummies. Beyond the previously mentioned categories of stock types, the distinction of whether a country belongs to the *OECD* deserves attention. The logic of the variable is to determine whether more advanced markets display different effects on returns. Emerging markets may exhibit higher volatility and a greater probability of large price changes than developed markets (De Santis *et al.* 1997).

In addition, there is a similar intuition behind the differentiation of continents. North America, Asia, Europe and Australia each have their own dummies. Based on studies conducted by Lee & Rui (2002) on the New York, London and Tokyo markets and Chen *et al.* (2001) on 9 developed countries across North America (US, Canada), Europe (UK, France, Italy, Switzerland, Netherlands) and Asia (Japan, Hong Kong) I do not expect any differences across the markets from developed countries.

On the other hand, when it comes to developing countries, the primary observations are more diverse. For instance, Saatcioglu & Starks (1998) studying Latin America countries conclude that "emerging markets with different institutions and information flows than the developed markets, do not present similar stock price-volume lead-lag relation to the preponderance of studies employing U.S. data". Similarly other authors discuss the difference between US and Asian markets and find difference mainly in the sense that Asian markets are usually dominated by individual investors. These investors are supposed to be less rational and suffer from overconfidence and cognitive bias (Kim *et al.* 2003; Kaniel *et al.* 2008). In which direction it affects the volume-return relationship is from the literature unclear (Kaniel *et al.* 2012; Zhong *et al.* 2018).

Furthermore, *Market size*, measured in terms of GDP at the midyear of the data, helps distinguish larger markets from smaller ones on each continent. It is important to do so, since on the larger markets, there degree of economic freedom is higher and hence larger markets should be more efficient. These variables together are sufficient to capture diversity in the origins of the scrutinized stocks.

Estimation techniques. Several different approaches to estimating the trading volume relationship have evolved over time. Since some of them are probably abandoned during the time, they might provide different results. That it might be so proves, for example, one of the recent meta-analysis about forward premium Zigraiova *et al.* (2021). They show that OLS estimation technique and regime-switching models provide lower estimates than other ones. Thus, I distinguish different models and estimation techniques in my research also. I proceed with the following groups.
The current workhorse model in this field is the *Fama-MacBeth* methodology. It has the same basis as the Fama-French models, and Chordia *et al.* (2001), Lewellen (2015) or Brandle (2010), for instance, have promoted it significantly. The *Fama-MacBeth* approach dominates among the estimation techniques in my sample, with 43% usage among the primary articles. The baseline equation for this group of models is $Ret = \alpha_0 + \alpha_1 Vol + \alpha_2 Size + \alpha_3 BM + \alpha_4 Price + \alpha_5 Ret_{2-3} + \alpha_6 Ret_{4-6} + \alpha_7 X + e, e \sim N(0, \sigma^2)$,, where *Ret* represents excess return, *Vol* stands for trading volume, *Size* expresses the natural logarithm of the firm's market value of equity and *BM* indicates the natural logarithm of the book value of equity to the market value of equity. Moreover, Ret_{2-3} and Ret_{4-6} record the returns in previous periods, and X covers a set of other variables added into the model, such as yield, firm beta, market beta and firm size. I initially included the X variables in my estimation as well, but they appear insignificant, so I run the final estimation without them.
Bivariate VAR approaches comprise the second group of models. They again have small differences among themselves with respect to the number of lags included. All the VAR models belong to one group, since different bivariate VAR models produced similar results. VAR models in their estimation use, for example, Saatcioglu & Starks (1998), Lee & Rui (2002) and Ciner (2002) use. The baseline VAR equation is as follows:

$$\begin{aligned} Ret_t &= \alpha_0 + \alpha_1 Vol_t + \alpha_2 Vol_{t-1} + \alpha_3 Ret_{t-1} + e_t, e_t \sim N(0, \sigma^2), \\ Vol_t &= \beta_0 + \beta_1 Ret_t + \beta_2 Ret_{t-1} + \beta_3 Vol_{t-1} + u_t, u_t \sim N(0, \sigma^2). \end{aligned} \tag{4.5}$$

The *Simple* model group is a similarly large group, used in 21% of all the estimates. This group uses the following equation: $Ret_t = \alpha_0 + \alpha_1 Vol_t + e_t, e_t \sim$

$N(0, \sigma^2)$. The *simple* model employ, for example, Shu *et al.* (2004) or Tapa & Hussein (2016). In some cases, GARCH improves the variance equation by capturing heteroskedasticity: $h_t = \sigma_t^2 = \beta_0 + \beta_1 e_{t-1}^2 + \beta_2 h_{t-1}$. These models, used by Sana Hsieh (2014) and Tahir *et al.* (2016), among others, represent the fourth and last model group.
Another differentiation is based on the estimation methodology. Epps & Epps (1976) suggest that *OLS* estimates may have an upward bias; thus, they estimate the equations not only by *OLS* but also by maximum likelihood estimation (*MLE*). Beyond these two estimation techniques, newer articles prefer generalized method of moments (*GMM*) estimation. Among these are the papers of, for instance, Lee & Rui (2002) and Ciner (2003). Some articles do not mention the estimation technique; thus, these three groups are supplemented by the *Other methods* category when the method used is not clear or specified.
Moreover, some estimates relate to Monday trading only (Pisedtasalasai & Gunasekarage 2007). After weekends, stock markets are supposedly calmer (French 1980; Gibbons & Hess 1981). Therefore, these estimates are designated with the dummy variable *Monday*. A similar effect relates to January estimates in monthly data (Hu 1997). Thus, the January effect is joined with the Monday effect into one dummy. *Monday* estimates are expected to be more significant. On the other hand, *January excluded* estimates are also distinguished. Finally, I control for whether the primary data are *Trimmed*.

Publication characteristics. Last but not least among the variables capturing differences in the primary estimates are those related to publications. The impact of the quality of any publication may be studied from several perspectives. I employ three of them. One may expect these factors to be correlated with the unobserved features of the paper. In addition, these three variables are useful for the detection of potential publication bias. This fact increases their importance for the study. With the dummy variable *Published*, I obtain a systematic overview of whether published studies display different results from those in unpublished articles. It helps me to clarify whether the journals tend to publish some kind of results as it is discussed in the Section 4.3.
Furthermore, I control for quality across published and unpublished studies. The quality of the outlet distinguishes the *Impact factor* variable represented by the discounted recursive RePEc impact factor of the primary study. The advantage of its use arises from its availability for both working paper series and journals. This variable has been used in previous meta-analyses, for example,

Valickova *et al.* (2015a), who find it the variable *Impact factor* positive and significant in area of financial development and economic growth.

As the last publication characteristic, I choose the logarithm of the number of Google Scholar *Citations* normalized by the number of years since the first version of the study appeared. This reflects each article's relevance in the literature. This variable has been used among others by Bajzik *et al.* (2020) again with positive and significant signs in meta-analysis of Armington elasticity. Thus, I want to search whether the financial markets are likewise affected.

Besides one may suppose that effects in top-tier journals may be more rigorous than the estimation from lower quality journals or working papers (Astakhov *et al.* 2019). Thus, in one of the robustness checks I add the dummy variable *Top* for estimates from top-tier journals instead of the variables *Published*, *Impact factor*, *Citations*. The variable *Top* equals one, if the estimate was published in journal with article influence score the same or higher as the Financial Analysts Journal has. I discard the three mentioned ones, since their information and the information from the *Top* variable might be highly correlated or doubled.

4.4.1 Model Averaging

Model averaging techniques account for the outlined heterogeneity. Namely, the analysis deploys Bayesian model averaging in several specifications together with frequentist model averaging. Model averaging approaches have several merits in comparison with best-model approaches. First, they address model uncertainty in a systematic manner. Second, they deal with potential problems arising from mental conflict when one is faced with several competing model specifications. Third, they treat endogeneity problem and omitted variable bias methodically. Fourth, the model averaging techniques reveal the importance and magnitude of every collected control variable and thus compare all possible heterogeneity drivers qualitatively.

Raftery (1995) and Raftery *et al.* (1997) pioneered the deployment of BMA in social sciences. The widespread usage of the BMA approach, even in economics, is testified to in a summarizing article written by Moral-Benito (2015). In contrast, the usage of FMA in economics does not have such a long history. This branch of techniques was thoroughly described just a decade ago by Magnus *et al.* (2010) and Amini & Parmeter (2012). In economics, its usage has developed only recently (e.g., Havranek *et al.* 2017; Steel 2020; Gechert *et al.* 2020; Bajzik *et al.* 2020; Ehrenbergerova *et al.* 2021).

The base equation for both model averaging techniques consists of regressing an estimate on its standard error plus on the set of all control variables. Equation 4.6 clarifies the approach:

$$r_{ij} = \alpha_0 + \beta_0 SE + \sum_{k=1}^{39} \beta_k X_{k,ij} + e_{ij}, \tag{4.6}$$

where SE represents the standard error of the primary estimate and $X_{k,ij}$ is the value of the kth explanatory variable for the ith estimate from the jth study. Based on the definition, model averaging techniques do not exclude any variables in advance. This fact is of considerable importance when the aim is to explain heterogeneity among the studies. In my case, model averaging could potentially imply running 2^{39} regressions stemming from all the possible model combinations. Since such a process would be time consuming, BMA deploys a Markov chain Monte Carlo process with the Metropolis-Hastings algorithm (Zeugner & Feldkircher 2015) to avoid it.

This algorithm walks through the most probable models and assigns a posterior model probability (PMP) to each of them. The PMP expresses the probability of employment of the particular model. Based on the different PMPs, the posterior inclusion probability (PIP) of each variable arises. The PIP is a weighted average of the estimated coefficient of the variable, where the weights are the PMPs of the models. In comparison, FMA uses the orthogonalization of the covariate space (Amini & Parmeter 2012).

The easily interpretable posterior model probabilities and posterior inclusion probabilities make BMA preferable to FMA (Steel 2020). These statistics show more information than the simple point estimates with confidence intervals from FMA. Moreover, the scale at which one performs BMA does not matter based on the transformation invariance of the approach. This represents the second distinctive advantage of BMA over FMA (Fletcher 2018).

Both model averaging techniques require the setting of some prior knowledge. In the baseline BMA setting, I prefer the unit-information g-prior suggested by Eicher *et al.* (2011). It assigns each model the same prior weight and hence provides a convenient setting when there is a lack of knowledge of the parameter values. In addition to the g-prior, BMA requires model prior setting. Due to the small sample size, I prefer the dilution prior recommended by George (2010) and by Hasan *et al.* (2018).

It multiplies the prior model probabilities by the determinant of the model's correlation matrix. When the considered model is highly collinear, the determinant goes to zero. Thus, the model is given a small weight. For models with little collinearity, the opposite holds. Thus, the dilution prior deals with potential collinearity problems.
The robustness checks examine several BMA setting alternatives. Namely, I examine the data by combining a uniform g-prior with a uniform model prior and BRIC g-prior with a beta-binomial random model prior Fernandez *et al.* (2001); Ley & Steel (2009). Last, Mallowâ€™s criterion for model averaging (Hansen 2007) is used to deal with prior knowledge in the FMA setting.

4.4.2 Results

Turning to the results, an early presentation of the BMA conclusions appears in Figure 4.5. The model ordering goes from left to right, from the most significant to the least significant. The PMP of each model is captured by the corresponding column width. In a similar manner, the rows sort the explanatory variables. The variables with the highest PIP appear at the top. The cells at the nexus of the rows and the columns capture the effect that a variable has in a particular model. Red indicates a negative effect on the coefficient of interest and blue a positive effect, and the cell remains blank when the model does not include the variable. Hence, the stable red and blue variables are the ones of main interest.
Table 4.7 converts the Figure 4.5 into numbers. The PIP is now expressed in decimal numbers. According to Eicher *et al.* (2011), decisive variables are those with a PIP between 0.99 and 1, strong variables those with a PIP between 0.95 and 0.99, substantial variables those with a PIP from 0.75 to 0.95 and weak variables those with a PIP in the range of 0.5 to 0.75. Furthermore, the OLS frequentist check on at least weak variables completes the table.
Finally, I classify seven variables (without intercept) as decisive. Moreover, these variables display high stability, as is clear in Figure 4.5. Of decisive importance for the trading volume-stock return relationship are the *Abnormal returns*, *Data size*, *Midyear*, *Monthly* data frequency, *VAR* model, *Other methods* and *Other continents* variables. Altogether, variables from three of the four categories (data characteristics, structural variation and estimation technique) drive the differences across the estimated coefficients. The main discussion of the results focuses on these difference-makers.

Data characteristics. This category is represented the most represented among the decisive variables. The most important findings relate to data frequency and data type. The estimates originating from *Monthly* data are significantly lower than those based on higher-frequency (*Hourly, Daily, Weekly*) data. The difference between the average *Hourly* and *Monthly* estimates is, ceteris paribus, -0.168 (0.028). This outcome reveals data frequency to be the major driver of heterogeneity. This result is completely in line with conclusions of Kim *et al.* (2019). These authors after studying Asian ans Australasian market suggest that lower-frequency data are subject more to market frictions.
Next, the decreasing value of the estimates with the newer data has already been indicated in Figure 4.2 and in the discussion of the so-called "Prometheus effect" by Ioannidis (2008) and Kim *et al.* (2019). A higher volatility of *Abnormal returns* caused by differences in trading volume is anticipated by Lin & Liu (2017). The same holds for the sign and magnitude of *Absolute returns*, which are completely in line with the predictions of (Lin & Liu 2017; Mahajan & Singh 2009a). On the other hand, *Excess returns* by definition should behave similarly to *Absolute returns*, but the results contradict this idea. Indeed, Table 4.7 suggests the opposite: the *Excess* variable has a negative impact.

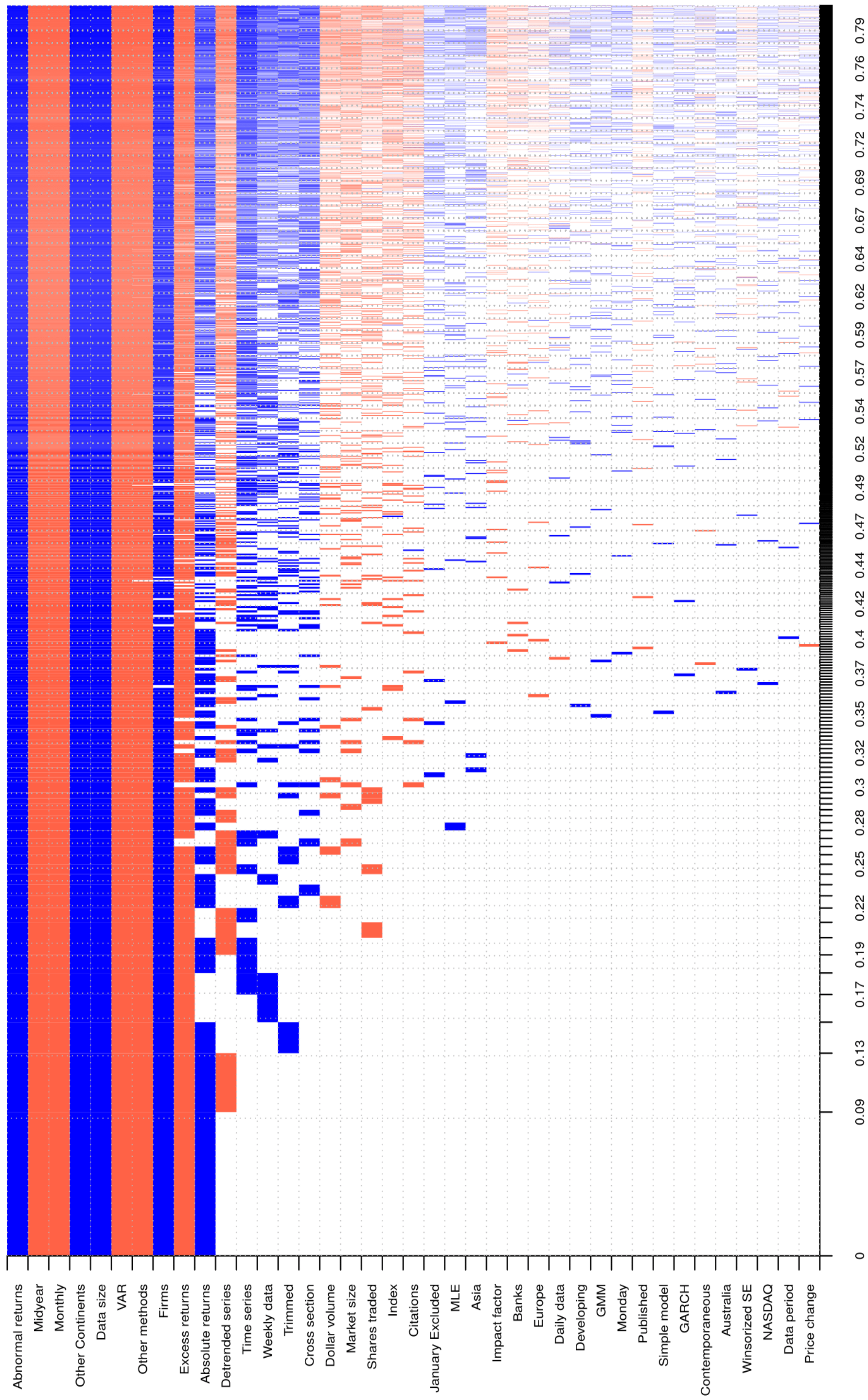


Figure 4.5: Bayesian model averaging

Structural variation. In line with the intuition of De Santis *et al.* (1997), the results show that emerging markets, especially those in Latin America and Africa, exhibit higher volatility and a greater probability of large price changes. The coefficients of 0.272 and 0.045 emphasize this effect. Similar conclusion draw Kim *et al.* (2019) for Asian and Australasian markets.
Moreover, stock markets across North America, Europe and Australia display no differences at all. This indicates that stock exchanges in emerging markets behave differently from those in developed markets regardless of the size of the market. The similar efficiency of these markets was expected.
Furthermore, the results contradict the conclusions of Chordia *et al.* (2001) about the stronger effect of trading volume on stock returns on the *NASDAQ* stock exchange. *NASDAQ* stocks, together with *Banks* and *Index* stocks, exhibit similar results to those based on *All* stocks from a given stock exchange. On the other hand, *Firms* stocks rise significantly higher (0.067). This finding uncovers previously hidden dynamics.

Table 4.7: Explaining heterogeneity – BMA dilution prior and frequentist check

	BMA – Dilution prior			Frequentist check – OLS		
Response variable:	Post. Mean	Post. SD	PIP	Estimate	SE	p-value
Constant	0.155	NA	1.000	0.137	0.017	0.000
Winsorized SE	-0.001	0.029	0.016			
Data characteristics						
Contemporaneous	0.000	0.002	0.017			
Price change	0.000	0.004	0.014			
Absolute returns	0.040	0.037	0.598	0.074	0.024	0.002
Abnormal returns	0.197	0.039	1.000	0.188	0.107	0.080
Excess returns	-0.073	0.030	0.908	-0.081	0.044	0.064
Dollar volume	-0.005	0.015	0.122			
Shares traded	-0.005	0.015	0.118			
Detrended series	-0.023	0.031	0.407			
Data period	0.000	0.000	0.014			
Data size	0.020	0.003	1.000	0.020	0.003	0.000
Midyear	-0.054	0.012	1.000	-0.049	0.011	0.000
Daily data	0.000	0.005	0.023			
Weekly data	0.016	0.034	0.226			
Monthly	-0.166	0.028	1.000	-0.166	0.042	0.000
Time series	0.024	0.036	0.358			
Cross section	0.023	0.052	0.194			
Structural variation						
Index	-0.005	0.018	0.087			
NASDAQ	0.000	0.005	0.014			

Continued on next page

Banks	-0.001	0.010	0.031			
Firms	0.067	0.023	0.943	0.080	0.024	0.001
Developing	0.000	0.003	0.022			
Market size	-0.001	0.004	0.121			
Asia	0.001	0.004	0.036			
Europe	0.000	0.006	0.024			
Australia	0.000	0.006	0.017			
Other Continents	0.272	0.045	1.000	0.293	0.049	0.000
Estimation technique						
VAR	-0.117	0.026	0.998	-0.134	0.027	0.000
Simple model	0.000	0.003	0.018			
GARCH	0.000	0.006	0.018			
Monday	0.000	0.006	0.020			
Trimmed	0.011	0.024	0.210			
January Excluded	0.001	0.007	0.041			
MLE	0.001	0.009	0.037			
GMM	0.000	0.006	0.021			
Other methods	-0.091	0.020	0.993	-0.098	0.015	0.000
Publication characteristics						
Impact factor	0.000	0.003	0.036			
Citations	-0.001	0.005	0.082			
Published	0.000	0.003	0.019			
Studies	44			44		
Observations	468			468		

Notes: Response variable = partial correlation between trading volume and stock returns. SD = standard deviation. PIP = posterior inclusion probability. SE = standard error. The UIP g-prior and dilution model prior are deployed in BMA, as suggested by George (2010). The frequentist check (OLS) includes only variables with PIP>0.5 to form the best model. SEs are clustered at the study and country levels, as proposed by (Cameron *et al.* 2012). Table 4.6 describes all variables used.

Estimation techniques. Another potential explanation of the heterogeneity in the estimates from the primary studies relates to the use of the various estimation techniques. *VAR* models provide substantially lower results (-0.117) than other estimation techniques. This outcome should be taken into account by anyone considering employing a VAR model in a future analysis.

Similarly, the usage of *other methods* produces more strongly negative estimates than those produced by *OLS*, *MLE*, or *GMM*. The results do not support questions over the reliability of OLS estimates and thus contradict the concerns raised by Epps & Epps (1976). Moreover, neither trimming the primary dataset nor considering the Monday or January effect impacts the primary findings, contradicting the predictions of French (1980); Gibbons & Hess (1981); Hu (1997).

Publication characteristics. The results indicate no strong association between publication characteristics and the magnitude of the reported results. The number of citations, the impact factor of the series and publication in a peer-reviewed journal do not substantially affect the results. This conclusion is in line with different findings for contemporaneous and dynamic volume-return relationship from Section 4.3. This conclusion does credit to journals as well as authors. The financial markets studies do not follow those from financial development, or Armington elasticity fields as indicated Valickova *et al.* (2015a); Bajzik *et al.* (2020).

4.4.3 Robustness check

The stability of the results underscores the complex robustness checks. Even Table 4.7 provides a glimpse of the robustness of the results. I run simple OLS regressions on the variables with a PIP over 0.5 in the BMA baseline results. The OLS coefficients accord with the baseline results in both sign and magnitude. Moreover, the variables remain significant according to their t-statistics as indicated by the PIPs in BMA.
In addition, the robustness check deploys several different BMA specifications. The combinations of the UIP g-prior with the uniform model prior and the BRIC g-prior with the random model prior are chosen. A comparison of these two specifications with respect to the baseline setting of the UIP g-prior with dilution model prior depicts Figure 4.6. The results further captured numerically in Table B2 indicate stable PIPs across the priors.
The UIP estimation removes the significance of *Absolute returns* and elevates the PIP of the *Detrended* series and *time series* data above 0.5. This indicates that the use of *Detrended* series lowers the estimates by -0.035. On the other hand, the use of *Time series* data increases the estimates by 0.042. This means that detrended series yield a less pronounced relationship between returns and volume, which is what one should expect from detrending. The *Time series* effect acts as a counterweight to *Data size*, the effect of which grows with *Panel* data. The BRIC estimation duplicates the baseline dilution conclusions entirely.

Table B2 provides even the next two robustness checks. First of them is the FMA. Even the results of this robustness check line up with the baseline results. They also suggest the significance of *Detrended* series (a negative effect), *Trimmed* data (a positive effect) and the *January excluded* variable (a negative effect). The negative coefficient of the *Trimmed* variable indicates the presence of negative outliers in the data of the primary articles. The significance of *January excluded* captures investor sentiment in January. This result contradicts (Datar *et al.* 1998)'s finding of no effect of sentiment in January.

The last but not least robustness check included in Table B2 is again BMA model. This one uses the dilution prior model setting as the baseline specification, but the publication characteristics variables are captured differently. In this last specification I employ instead of variables *Impact factor*, *Citations* and *Published* the variable *Top* signalizing that the primary estimate was published in the top-tier journal. But in the analysis even this variable turns insignificant. Other variables remain stable.

This model averaging technique suggests the same conclusions as BMA except that the importance in the data characteristics category shifts from the data type variable to that of measurement of returns and volume. The main findings regarding the significance of the standard error effect and the effects of *Abnormal returns*, *Excess returns*, usage of *Monthly* data, *Data size*, *Midyear* of the data, *VAR* models, *Other methods* and *Other continents* remain the same. This validates the robustness of the baseline model.

In summary, the robustness checks testify to the decisive importance of the choice of type data and country of interest. Moreover, the results provide substantial evidence of the decisiveness of the effects related to *Abnormal returns*, *Other continents*, *VAR* models, and *Other methods*. All these findings are evident across the models and specifications.

4.4.4 Implied Effect of Trading Volume on Stock Returns

What are the implications for the effect of trading volume on stock returns? Although the estimation unveils several drivers of heterogeneity in the estimates, the question about the "true" underlying effect remains open. Since the conducted estimations suggest several key factors determining stock returns, I can propose a preferred estimation specification for future research based on current knowledge.

The presented results indicate the following: i) publication bias affects the results on the contemporaneous effect but not those on the dynamic effect; ii) the results differ based on the type of stocks and continent of origin; and iii) different data employed and estimation methodology used cause some differences

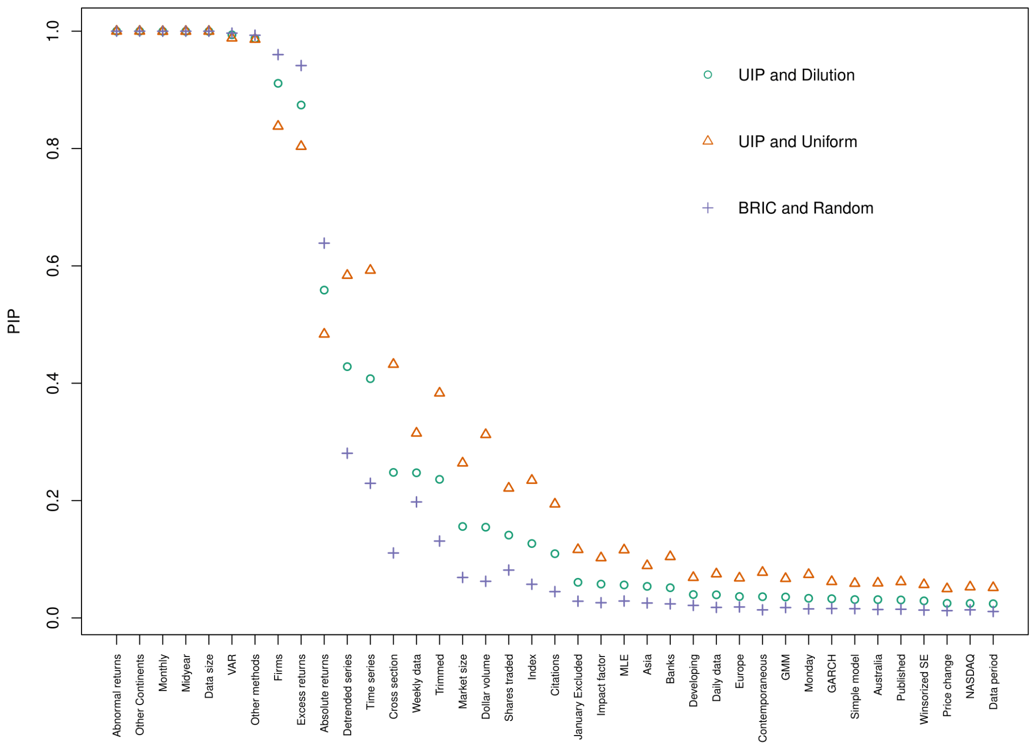


Figure 4.6: Bayesian model averaging - a comparison of the different priors

in findings. I consider all of these major findings when suggesting the "true" effect.

First, I choose workhorse models in this area to provide the baseline model and data specification settings. I choose the models of Brennan *et al.* (1998) and Chordia *et al.* (2001) for the dynamic estimates and that of Datar *et al.* (1998) for the contemporaneous ones. These are perhaps not the newest studies, but they remain seminal in this area of research, and their findings have not yet been overturned. Furthermore, they have been published in journals with high-quality peer review (the first two in the Journal of Financial Economics and the third in the Journal of Financial Markets).

The implied estimates captured in Table 4.8 and 4.9 are based on a linear combination of the model characteristics from these three papers except in regard to the variables distinguishing between dynamic and contemporaneous effects and the factors related to structural heterogeneity. Besides for the data frequency I choose daily data instead of monthly, since monthly data proves to be not the convenient throughout my analysis. Moreover, in the case of the standard error, the coefficient is zero, indicating no presence of publication bias.

The characteristics from Brennan *et al.* (1998), Chordia *et al.* (2001) and Datar *et al.* (1998) correspond to the dynamic and contemporaneous effects, and the estimates by continent indicate whether the continent is developing (*Asia* and *other continents*, that is, Africa and Latin America) or developed (*Europe*, *North America*, and *Australia*). The 90th percentile settings for *market size* and *midyear* indicate that newer datasets are preferred. Furthermore, stock characteristics serve as the last but not the least important input in the implied estimates.

Table 4.8: Implied estimates – contemporaneous effect

			Stock type		
Continent	All	Index	Banks	Firms	NASDAQ
North America	0.113	0.071	0.062	0.135*	0.138
	(0.091)	(0.087)	(0.074)	(0.069)	(0.089)
Europe	0.124	0.082	0.074	0.147*	
	(0.100)	(0.096)	(0.079)	(0.081)	
Australia	0.124	0.082	0.074	0.147	
	(0.123)	(0.122)	(0.105)	(0.109)	
Asia	0.123	0.081	0.072	0.145*	
	(0.095)	(0.092)	(0.78)	(0.076)	
Other Continents	0.431***	0.389***	0.380***	0.453***	
	(0.137)	(0.128)	(0.130)	(0.121)	
Best practice	0.116				
	(0.092)				

Notes: The values suggest the best practice estimates of trading-volume relationship across different continents and stocks implied by study design of Datar *et al.* (1998). Standard errors reported in parentheses are derived from OLS estimates and clustered at the study and country level as suggested Cameron *et al.* (2012)). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The implied estimate results reveal two points. When one considers only overall best practice estimates, there is a negligible and insignificant contemporaneous and even a negative dynamic effect of trading volume on stock returns. But after a closer look at particular elements of structural heterogeneity, important findings for developing profitable trading strategies appear. First, there is a positive significant effect of trading volume on stock the *Firms* stocks that ranges from 0.135 to 0.147 for the contemporaneous case.

Second, even the conclusions on individual continents vary. The developing

Table 4.9: Implied estimates – dynamic effect

			Stock type		
Continent	All	Index	Banks	Firms	NASDAQ
North America	0.085	0.043	0.035	0.107	0.110
	(0.087)	(0.089)	(0.073)	(0.069)	(0.085)
Europe	0.096	0.054	0.046	0.119	
	(0.097)	(0.099)	(0.078)	(0.081)	
Australia	0.097	0.055	0.046	0.119	
	(0.0125)	(0.128)	(0.109)	(0.112)	
Asia	0.095	0.053	0.045	0.117	
	(0.092)	(0.094)	(0.077)	(0.075)	
Other Continents	0.403^{***}	0.361^{***}	0.353^{***}	0.426^{***}	
	(0.132)	(0.126)	(0.126)	(0.117)	
Best practice	0.088				
	(0.089)				

Notes: The values suggest the best practice for estimating the trading volume relationship across different continents and stocks, as implied by the study design of Brennan *et al.* (1998) and Chordia & Swaminathan (2000). Standard errors, reported in parentheses, are derived from OLS estimates and clustered at the study and country level, as suggested by Cameron *et al.* (2012)). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

continents, especially South America and Africa, display a significant effect of trading volume on stocks. The effect is positive and significant in both the contemporaneous and dynamic cases across all stock types. The implied estimates are all significant and range from 0.353 to 0.453. This indicates that these stock markets are still evolving and provide more arbitrage opportunities. These findings clarify why the results in the literature on stock returns and trading volume diverge. Authors use different data from different countries, and naturally this leads to different conclusions. Thus, one should bear it in mind when proposing conclusions that for different countries in varying circumstances, findings differ widely.

4.5 Conclusions

The first quantitative synthesis of the broad economic literature on the relationship between trading volume and stock returns comes into existence. This relationship, crucial for building profitable trading strategies, conducting event studies, futures markets investigation or for confirming efficient market hypothesis, subdues thorough examination. A total of 468 estimates collected from 44 studies reveal several significant outcomes.

First of all, I investigate the publication bias by common approaches such as a funnel plot and OLS and WLS estimations. Then I deploy recently developed nonlinear estimators proposed by (Furukawa 2019) and Bom & Rachinger (2019). The study indicates the presence of publication bias, at least in cases when the investigated relationship is contemporaneous. The mean effect after correction for publication bias has a negligible value. Furthermore, based on a a caliper test (Gerber *et al.* 2008; Gerber & Malhotra 2008), even the p-hacking and HARKing affect the results from primary studies.

Moreover, other study-specific factors affect the corrected mean. For this heterogeneity investigation, Bayesian (Raftery 1995) and frequentist (Amini & Parmeter 2012) model averaging is deployed. The results show that data characteristics, structural variation and different methodological approaches explain a large part of the inconsistency in the primary results. Concretely, one has to be cautious in using *Monthly* data or *VAR* models. These variables are associated with a substantially more negative effect of trading volume on returns.

Moreover, the type of stock and country of estimate matter. For instance, the trading volume may predict the development of the stocks of firms. On the contrary, the effect of trading volume on the stocks of the banks or overall indexes turns to be insignificant. These conclusions remain the same even across the continents. The same findings hold for markets in the North America, Europe, Australia and Asia. Just estimates for developing countries differ. Trading volume predicts the stock return on emerging markets better than on the developed ones. So on the developing markets one can partially predict via trading volume development of return of any kind of stocks. Final observation I have is the one regarding the journal quality. The publication characteristics of the article do not affect the results.

All in all, one should bear in mind the specifics of each stock type and avoid dangerously relying on some overall conclusion. All these results can serve as a baseline for model calibration or can directly help in traders' strategies.

Chapter 5

Conclusion

In my thesis I study topics in international trade, macroprudential policy, and financial markets. I use a meta-analytical approach that allows me to see the entirety of the literature in its full context, comparing inter-study differences, and helps me understand the topic fully in both historical and recent context. The conclusions I draw have the potential to help the entire economics research community with model calibration, policy implementation, and market understanding.

In the first thesis I together with my colleagues from Charles University focus on 50 years' worth of research into Armington elasticity, a key parameter in international trade (Armington 1969). This study was inspired by the discussion on the welfare effect of globalization, trade balance adjustments, and the exchange rate pass-through of monetary policy, all of which are heavily affected by Armington elasticity (Costinot & Rodriguez-Clare 2014; Imbs & Mejean 2015; Auer & Schoenle 2016). For example, Schurenberg-Frosch (2015), who recomputes the results of 50 previously published models using different values of the elasticity, finds that with plausible changes in this value, the results change qualitatively in more than half of all cases. Besides, Hillberry & Hummels (2013) state that the Armington elasticity is without exaggeration the most important parameter in modern trade history. Therefore, given the 50-year history of research and the importance of the parameter, I believe that meta-analysis was the ideal tool for this context. It has allowed me to work with results from the entire history of research into the parameter, which I could then assess using the newest methodologies. Hence, I was able to avoid repeating the same mistakes that plagued the historical studies, and base my results on the most recent methods and approaches. I find evidence of strong publication bias in the estimates of the long-term Armington elasticity, resulting in an exaggeration of the mean estimate by more than 50%. Our findings indicate that study characteristics are systematically associated with reported results. These characteristics were examined only on long-run estimates of the sample, so there is no danger that the short-run and long-run estimates are correlated in this part of the analysis. Besides we use the dilution prior in the BMA alleviate the possible multi-collinearity problem greatly, as it gives less weight to more collinear variables. Among the 32 variables we constructed, the most important in model averaging are the ones related to the data used in the estimation: data frequency (monthly, quarterly, or annual), data dimension (time series, cross-section, or panel), and dataset size. We also find a systematic correlation between quality measures and the reported magnitude of the elasticity. These findings are of major importance – the Armington elasticity is used in analyzing the welfare effects of globalization, trade balance adjustments, and the exchange rate pass-through of monetary policy. Any attempt to evaluate the effect of tariffs also depends crucially on Armington elasticity. For example, Engler & Tervala (2018) mention that changing the elasticity value from 3 to 8 more than doubles the estimated welfare gains from the Transatlantic Trade and Investment Partnership.

In the second chapter, I together with my colleagues from Czech National Bank delved into macroprudential policy, namely the relationship between bank capital and lending. Quantifying this relationship has been among the most pivotal research questions in the field for almost two decades. The topic was given particular attention following the onset of the 2007-2009 Global Financial Crisis (GFC), when the likelihood of a credit crunch was under debate, and again when the first quantitative easing programs were being implemented. The question has reemerged more recently with the gradual implementation of Basel III and an increasing use of macroprudential policy instruments. Following the implementation of Basel III, the observed minimum capital requirements effectively rose from 8% to 10.5%. However, due to all the additional prudential buffers, the capital requirements were able to reach as high as 20% (BCBS 2010). This study add to the recent economic and academic discussion Peek & Rosengren (1997); Houston *et al.* (1997); Berrospide & Edge (2010b); Gambacorta & Marques-Ibanez (2011a); De Jonghe *et al.* (2020). The discussion brings a wide range of possible outcomes when quantifying the impact of changes in bank capital on bank lending. On the one hand, an increase in the bank capital ratio due to the introduction of a new capital regulation may dampen bank lending activities as a bank would try to avoid the higher costs of financing loans by capital (De Jonghe *et al.* 2020). On the other hand, a general increase in the bank capital (equity) ratio due to, for example, bank profit accumulation should be reflected in an increase in lending, suggesting a positive effect (Berrospide & Edge 2010b). In this study we also rely on meta-analytical tools, which allow me compare the existing literature both qualitatively and quantitatively, while allowing us to control for different political conditions in different countries. My colleagues and I find that various study characteristics are systematically associated with the reported results. Among the 40 variables we construct, the most important for model averaging are those related to data, the estimation technique and cross-country or regional differences. Moreover, we tried to capture the supply and demand sides of the respective regions and periods as close as possible. Thus, we include in the analysis external factors such as spread, housing price growth, monetary policy rate or unemployment. We believe we captured the surroundings as closely as possible, but there is still the question of whether we cannot capture the demand side in a better way. We find that single-country studies with larger sample sizes positively correlate with the collected semi-elasticities, while studies shielded from omitted variable bias with more favorable publication characteristics are generally negatively correlated with the reported estimates. Apart from data characteristics, estimates of the effect of changes to the simple capital-to-asset ratio are also dependent on the variables reflecting the macro-financial characteristics of the countries analyzed. The heterogeneity in the estimates based on the regulatory capital ratio can thus mostly be explained by model specification. In the case of the literature on capital requirements, the standard error is the most important variable in explaining the variation in the reported estimates. Large standard errors are associated with more negative estimates, supporting the existence of publication bias in this category. We perceive the contribution of this paper to be threefold. First, quantifying the effect of changes to bank capital on the supply of credit is of utmost importance to policymakers. Obtaining a comprehensive overview of the literature's findings goes well beyond the scope of individual studies that are, by nature, very selective. Second, we show the caveats associated with modelling the relationship between bank capital and lending and inform about the most commonly employed practices. Third, we present some indications that the relationship is changing over time and discuss the implications this might have for correctly estimating and assessing the impact of capital regulation.

In the third and last chapter, I focused on the trading volume and returns relationship. Studying the return-volume relationship is of both theoretical and practical interest. While some of the authors doubt the importance and direction of the relationship (Lee & Rui 2000; 2002), citebeaver1968 writes: “An important distinction between the price and volume tests is that the former reflects changes in the expectations of the market as a whole, while the latter reflects changes in the expectations of individual investor.” Morgan (1976) continues with the suggestion that volume is connected with systemic risk and thus with stock returns. Trading volume thus enters the discussion of stock returns beside the well-known factors proposed by Fama & French (1992; 1993; 1996) and Jegadeesh & Titman (1993; 1995). Moreover, Karpoff (1987) added several other reasons to study the trading-volume effect. First, this type of research provides insight into financial market structure. Second, it is seminal for event studies, which use price and volume data to draw conclusions. Third, the return-volume relationship has significant implications for futures markets research. These factors make findings in this area even more valuable. And since many new traders have been entering the market via online platforms - at least during the coronavirus crisis – studying the return-volume relationship remains in the forefront of interest. For example, Chiah & Zhong (2020) found a large spike in trading volume in 37 international equity markets during the coronavirus crisis. During my study I found the topic to be very broad. Thus, I decided to once again employ meta-analysis. Specifically, the partial correlation coefficients technique allows me to compare otherwise incomparable estimates, though at the cost of losing the economic interpretation of the results (the statistical significance remains). It is a common approach in the meta-analysis of broad topics, as shown by Valickova *et al.* (2015b). With this data transformation I found publication bias, at least in the more recent papers. The mean after correction for publication bias has a negligible value. Moreover, other study-specific aspects affect the corrected mean. The results of both Bayesian model averaging (BMA) and frequentist model averaging (FMA) indicate that data characteristics, structural variation, and different methodological approaches explain a large part of the inconsistency in the primary results. For example, usage of monthly data or VAR models makes the effect of trading volume on returns substantially more negative. Other sources of variation include the type of stock and country of origin. For instance, my analysis shows that trading volume may predict the stock returns of non-financial firms. On the other hand, the effects of trading volume on bank stocks or general indices turn out to be insignificant. These conclusions were found worldwide. The same holds for markets in North America, Europe, Australia, and Asia. Only the estimates for developing countries differ. Trading volume predicts stock returns in emerging markets better than in developed ones. So, in developing markets, one can partially predict the returns development of any stock via trading volume. Thus, one should keep in mind the specifics of each stock when forming a portfolio, calibrating a model, preparing a trading strategy (Chiah & Zhong 2020), or conducting research, e.g. for futures markets (Karpoff 1987), or for changes in the expectations of individual investors (Beaver 1968). It remains open to discussion whether the direction of the causal relationship is from trading volume to returns or vice versa, whether the relationship is contemporaneous or lagged, or whether trading volume should be considered important in determining stock returns as proposed by the Fama-French models.

While writing and defending the thesis and the individual articles, several points came to my mind regarding the recent meta-analytical debates, as well as the issues and contributions of my work. For example, while I was collecting the data for each respective meta-analysis, I noticed that some authors published several articles within a short time period (e.g., Lundmark & Shahrammehr 2011a;b; 2012). While the focus of each of these studies differed somewhat (e.g. Armington elasticity for roundwood or biomass), there is still a substantial risk that the study design and approach would remain very similar. The question for me was whether to include all such papers, potentially giving undue weight to their shared approaches, or whether to discard some of them, introducing discretionary bias. I ultimately decided to include all papers, as the meta-analytic paradigm is to work with "all relevant articles". To mitigate the issue I rely on a dilution prior in the Bayesian model averaging to give less weight to such potentially collinear studies (as Bayesian statistics does not use fixed effects). Next, there is discussion in the field whether it is appropriate to include both working and published papers in the analysis, or to drop the working papers. Doucouliagos & Stanley (2013) suggest that results of meta-analyses working with published articles only do not statistically differ from those including both journals and working papers. Nevertheless, we tried to mine this issue for data in the second article, where we differentiated whether the working paper version of a later published article differed in the model used, the number of countries or time span, or in neither of those. Then we included all original estimates regardless of whether the journal version was included or not. Finally, we found no statistical difference, allowing us to conclude that in the area of banks, capital, and lending the results are not biased by editor or referee preference for a particular methodology or country and period framework. On the contrary, I believe that similar research should be included and expanded in other meta-analyses as well.

Next, I studied recent seminal articles about the central bank bias (Fabo *et al.* 2021). As two of my articles were already published when the discussion arose, so this issue is only included in our macroprudential policy paper. We tracked whether the original working paper was published under the name of a central bank or not (as a proxy for the authors' affiliation), but the variable turned out to be insignificant. We believe this is a sufficient proxy for the set of author characteristics. I have also studied Fabo *et al.* (2021; 2024); Weale & Wieladek (2022), wherein a discussion was carried on about the inclusion of outliers in meta-analysis. Unfortunately, by the time this debate concluded, all my research included in the dissertation was already published, so I was not able to incorporate their findings into my thesis. With respect to the proposed research question, I used meta-analytical approaches that represented the state of the art at the time for this type of application. For example, I used winsorization instead of trimming. Next, whenever I referred to a study for illustration, I used medians (e.g. per study, or per year) instead of means, to avoid possible suspicion of excessive visual manipulation of the data after winsorization. Still, from my understanding of the discussion between Fabo *et al.* (2021; 2024); Weale & Wieladek (2022), I agree that the topic of the outliers should be given more attention in future meta-analyses. Furthermore, I am aware of the fact that each of my studies only uses the "midpoint and "length" of the primary data concerning the time span and length of the primary articles, suggesting the relationship of the data midyear is linear. In reality, one might suggest that, for instance, credit reactions under Basel I, II, and III might be different and hence, separate dummies for each decade might fit the data better. I considered this issue deeply and made the change in several of my studies, but with results not significantly different from those published in the thesis. Thus, I ultimately decided not to include this change in my articles.

Moreover, I was confronted by the problem of endogeneity. Endogeneity should not be issue in the meta-analytical estimation itself, but might appear in the primary articles (Gambacorta & Shin 2018; Roulet 2018). It is thus possible that the final data for meta-analysis included observations affected by endogeneity. To my knowledge, the best approach to deal with endogeneity in the meta-analysis is to capture the respective estimation methodologies from primary articles with dummies. This helps the researcher to discern whether methodologies that tackle endogeneity directly (e.g. GMM) provide systematically different estimates than those which do not. In this way, one can determine whether and to what extent endogeneity affects the primary estimates. An alternative way to deal with endogeneity would be to capture via dummy whether the primary study dealt with endogeneity or not. However, I find it redundant to include this treatment when I already control for different methodologies.

In general, I have striven to work with state-of-the-art meta-analytical methodologies throughout my academic career. But the field is dynamic and there is always room for improvement. Developments in the field allow me to see my work more critically and discover potential for future improvement. New discoveries also open questions that do not always have direct answers, but still motivate me to seek ways of enhancing my meta-analytical research and consider its shortcomings and benefits.

Bibliography

Aiyar, S., C. W. Calomiris, & T. Wieladek (2014a): "Does Macro-Prudential Regulation Leak? Evidence from a UK Policy experiment." *Journal of Money, Credit and Banking* **46(s1)**: pp. 181–214.

Aiyar, S., C. W. Calomiris, & T. Wieladek (2016): "How Does Credit Supply Respond to Monetary Policy and Bank Minimum Capital Requirements?" *European Economic Review* **82**: pp. 142–165.

Aiyar, S. S., C. W. Calomiris, & T. Wieladek (2014b): "How Does Credit Supply Respond to Monetary Policy and Bank Minimum Capital Requirements?" *Boe working paper no. 508*, Bank of England.

Akinci, O. & J. Olmstead-Rumsey (2018): "How Effective Are Macroprudential Policies? An Empirical Investigation." *Journal of Financial Intermediation* **33**: pp. 33–57.

Akpansung, A. O. & M. O. Gidigbi (2015): "The relationship between trading volumes and returns in the nigerian stock market." *International Research Journal of Finance and Economics* **132**: pp. 150–163.

Akram, Q. F. (2012): "Macro Effects of Capital Requirements and Macroprudential Policy." *Working paper no. 2012/21*, Norges Bank.

Akram, Q. F. (2014): "Macro Effects of Capital Requirements and Macroprudential Policy." *Economic Modelling* **42**: pp. 77–93.

Al-Jafari, M. K. & A. Tliti (2013): "An empirical investigation of the relationship between stock return and trading volume: Evidence from the Jordanian banking sector." *Journal of Applied Finance and Banking* **3(3)**: pp. 45–64.

Alam, Z., A. Alter, E. Eiseman, G. Gelos, H. Kang, M. Narita, E. Nier, & N. Wang (2019): "Digging Deeper – Evidence on the Effects of Macroprudential Policies from a New Database." *Imf working paper no. 19/66*, International Monetary Fund.

Alaouze, C. M. (1977): "Estimates of the elasticity of substitution between imported and domestically produced goods classified at the input-output level of aggregation." *Centre of Policy Studies/IMPACT Project Working Paper o-13/1977*, Victoria University, Centre of Policy Studies/IMPACT Centre, Melbourne.

Alaouze, C. M., J. S. Marsden, & J. Zeitsch (1977): "Estimates of the elasticity of substitution between imported and domestically produced commodities at the four digit ASIC level." *Centre of Policy Studies/IMPACT Project Working Paper o-11/1977*, Victoria University, Centre of Policy Studies/IMPACT Centre, Melbourne.

Albrizio, S., S. Choi, D. Furceri, & C. Yoon (2020): "International Bank Lending Channel of Monetary Policy." *Journal of International Money and Finance* **102**: p. 102124.

Almazan, A., A. Martín-Oliver, & J. Saurina (2015): "Securitization and Banks' Capital Structure." *The Review of Corporate Finance Studies* **4(2)**: pp. 206–238.

Altavilla, C., M. Boucinha, & J.-L. Peydró (2018): "Monetary Policy and Bank Profitability in a Low Interest Rate Environment." *Economic Policy* **33(96)**: pp. 531–586.

Amini, S. M. & C. F. Parmeter (2012): "Comparison of model averaging techniques: Assessing growth determinants." *Journal of Applied Econometrics* **27(5)**: pp. 870–876.

Anderson, E., M. A. J. d'Orey, M. Duvendack, & L. Esposito (2018): "Does Government Spending Affect Income Poverty? A Meta-regression Analysis." *World Development* **103(C)**: pp. 60–71.

Andrews, I. & M. Kasy (2019): "Identification of and Correction for Publication Bias." *American Economic Review* **109(8)**: pp. 2766–94.

Anundsen, A. K., K. Gerdrup, F. Hansen, & K. Kragh-Sørensen (2016): "Bubbles and Crises: The Role of House Prices and Credit." *Journal of Applied Econometrics* **31(7)**: pp. 1291–1311.

Araujo, J., M. Patnam, A. Popescu, F. Valencia, & W. Yao (2020): "Effects of Macroprudential Policy: Evidence from over 6,000 Estimates." *Imf working paper wp/20/67*, International Monetary Fund.

Armington, P. S. (1969): "A theory of demand for products distinguished by place of production." *IMF Staff Papers* **16(1)**: pp. 159–178.

Aspalter, L. (2016): "Estimating industry-level armington elasticities for EMU countries." *Working Paper 217/2016*, Department of Economics, Vienna University of Economics and Business, Vienna.

Assogbavi, T., J. Schell, & S. Fagnisse (2007): "Equity price-volume relationship on the Russian stock exchange." *International Business and Economics Research Journal* **6(9)**: pp. 107–116.

Astakhov, A., T. Havranek, & J. Novak (2019): "Firm Size And Stock Returns: A Quantitative Survey." *Journal of Economic Surveys* **33(5)**: pp. 1463–1492.

Auer, R. A. & R. S. Schoenle (2016): "Market structure and exchange rate pass-through." *Journal of International Economics* **98(C)**: pp. 60–77.

Auer, S., M. Ganarin, & P. Towbin (2018): "International Banking and Cross-Border Effects of Regulation: Lessons from Switzerland." *48th issue (March 2017) of the International Journal of Central Banking* .

Backus, D., P. Kehoe, & F. Kydland (1994): "Dynamics of the trade balance and the terms of trade: The j-curve?" *American Economic Review* **84(1)**: pp. 84–103.

Bahaj, S., J. Bridges, F. Malherbe, & C. O'Neill (2016): "What Determines How Banks Respond to Changes in Capital Requirements?" *Bank of england working paper no. 593*, Bank of England.

Bajzik, J., T. Havranek, Z. Irsova, & J. Schwarz (2020): "Estimating the Armington Elasticity: The Importance of Data Choice and Publication Bias." *Journal of International Economics* **127(C)**: p. 103383.

Balistreri, E. J., A. Al-Qahtani, & C. A. Dahl (2010): "Oil and petroleum product Armington elasticities: A new-geography-of-trade approach to estimation." *The Energy Journal* **31(3)**: pp. 167–179.

Basu, S. (1977): "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis." *The Journal of Finance* **32(3)**: pp. 663–682.

BCBS (2010): "An assessment of the long-term economic impact of stronger capital and liquidity requirements." *Report*, Bank for International Settlements: Basel Committee on Banking Supervision.

Beaver, W. H. (1968): "The information content of annual earnings announcements." *Journal of Accounting Research* pp. 67–92.

Beltratti, A. & R. M. Stulz (2012): "The Credit Crisis around the Globe: Why Did Some Banks Perform Better?" *Journal of Financial Economics* **105(1)**: pp. 1–17.

Berger, A. N. & C. H. Bouwman (2013): "How Does Capital Affect Bank Performance During Financial Crises?" *Journal of Financial Economics* **109(1)**: pp. 146–176.

Berger, A. N. & G. F. Udell (1994): "Did Risk-Based Capital Allocate Bank Credit and Cause a "Credit Crunch" in the United States?" *Journal of Money, Credit and Banking* **26(3)**: pp. 585–628.

Bergin, P. (2006): "How well can the new open economy macroeconomics explain the exchange rate and current account." *Journal of International Money and Finance* **25(5)**: pp. 675–701.

Bernanke, B. S., C. S. Lown, & B. M. Friedman (1991): "The Credit Crunch." *Brookings Papers on Economic Activity* **1991(2)**: pp. 205–247.

Berrospide, J., R. Correa, L. Goldberg, & F. Niepmann (2016): "International Banking and Cross-Border Effects of Regulation: Lessons from the United States." *Working paper 22645*, National Bureau of Economic Research.

Berrospide, J. M. & R. Edge (2010a): "The Effects of Bank Capital on Lending: What Do We Know, and What Does It Mean?" *Finance and* *economics discussion series no. 2010-44*, Board of Governors of the Federal Reserve System.

Berrospide, J. M. & R. M. Edge (2010b): "The Effects of Bank Capital on Lending: What Do We Know, and What Does It Mean?" *International Journal of Central Banking* **6(34)**: pp. 1–50.

Bilgic, A., S. King, A. Lusby, & D. F. Schreiner (2002): "Estimates of U.S. regional commodity trade elasticities of substitution." *Journal of Regional Analysis and Policy* **32(2)**: pp. 79–98.

Blanco-Perez, C. & A. Brodeur (2020): "Publication Bias and Editorial Statement on Negative Findings." *Economic Journal* **130(629)**: pp. 1226–1247.

Blonigen, B. A. & W. W. Wilson (1999): "Explaining Armington: What determines substitutability between home and foreign goods?" *The Canadian Journal of Economics* **32(1)**: pp. 1–21.

Bom, P. R. D. & H. Rachinger (2019): "A kinked meta-regression model for publication bias correction." *Research Synthesis Methods* **10(4)**: pp. 497–514.

Borio, C. & L. Gambacorta (2017): "Monetary Policy and Bank Lending in a Low Interest Rate Environment: Diminishing Effectiveness?" *Journal of Macroeconomics* **54**: pp. 217–231.

Borio, C., L. Gambacorta, & B. Hofmann (2017): "The Influence of Monetary Policy on Bank Profitability." *International Finance* **20(1)**: pp. 48–63.

Brandle, A. (2010): "Volume based portfolio strategies - analysis of the relationship between trading activity and expected returns in the cross-section of Swiss stocks." *Technical report*, University of St. Gallen.

Brei, M., L. Gambacorta, & G. von Peter (2011): "Rescue Packages and Bank Lending." *Bis working papers no. 357*, Bank for International Settlements.

Brei, M., L. Gambacorta, & G. Von Peter (2013): "Rescue Packages and Bank Lending." *Journal of Banking & Finance* **37(2)**: pp. 490–505.

Brennan, M. J., T. Chordia, & A. Subrahmanyam (1998): "Alternative factor specifications, security characteristics, and the cross-section of expected returns." *Journal of Financial Economics* **49**: pp. 345–373.

Bridges, J., D. Gregory, M. Nielsen, S. Pezzini, A. Radia, M. Spaltro *et al.* (2014): "The Impact of Capital Requirements on Bank Lending." *Boe working paper no. 486*, Bank of England.

Broda, C. & D. E. Weinstein (2006): "Globalization and the gains from variety." *The Quarterly Journal of Economics* **121(2)**: pp. 541–585.

Brodeur, A., N. Cook, & A. Heyes (2020): "Methods Matter: P-Hacking and Causal Inference in Economics." *American Economic Review* **(forthcoming)**.

Brodeur, A., M. Le, M. Sangnier, & Y. Zylberberg (2016): "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

Bruns, S. B., I. Asanov, R. Bode, M. Dunger, C. Funk, S. M. Hassan, J. Hauschildt, D. Heinisch, K. Kempa, J. König *et al.* (2019): "Reporting errors and biases in published empirical findings: Evidence from innovation research." *Research Policy* **48(9)**: p. 103796.

Bruns, S. B. & J. P. A. Ioannidis (2016): "p-Curve and p-Hacking in Observational Research." *PloS ONE* **11(2)**: p. e0149144.

Buch, C. M. & E. Prieto (2012): "Do Better Capitalized Banks Lend Less? Long-Run Panel Evidence from Germany." *Iaw diskussionspapiere no. 84*, Institute for Applied Economic Research.

Buch, C. M. & E. Prieto (2014): "Do Better Capitalized Banks Lend Less? Long-Run Panel Evidence from Germany." *International Finance* **17(1)**: pp. 1–23.

Cameron, A. C., J. B. Gelbach, & D. L. Miller (2012): "Robust inference with multiway clustering." *Journal of Business & Economic Statistics* .

Card, D., J. Kluve, & A. Weber (2018): "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* **16(3)**: pp. 894–931.

CARLSON, M., H. SHAN, & M. WARUSAWITHARANA (2011): "Capital Ratios and Bank Lending: A Matched Bank Approach." *Finance and economics discussion series no. 2011-34*, Board of Governors of the Federal Reserve System.

CARLSON, M., H. SHAN, & M. WARUSAWITHARANA (2013): "Capital Ratios and Bank Lending: A Matched Bank Approach." *Journal of Financial Intermediation* **22(4)**: pp. 663–687.

CASSONI, A. & M. FLORES (2008): "Methodological shortcomings in estimating Armington elasticities." *MPRA Working Paper 34544/2008*, Munich Personal RePEc Archive, Munich.

CASSONI, A. & M. FLORES (2010): "Armington elasticities: Estimates for Uruguayan manufacturing sectors." *Serie Documentos de Trabajo (working papers) 7/2010*, Departamento de EconomĂa, Facultad de Ciencias Sociales, Universidad de la RepĂşblica.

CAZACHEVICI, A., T. HAVRANEK, & R. HORVATH (2020): "Remittances and economic growth: A meta-analysis." *World Development* **134**: p. 105021.

CERUTTI, E., S. CLAESSENS, & L. LAEVEN (2017a): "The Use and Effectiveness of Macroprudential Policies: New Evidence." *Journal of Financial Stability* **28**: pp. 203–224.

CERUTTI, E., R. CORREA, E. FIORENTINO, & E. SEGALLA (2017b): "Changes in Prudential Policy Instruments – A New Cross-Country Database." *International Journal of Central Banking* **13(1)**: pp. 477–503.

CHANG, S. S. & F. A. WANG (2019): "Informed contrarian trades and stock returns." *Journal of Financial Markets* **42**: pp. 75–93.

CHEN, G.-m., M. FIRTH, & O. M. RUI (2001): "The dynamic relation between stock returns, trading volume, and volatility." *Financial Review* **36(3)**: pp. 153–174.

CHIAH, M. & A. ZHONG (2020): "Trading from home: The impact of COVID-19 on trading volume around the world." *Finance Research Letters* **37**: p. 101784.

CHINN, M. D. & H. ITO (2008): "A New Measure of Financial Openness." *Journal of Comparative Policy Analysis* **10(3)**: pp. 309–322.

Chordia, T., A. Subrahmanyam, & V. R. Anshuman (2001): "Trading activity and expected stock returns." *Journal of Financial Economics* **59(1)**: pp. 3–32.

Chordia, T. & B. Swaminathan (2000): "Trading volume and cross-autocorrelations in stock returns." *The Journal of Finance* **55(2)**: pp. 913–935.

Christensen, G. & E. Miguel (2018): "Transparency, Reproducibility, and the Credibility of Economics Research." *Journal of Economic Literature* **56(3)**: pp. 920–980.

Chuang, W.-I. & B.-S. Lee (2006): "An empirical evaluation of the overconfidence hypothesis." *Journal of Banking and Finance* **30(9)**: pp. 2489–2515.

Ciner, C. (2002): "The stock price-volume linkage on the Toronto stock exchange: Before and after automation." *Review of Quantitative Finance and Accounting* **19(4)**: pp. 335–349.

Ciner, C. (2003): "Dynamic linkages between trading volume and price movements: Evidence for small firm stocks." *The Journal of Entrepreneurial Finance* **8(1)**: pp. 87–102.

Clark, P. K. (1973): "A subordinated stochastic process model with finite variance for speculative prices." *Econometrica: Journal of the Econometric Society* pp. 135–155.

Cohen, B. H. (2013): "How Have Banks Adjusted to Higher Capital Requirements?" *Bis quarterly review, september 2013*, Bank for International Settlements.

Cohen, B. H. & M. Scatigna (2016): "Banks and Capital Requirements: Channels of Adjustment." *Journal of Banking & Finance* **69**: pp. S56–S69.

Copeland, T. (1976): "A model of asset trading under the assumption of sequential information arrival." *The Journal of Finance* **31**: pp. 1149–1168.

Corado, C. & J. de Melo (1983): "An ex-ante model for estimating the impact on trade flows of a country's accession to a customs union." *Departmental Working Paper DRD67/1983*, Development Research Department, World Bank.

Corbo, V. & C. Osbat (2013): "Trade adjustment in the European Union - A structural estimation approach." *ECB Working Paper 1535/2013*, European Central Bank, Frankfurt a. M.

Costinot, A. & A. Rodriguez-Clare (2014): "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In G. Gopinath, E. Helpman, & K. Rogoff (editors), "Handbook of International Economics," volume 4, chapter 4, pp. 197–261. Elsevier.

Couaillier, C., A. Reghezza, C. R. d'Acri, & A. Scopelliti (2022): "How to Release Capital Requirements During a Pandemic? Evidence from Euro Area Banks." *Ecb working paper no. 2720*, European Central Bank.

Crouch, R. L. (1970): "The volume of transactions and price changes on the New York stock exchange." *Financial Analysts Journal* **26(4)**: pp. 104–109.

Datar, V. T., N. Y. Naik, & R. Radcliffe (1998): "Liquidity and stock returns: An alternative test." *Journal of Financial Markets* **1(2)**: pp. 203–219.

De Jonghe, O., H. Dewachter, & S. Ongena (2016): "Bank Capital (Requirements) and Credit Supply: Evidence from Pillar 2 Decisions." *Nbb working paper no. 303*, National Bank of Belgium.

De Jonghe, O., H. Dewachter, & S. Ongena (2020): "Bank Capital (Requirements) and Credit Supply: Evidence from Pillar 2 Decisions." *Journal of Corporate Finance* **60**: p. 101518.

De Meiros, O. R. & B. F. N. Van Doornik (2008): "The empirical relationship between stock returns, return volatility and trading volume in the Brazilian stock market." *Brazilian Business Review (English Edition)* **5(1)**.

De Nicolò, G. (2015): "Revisiting the Impact of Bank Capital Requirements on Lending and Real Activity." *Available at SSRN 2688110* .

De Santis, G. *et al.* (1997): "Stock returns and volatility in emerging financial markets." *Journal of International Money and Finance* **16(4)**: pp. 561–579.

Deli, Y. D. & I. Hasan (2017): "Real Effects of Bank Capital Regulations: Global Evidence." *Journal of Banking & Finance* **82**: pp. 217–228.

Deli, Y. D., I. Hasan *et al.* (2017): "Real Effects of Bank Capital Regulations: Global Evidence." *Bank of finland research discussion paper no. 23/2017*, Bank of Finland.

Demirguc-Kunt, A., E. Detragiache, & O. Merrouche (2013): "Bank Capital: Lessons from the Financial Crisis." *Journal of Money, Credit and Banking* **45(6)**: pp. 1147–1164.

Devanadhen, K., K. Srnivasan, & M. Deo (2010): "Price changes, trading volume and time-varying conditional volatility: Evidence from Asia Pacific stock market." *International Review of Applied Financial Issues and Economics* **2(2)**: p. 379.

Disyatat, P. (2011): "The Bank Lending Channel Revisited." *Journal of Money, Credit and Banking* **43(4)**: pp. 711–734.

Doucouliagos, C. (2005): "Publication bias in the economic freedom and economic growth literature." *Journal of Economic Surveys* **19(3)**: pp. 367–387.

Doucouliagos, C. (2011): "How large is large? preliminary and relative guidelines for interpreting partial correlations in economics." *Economics series working paper 05*, Deakin University.

Doucouliagos, C. & T. D. Stanley (2013): "Are all economic facts greatly exaggerated? Theory competition and selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

Drehmann, M. & L. Gambacorta (2012): "The Effects of Countercyclical Capital Buffers on Bank Lending." *Applied Economics Letters* **19(7)**: pp. 603–608.

Egger, M., G. D. Smith, & C. Minder (1997a): "Bias in meta-analysis detected by a simple, graphical test." *Journal of Economic Surveys* **316**: pp. 629–634.

Egger, M., G. D. Smith, M. Schneider, & C. Minder (1997b): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **315(7109)**: pp. 629–634.

Ehrenbergerova, D., J. Bajzik, & T. Havranek (2021): "When Does Monetary Policy Sway House Prices? A Meta-Analysis." *Working paper*, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default priors and predictive performance in Bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Engler, P. & J. Tervala (2018): "Welfare Effects of TTIP in a DSGE Model." *Economic Modelling* **70**: pp. 230–238.

Epps, T. W. & M. L. Epps (1976): "The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis." *Econometrica: Journal of the Econometric Society* pp. 305–321.

Fabo, B., M. Jančoková, E. Kempf, & L. Pástor (2021): "Fifty Shades of QE: Comparing Findings of Central Bankers and Academics." *Journal of Monetary Economics* .

Fabo, B., M. Jančoková, E. Kempf, & L. Pástor (2024): "Fifty shades of QE: Robust evidence." *Journal of Banking & Finance* **159**: p. 107065.

Fama, E. F. (1965): "The behavior of stock-market prices." *The Journal of Business* **38(1)**: pp. 34–105.

Fama, E. F. & K. R. French (1992): "The cross-section of expected stock returns." *The Journal of Finance* **47(2)**: pp. 427–465.

Fama, E. F. & K. R. French (1993): "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* **33**: pp. 3–56.

Fama, E. F. & K. R. French (1996): "Multifactor explanations of asset pricing anomalies." *The Journal of Finance* **51(1)**: pp. 55–84.

Faria, W. R. & E. A. Haddad (2014): "Estimacao das Elasticidades de Substituicao do comercio regional do Brasil [Estimations of Elasticities of regional trade in Brazil]." *Nova Economia* **24(1)**: pp. 141–168.

Feenstra, R. C., P. Luck, M. Obstfeld, & K. N. Russ (2018): "In search of the Armington elasticity." *The Review of Economics and Statistics* **100(1)**: pp. 135–150.

Feldkircher, M. & S. Zeugner (2012): "The impact of data revisions on the robustness of growth determinantsâ€"a note on 'determinants of economic growth: Will data tell?" *Journal of Applied Econometrics* **27(4)**: pp. 686–694.

Fernandez, C., E. Ley, & M. F. J. Steel (2001): "Benchmark priors for Bayesian Model Averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

Fidrmuc, J. & I. Korhonen (2006): "Meta-Analysis of the Business Cycle Correlation Between the Euro Area and the CEECs." *Journal of Comparative Economics* **34(3)**: pp. 518–537.

Fidrmuc, J. & R. Lind (2020): "Macroeconomic Impact of Basel III: Evidence from a Meta-Analysis." *Journal of Banking & Finance* **112**: p. 105359.

Fisher, R. (1954): *Statistical Methods for Research Workers, 12th edn.* Edinburgh, UK: Oliver and Boyd.

Fitzpatrick, T. & K. McQuinn (2007): "House Prices and Mortgage Credit: Empirical Evidence for Ireland." *The Manchester School* **75(1)**: pp. 82–103.

Fletcher, D. (2018): "Bayesian model averaging." In "Model Averaging," pp. 31–55. Springer.

Francois, J. F. & S. Kaplan (1996): "Aggregate Demand Shifts, Income Distribution, and the Linder Hypothesis." *The Review of Economics and Statistics* **78(2)**: pp. 244–250.

French, K. R. (1980): "Stock returns and the weekend effect." *Journal of Financial Economics* **8(1)**: pp. 55–69.

Furukawa, C. (2019): "Publication bias under aggregation frictions: Theory, evidence, and a new correction method." *Working paper*, MIT.

Galac, T. *et al.* (2010): "The Central Bank as Crisis-Manager in Croatia – A Counterfactual Analysis." *Technical report*, Croatian National Bank.

Gallaway, M. P., C. A. McDaniel, & S. A. Rivera (2003): "Short-run and long-run industry-level estimates of US Armington elasticities." *The North American Journal of Economics and Finance* **14(1)**: pp. 49–68.

Gambacorta, L. & D. Marques-Ibanez (2011a): "The Bank Lending Channel: Lessons from the Crisis." *Economic Policy* **26(66)**: pp. 135–182.

Gambacorta, L. & D. Marques-Ibanez (2011b): "The Bank Lending Channel: Lessons from the Crisis." *Bis working paper no. 345*, Bank for International Settlements.

Gambacorta, L. & A. Murcia (2020): "The Impact of Macroprudential Policies in Latin America: An Empirical Analysis Using Credit Registry Data." *Journal of Financial Intermediation* **42**: p. 100828.

Gambacorta, L. & H. S. Shin (2016): "Why Bank Capital Matters for Monetary Policy." *Bis working paper no. 558*, Bank for International Settlements.

Gambacorta, L. & H. S. Shin (2018): "Why Bank Capital Matters for Monetary Policy." *Journal of Financial Intermediation* **35**: pp. 17–29.

Gan, J. (2006): "Substitutability between US domestic and imported forest products: The Armington approach." *Forest Science* **52(1)**: pp. 282–306.

Gechert, S., T. Havranek, Z. Irsova, & D. Kolcunova (2020): "Measuring capital-labor substitution: The importance of method choices and publication bias." *CNB WP* pp. 1–49.

Gechert, S., T. Havranek, Z. Irsova, & D. Kolcunova (2022): "Measuring Capital-Labor Substitution: The Importance of Method Choices and Publication Bias." *Review of Economic Dynamics* **45**: pp. 55–82.

George, E. I. (2010): "Dilution priors: Compensating for model space redundancy." In "IMS Collections Borrowing Strength: Theory Powering Applications - A Festschrift for Lawrence D. Brown," volume 6, pp. 158–165. Institute of Mathematical Statistics.

Gerber, A., N. Malhotra *et al.* (2008): "Do statistical reporting standards affect what is published? Publication bias in two leading political science journals." *Quarterly Journal of Political Science* **3(3)**: pp. 313–326.

Gerber, A. S. & N. Malhotra (2008): "Publication bias in empirical sociological research: Do arbitrary significance levels distort published results?" *Sociological Methods & Research* **37(1)**: pp. 3–30.

Gervais, S., R. Kaniel, & D. H. Mingelgrin (2001): "The high volume return premium." *The Journal of Finance* **56**: pp. 877–919.

Gibbons, M. R. & P. Hess (1981): "Day of the week effects and asset returns." *Journal of Business* pp. 579–596.

Gibson, K. L. (2003): "Armington elasticities for South Africa: Long- and short-run industry level estimates." *TIPS Working Paper Series 12/2003*, Trade and Industrial Policy Strategies, Johannesburg.

Godfrey, M. D., C. W. Granger, & O. Morgenstern (1964): "The random-walk hypothesis of stock market behavior." *Kyklos* **17(1)**: pp. 1–30.

Goldstein, M. & M. S. Khan (1985): "Income and price effects in foreign trade." In R. W. Jones & P. B. Kenen (editors), "Handbook of International Economics," volume 2, chapter 20, pp. 1041–1105. Amsterdam: North-Holland.

Granger, C. W. (1969): "Investigating causal relations by econometric models and cross-spectral methods." *Econometrica: Journal of the Econometric Society* pp. 424–438.

Granger, C. W. & O. Morgenstern (1963): "Spectral analysis of New York stock market prices 1." *Kyklos* **16(1)**: pp. 1–27.

Greene, W. H. (2008): *Econometric Analysis*. Upper Saddle River, New Jersey: Pearson International in Prentice-Hall International Editions.

Guidara, A., I. Soumaré, F. T. Tchana *et al.* (2013): "Banks' Capital Buffer, Risk and Performance in the Canadian Banking System: Impact of Business Cycles and Regulatory Changes." *Journal of Banking & Finance* **37(9)**: pp. 3373–3387.

Gurgul, H., P. Majdosz, & R. Mestel (2007): "Price volume relations of dax companies." *Financial Markets and Portfolio Management* **21**: pp. 353–379.

Hafner, C. M. (2005): "Durations, volume and the prediction of financial returns in transaction time." *Quantitative Finance* **5(2)**: pp. 145–152.

Han, Y., D. Huang, D. Huang, & G. Zhou (2018): "Volume and return: The role of mispricing." *SSRN*, University of North Carolina.

HANCOCK, D. & J. A. WILCOX (1993): "Has There Been A "Capital Crunch" in Banking? The Effects on Bank Lending of Real Estate Market Conditions and Bank Capital Shortfalls." *Journal of Housing Economics* **3(1)**: pp. 31–50.

HANCOCK, D. & J. A. WILCOX (1994): "Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations." *Real Estate Economics* **22(1)**: pp. 59–94.

HANSEN, B. (2007): "Least Squares Model Averaging." *Econometrica* **75(4)**: pp. 1175–1189.

HARVEY, C. R. (2017): "Presidential address: The scientific outlook in financial economics." *The Journal of Finance* **72(4)**: pp. 1399–1440.

HARVEY, C. R., Y. LIU, & H. ZHU (2016): "... and the cross-section of expected returns." *The Review of Financial Studies* **29(1)**: pp. 5–68.

HASAN, I., R. HORVATH, & J. MARES (2018): "What Type of Finance Matters for Growth? Bayesian Model Averaging Evidence." *World Bank Economic Review* **32(2)**: pp. 383–409.

HAVRANEK, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

HAVRANEK, T., R. HORVATH, Z. IRSOVA, & M. RUSNAK (2015): "Cross-country heterogeneity in intertemporal substitution." *Journal of International Economics* **96(1)**: pp. 100–118.

HAVRANEK, T. & Z. IRSOVA (2017): "Do borders really slash trade? A meta-analysis." *IMF Economic Review* **65(2)**: pp. 365–396.

HAVRANEK, T., Z. IRSOVA, & K. JANDA (2012): "Demand for gasoline is more price-inelastic than commonly thought." *Energy Economics* **34(1)**: pp. 201–207.

HAVRANEK, T., Z. IRSOVA, L. LASLOPOVA, & O. ZEYNALOVA (2021): "Skilled and Unskilled Labor Are Less Substitutable than Commonly Thought." *CEPR Discussion Papers 15724*, CEPR.

HAVRANEK, T. & M. RUSNAK (2013): "Transmission Lags of Monetary Policy: A Meta-Analysis." *International Journal of Central Banking* **9(4)**: pp. 39–76.

HAVRANEK, T., M. RUSNAK, & A. SOKOLOVA (2017): "Habit formation in consumption: A meta-analysis." *European Economic Review, Elsevier* **95(C)**: pp. 142–167.

HAVRANEK, T. & A. SOKOLOVA (2020): "Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say 'Probably Not'." *Review of Economic Dynamics* **35**: pp. 97–122.

HEATHCOTE, J. & F. PERRI (2002): "Financial Autarky and International Business Cycles." *Journal of Monetary Economics* **49(3)**: pp. 601–627.

HEDGES, L. V. (1992): "Modeling publication selection effects in meta-analysis." *Statistical Science* **7(2)**: pp. 246–255.

HERNANDEZ, G. (1998): "Elasticidades de sustitucion de las importaciones para la Economia Colombiana." *Revista de Economia del Rosario* **1(2)**: pp. 79–89.

HERTEL, T. W., E. IANCHOVICHINA, & B. J. MCDONALD (1997): "Multi-Region General Equilibrium Modeling." In J. F. FRANCOIS & K. A. REINERT (editors), "Applied Methods for Trade Policy Analysis: A Handbook," chapter 9, pp. 258–299. Cambridge: Cambridge University Press.

HIEMSTRA, C. & J. D. JONES (1994): "Testing for linear and nonlinear Granger causality in the stock price-volume relation." *The Journal of Finance* **49(5)**: pp. 1639–1664.

HIGGINS, J. P. T. & S. G. THOMPSON (2002): "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* **21(11)**: pp. 1539–58.

HILLBERRY, R. H., M. A. ANDERSON, E. J. BALISTRERI, & A. K. FOX (2005): "Taste parameters as model residuals: Assessing the 'fit' of an Armington trade model." *Review of International Economics* **13(5)**: pp. 973–984.

HILLBERRY, R. H. & D. HUMMELS (2013): "Trade elasticity parameters for a Computable General Equilibrium model." In P. B. DIXON & D. W. JORGENSON (editors), "Handbook of Computable General Equilibrium Modeling," volume 1, chapter 18, pp. 1213–1269. Elsevier.

Houston, J., C. James, & D. Marcus (1997): "Capital Market Frictions and the Role of Internal Capital Markets in Banking." *Journal of Financial Economics* **46(2)**: pp. 135–164.

Hu, S.-Y. (1997): "Trading turnover and expected stock returns: The trading frequency hypothesis and evidence from the Tokyo stock exchange." *Working paper*, University of Chicago.

Huang, X. & Q. Xiong (2015): "Bank Capital Buffer Decisions under Macroeconomic Fluctuations: Evidence for the Banking Industry of China." *International Review of Economics & Finance* **36**: pp. 30–39.

Hubler, M. & F. Pothen (2017): "Trade-induced productivity gains reduce incentives to impose strategic tariffs." *Economic Modelling* **61**: pp. 420–431.

Huchet-Bourdon, M. & E. Pishbahar (2009): "Armington elasticities and tariff regime: An application to European Union rice imports." *Journal of Agricultural Economics* **60(3)**: pp. 586–603.

Hummels, D. (1999): "Toward a Geography of Trade Costs." *GTAP Working Paper 17/1999*, Purdue University, West Lafayette, Indiana.

IBRD (2009): "2009 Information and communications for development: Extending reach and increasing impact." *Technical report*, International Bank for Reconstruction and Development, The World Bank, Washington DC.

Imai, T., T. A. Rutter, & C. F. Camerer (2021): "Meta-analysis of present-bias estimation using convex time budgets." *The Economic Journal* **131(636)**: pp. 1788–1814.

Imbierowicz, B., J. Kragh, & J. Rangvid (2018): "Time-Varying Capital Requirements and Disclosure Rules: Effects on Capitalization and Lending Decisions." *Journal of Money, Credit and Banking* **50(4)**: pp. 573–602.

Imbs, J. M. & I. Mejean (2015): "Elasticity optimism." *American Economic Journal: Macroeconomics* **7(3)**: pp. 43–83.

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, & B. Puranen (2014): "World Values Survey: All rounds - Country-pooled datafile at http://www.worldvaluessurvey.org/ WVSDocumentationWVL.jsp." *Technical report*, Madrid: JD Systems Institute.

IOANNIDIS, J. P. (1998): "Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials." *Jama* **279(4)**: pp. 281–286.

IOANNIDIS, J. P. (2008): "Why most discovered true associations are inflated." *Epidemiology* pp. 640–648.

IOANNIDIS, J. P., T. D. STANLEY, & H. DOUCOULIAGOS (2017a): "The power of bias in economics research." *The Economic Journal* **127(605)**: pp. F236–F265.

IOANNIDIS, J. P., T. D. STANLEY, & H. DOUCOULIAGOS (2017b): "The Power of Bias in Economics Research." *Economic Journal* **127(605)**: pp. F236–F265.

IVANOVA, N. (2005): "Estimation of own- and cross-price elasticities of disaggregated imported and domestic goods in Russia." *EERC Working Paper Series 13/2005*, Economics Education and Research Consortium Research Network, Moscow, Russia.

JAIN, P. C. & G.-H. JOH (1988): "The dependence between hourly prices and trading volume." *The Journal of Financial and Quantitative Analysis* **23**: pp. 269–283.

JEFFREYS, H. (1961): *Theory of Probability. Classic Texts in the Physical Sciences.* Oxford: Oxford University Press, third edition.

JEGADEESH, N. & S. TITMAN (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of Finance* **48(1)**: pp. 65–91.

JEGADEESH, N. & S. TITMAN (1995): "Overreaction, delayed reaction, and contrarian profits." *The Review of Financial Studies* **8(4)**: pp. 973–993.

JENNINGS, R. H., L. T. STARKS, & F. J. C. (1981): "An equilibrium model of asset trading with sequential information arrival." *The Journal of Finance* **36(1)**: pp. 143–161.

JOKIPII, T. & A. MILNE (2011): "Bank Capital Buffer and Risk Adjustment Decisions." *Journal of Financial Stability* **7(3)**: pp. 165–178.

JORDÀ, Ò., B. RICHTER, M. SCHULARICK, & A. M. TAYLOR (2021): "Bank Capital Redux: Solvency, Liquidity, and Crisis." *The Review of Economic Studies* **88(1)**: pp. 260–286.

JORDÀ, Ò., M. SCHULARICK, & A. M. TAYLOR (2011): "Financial Crises, Credit Booms, and External Imbalances: 140 Years of Lessons." *IMF Economic Review* **59(2)**: pp. 340–378.

JOYCE, M. & M. SPALTRO (2014): "Quantitative easing and bank lending: A panel data approach." *Boe working paper no.504*, Bank of England.

KANIEL, R., A. OZOGUZ, & L. STARKS (2012): "The high volume return premium: Cross-country evidence." *Journal of Financial Economics* **103(2)**: pp. 255–279.

KANIEL, R., G. SAAR, & S. TITMAN (2008): "Individual investor trading and stock returns." *The Journal of Finance* **63(1)**: pp. 273–310.

KANNGIESSER, D., R. MARTIN, L. MAURIN, & D. MOCCERO (2017): "Estimating the Impact of Shocks to Bank Capital in the Euro Area." *Ecb working paper no.2077*, European Central Bank.

KAPUSCINSKI, C. A. & P. G. WARR (1999): "Estimation of Armington Elasticities: An Application to the Philippines." *Departmental Working Papers 1996-08*, Australian National University, Economics RSPAS.

KARPOFF, J. M. (1987): "The relation between price changes and trading volume: A survey." *Journal of Financial and Quantitative Analysis* **22(1)**: pp. 109–126.

KASHYAP, A. K., J. C. STEIN, & S. HANSON (2010): "An Analysis of the Impact of 'Substantially Heightened' Capital Requirements on Large Financial Institutions." *Booth School of Business, University of Chicago, mimeo* **2**: pp. 1–47.

KAWASHIMA, S. & D. A. P. SARI (2010): "Time-varying Armington elasticity and country-of-origin bias: from the dynamic perspective of the Japanese demand for beef imports." *Australian Journal of Agricultural and Resource Economics* **54(1)**: pp. 27–41.

Kehoe, T. J., P. S. Pujolas, & J. Rossbach (2017): "Quantitative trade models: Developments and challenges." *Staff Report 537*, Federal Reserve Bank of Minneapolis.

Kim, D. & W. Sohn (2017): "The Effect of Bank Capital on Lending: Does Liquidity Matter?" *Journal of Banking & Finance* **77**: pp. 95–107.

Kim, J., H. Doucouliagos, & T. Stanley (2019): "Market efficiency in Asian and Australasian stock markets: A fresh look at the evidence." In "International financial markets," pp. 382–419. Routledge.

Kim, S. H., K. A. Kim, & J. R. Nofsinger (2003): "The behavior and performance of individual investors in Japan." *Pacific Basin Finance* pp. 1612–1656.

Kishan, R. P. & T. P. Opiela (2000): "Bank Size, Bank Capital, and the Bank Lending Channel." *Journal of Money, Credit and Banking* pp. 121–141.

Klomp, J. & J. De Haan (2009): "Central Bank Independence and Financial Instability." *Journal of Financial Stability* **5(4)**: pp. 321–338.

Kočenda, E. & I. Iwasaki (2022): "Bank Survival Around the World: A Meta-Analytic Review." *Journal of Economic Surveys* **36(1)**: pp. 108–156.

Kolcunová, D. & S. Malovaná (2019): "The Effect of Higher Capital Requirements on Bank Lending: The Capital Surplus Matters." *Cnb working paper no. 2/19*, Czech National Bank.

Košak, M., S. Li, I. Lončarski, & M. Marinč (2014): "Quality of Bank Capital and Bank Lending Behavior during the Global Financial Crisis." *Working paper*, University of Ljubljana.

Košak, M., S. Li, I. Lončarski, & M. Marinč (2015): "Quality of Bank Capital and Bank Lending Behavior During the Global Financial Crisis." *International Review of Financial Analysis* **37**: pp. 168–183.

Kragh, J. & J. Rangvid (2016): "Time-Varying Capital Requirements and Disclosure Rules: Effects on Capitalization and Lending Decisions of Banks." *Working paper no. 41*, FRIC, Copenhagen Business School.

KUIPER, M. & F. VAN TONGEREN (2006): "Using gravity to move Armington: An empirical approach to the small initial trade share problem in general equilibrium models." *Technical report*, Global Trade Analysis Project (GTAP). Paper prepared for the Ninth Annual Conference on Global Economic Analysis, June 15-17, 2006 in Addis Ababa, Ethiopia.

LABONNE, C. & G. LAMÉ (2014): "Credit growth and bank capital requirements: Binding or not?" *Banque de france working paper series no. 481*, Banque de France.

LACHLER, U. (1984): "The elasticity of substitution between imported and domestically produced goods in Germany." *Kiel Working Papers 200/1984*, Kiel Institute of World Economics, Kiel.

LAMBERTINI, L. & A. MUKHERJEE (2016): "Is Bank Capital Regulation Costly for Firms? Evidence from Syndicated Loans." *Working paper*, The École Polytechnique Fédérale de Lausanne.

LE, Q. T. & M. MEHMED (2009): "The relationship between trading volume, stock index returns and volatility: Empirical evidence in Nordic countries." *Master thesis in finance*, Lund University.

LE LESLÉ, V. & S. Y. AVRAMOVA (2012): "Revisiting Risk-Weighted Assets." *Imf working paper no. 90*, International Monetary Fund.

LEE, C. & O. RUI (2000): "Does trading volume contain information to predict stock returns? Evidence from China's stock markets." *Review of Quantitative Finance and Accounting* **14**: pp. 341–360.

LEE, C. & O. RUI (2002): "The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence." *Journal of Banking and Finance* **26**: pp. 51–78.

LEWELLEN, J. (2015): "The cross section of expected stock returns." *Critical Financial Review* **4(1)**: pp. 1–44.

LEY, E. & M. F. STEEL (2009): "On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression." *Applied Econometrics* **24**: pp. 651–674.

LIN, T.-C. & X. LIU (2017): "Skewness, individual investor preference, and the cross-section of stock returns." *Review of Finance* pp. 1–36.

LINDER, S. B. (1961): *An essay on trade and transformation.* Almqvist & Wiksells boktryckeri: Uppsala.

LO, A. W. & J. WANG (2000): "Trading volume: Definitions, data analysis, and implications of portfolio theory." *The Review of Financial Studies* **13**: pp. 257–300.

LONG, H., Y. JINAG, & Y. ZHU (2018): "Idiosyncratic tail risk and expected stock returns: Evidence from the Chinese stock markets." *Finance Research Letters* **24**: pp. 129–136.

LOPEZ, E. & E. PAGOULATOS (2002): "Estimates and Determinants of Armington Elasticities for the U.S. Food Industry." *Journal of Industry, Competition and Trade* **2(3)**: pp. 247–258.

LOUHICHI, W. (2012): "Does trading activity contain information to predict stock returns? Evidence from Euronext Paris." *Applied Financial Economics* **22(8)**: pp. 625–632.

LOUKIL, N., M. B. ZAYANI, & A. OMRI (2010): "Impact of liquidity on stock returns: an empirical investigation of the Tunisian stock market." *Macroeconomics and Finance in Emerging Market Economies* **3(2)**: pp. 261–283.

LOZANO KARANAUSKAS, C. (2004): "Elasticidades de sustitucion Armington para Colombia." *Archivos de Economia 271/2004*, Direccion de Estudios Economicos, Departamento Nacional de Planeacion, Republica de Colombia.

LUNDMARK, R. & S. SHAHRAMMEHR (2011a): "Forest biomass and Armington elasticities in Europe." *Biomass and Bioenergy* **35(1)**: pp. 415–420.

LUNDMARK, R. & S. SHAHRAMMEHR (2011b): "Sweden's import substitution possibilities for roundwood." *Scandinavian Journal of Forest Research* **26(2)**: pp. 146–153.

LUNDMARK, R. & S. SHAHRAMMEHR (2012): "Armington elasticities and induced industrial roundwood specialization in Europe." *Journal of Natural Resources Policy Research* **4(3)**: pp. 161–170.

MAGNUS, J. R., O. POWELL, & P. PRUFER (2010): "A comparison of two model averaging techniques with an application to growth empirics." *Journal of econometrics* **154(2)**: pp. 139–153.

Mahajan, S. & B. Singh (2008): "An empirical analysis of stock price-volume relationship in Indian stock market." *Vision* **12(3)**: pp. 1–13.

Mahajan, S. & B. Singh (2009a): "The empirical investigation of relationship between return, volume and volatility dynamics in Indian stock market." *Eurasian Journal of Business and Economics* **2(4)**: pp. 113–137.

Mahajan, S. & B. Singh (2009b): "Return-volume dynamics in Indian stock market." *Asia-Pacific Business Review* **5(3)**: pp. 63–70.

Malkiel, B. G. (2003): "The efficient market hypothesis and its critics." *Journal of Economic Perspectives* **17(1)**: pp. 59–82.

Malovaná, S., J. Bajzík, D. Ehrenbergerová, & J. Janků (2022): "A Prolonged Period of Low Interest Rates in Europe: Unintended Consequences." *Journal of Economic Surveys* **37(2)**.

Malovaná, S. & J. Frait (2017): "Monetary Policy and Macroprudential Policy: Rivals or Teammates?" *Journal of Financial Stability* **32**: pp. 1–16.

Marquez, J. (2002): *Estimating Trade Elasticities*. Boston: Kluwer.

Marshall, B. R. & M. Young (2003): "Liquidity and stock returns in pure order-driven markets: evidence from the Australian stock market." *International Review of Financial Analysis* **12(2)**: pp. 173–188.

Mathur, A., M. Naylor, & A. Rajan (2023): "Useful, Usable, and Used? Buffer Usability During the Covid-19 Crisis." *Boe staff working paper no. 1011*, Bank of England.

McCallum, J. (1995): "National borders matter: Canada-U.S. regional trade patterns." *American Economic Review* **85**: pp. 615–623.

McCloskey, D. N. & S. T. Ziliak (2019): "What Quantitative Methods Should We Teach to Graduate Students? A Comment on Swann's Is Precise Econometrics an Illusion?" *The Journal of Economic Eduation* **50(4)**: pp. 356–361.

McDaniel, C. A. & E. J. Balistreri (2003): "A discussion on Armington trade substitution elasticities." *Economie Internationale* **94-95(2)**: pp. 301–313.

McGowan, C. B. & J. Muhammad (2012): "The relationship between price and volume for the Russian trading system." *The International Business and Economics Research Journal (Online)* **11(9)**: pp. 963–970.

McMillan, D. & A. Speight (2002): "Return-volume dynamics in UK futures." *Applied Financial Economics* **12**: pp. 707–713.

Meeks, R. (2017): "Capital Regulation and the Macroeconomy: Empirical Evidence and Macroprudential Policy." *European Economic Review* **95**: pp. 125–141.

Memmel, C. & P. Raupach (2010): "How Do Banks Adjust Their Capital Ratios?" *Journal of Financial Intermediation* **19(4)**: pp. 509–528.

Mendicino, C., K. Nikolov, J. Suarez, & D. Supera (2018): "Optimal Dynamic Capital Requirements." *Journal of Money, Credit and Banking* **50(6)**: pp. 1271–1297.

Mésonnier, J.-S. & D. Stevanovic (2012): "Bank Leverage Shocks and the Macroeconomy in a Data-Rich Environment." *Banque de france working paper series no. 394*, Banque de France.

Mésonnier, J.-S. & D. Stevanovic (2017): "The Macroeconomic Effects of Shocks to Large Banks' Capital." *Oxford Bulletin of Economics and Statistics* **79(4)**: pp. 546–569.

Mestel, R., H. Gurgul, & P. Majdosz (2003): "The empirical relationship between stock returns, return volatility and trading volume on the Austrian stock market." *Research paper*, University of Graz, Institute of Banking and Finance.

Miles, D., J. Yang, & G. Marcheggiano (2013): "Optimal Bank Capital." *The Economic Journal* **123(567)**: pp. 1–37.

Mohler, L. & M. Seitz (2012): "The gains from variety in the European Union." *Review of World Economics* **148(3)**: pp. 475–500.

Mora, N. & A. Logan (2010): "Shocks to Bank Capital: Evidence from UK Banks at Home and Away." *Bank of england working paper no. 387*, Bank of England.

Mora, N. & A. Logan (2012): "Shocks to Bank Capital: Evidence from UK Banks at Home and Away." *Applied Economics* **44(9)**: pp. 1103–1119.

Moral-Benito, E. (2015): "Model Averaging in Economics: An Overview." *Journal of Economic Surveys* **29(1)**: pp. 46–75.

Morgan, I. G. (1976): "Stock prices and heteroscedasticity." *The Journal of Business* **49(4)**: pp. 496–508.

Naceur, M. S. B., J. Pépy, & C. Roulet (2017): "Basel III and Bank-Lending: Evidence from the United States and Europe." *Imf working paper no. wp/17/245*, International Monetary Fund.

Naceur, S. B., K. Marton, & C. Roulet (2018): "Basel III and Bank-Lending: Evidence from the United States and Europe." *Journal of Financial Stability* **39**: pp. 1–27.

Narayan, P. K. & X. Zheng (2010): "Market liquidity risk factor and financial market anomalies: Evidence from the Chinese stock market." *Pacific-Basin Finance Journal* **18(5)**: pp. 509–520.

Nemeth, G., L. Szabo, & C. Juan-Carlos (2011): "Estimation of Armington elasticities in a CGE economy-energy-environment model for Europe." *Economic Modelling* **28(4)**: pp. 1993–1999.

Nganou, J.-P. (2005): "Estimates of Armington parameters for a landlocked economy." *MPRA Working Paper 31459/2005*, Munich Personal RePEc Archive, Munich.

Noss, J. & P. Toffano (2014): "Estimating the Impact of Changes in Aggregate Bank Capital Requirements During an Upswing." *Boe working paper no. 494*, Bank of England.

Noss, J. & P. Toffano (2016): "Estimating the Impact of Changes in Aggregate Bank Capital Requirements on Lending and Growth During an Upswing." *Journal of Banking & Finance* **62**: pp. 15–27.

Ochere, O. G., M. T. Nasieku, & T. O. Olweny (2018): "Trading volume and fama-french three factor model on excess return. Empirical evidence from nairobi security exchange." *European Scientific Journal* **14(22)**: pp. 276–289.

Ogundeji, A. A., J. andre, & D. Uchezuba (2010): "Econometric estimation of Armington elasticities for selected agricultural products in South Africa." *South African Journal of Economic and Management Sciences* **13(2)**: pp. 123–134.

Olekseyuk, Z. & H. Schurenberg-Frosch (2016): "Are Armington elasticities different across countries and sectors? A European study." *Economic Modelling* **55**: pp. 328–342.

Olszak, M., M. Pipien, S. Roszkowska, & I. Kowalska (2014): "The Effects of Capital on Bank Lending in Large EU Banks – The Role of Procyclicality, Income Smoothing, Regulations and Supervision." *Working paper no. 5/2014*, University of Warsaw Faculty of Management.

Parsley, D. C. & S.-J. Wei (2001): "Explaining the border effect: the role of exchange rate variability, shipping costs, and geography." *Journal of International Economics* **55(1)**: pp. 87–105.

Peek, J. & E. S. Rosengren (1997): "The International Transmission of Financial Shocks: The Case of Japan." *American Economic Review* pp. 495–505.

Pisedtasalasai, A. & A. Gunasekarage (2007): "Causal and dynamic relationships among stock returns, return volatility and trading volume: Evidence from emerging markets in South-East Asia." *Asia-Pacific Financial Markets* **14(4)**: pp. 277–297.

Poudel, R. B. & S. R. Shrestha (2019): "Stock return and trading volume relation in Nepalese stock market: An ardl approach." *SEBON JOURNAL* .

Raftery, A. E. (1995): "Bayesian model selection in social research." *Sociological methodology* pp. 111–163.

Raftery, A. E., D. Madigan, & J. Hoeting (1997): "Bayesian model averaging for linear regression models." *Journal of the American Statistical Association* **92**: pp. 179–191.

Reinert, K. A. & D. W. Roland-Holst (1992): "Armington elasticities for United States manufacturing sectors." *Journal of Policy Modelling* **14(5)**: pp. 631–639.

Reinert, K. A. & C. R. Shiells (1991): "Estimated elasticities of substitution for analysis of a North American Free Trade Area." *Technical report*, Presented at the Conference on Modeling North American Economic Integration, March 1-2, 1991, Federal Reserve Bank of Minneapolis.

Richter, B., M. Schularick, & I. Shim (2019): "The Costs of Macroprudential Policy." *Journal of International Economics* **118**: pp. 263–282.

Rogalski, R. J. (1978): "The dependence of prices and volume." *The Review of Economics and Statistics* pp. 268–274.

Roodman, D. (2020): "Boottest: Stata module to provide fast execution of the wild bootstrap with null imposed." .

Rossi, P. E., G. M. Allenby, & R. McCulloch (2005): "Markov Chain Monte Carlo Methods." In V. Barnett, J. S. Hunter, & D. G. Kendall (editors), "Bayesian Statistics and Marketing," volume 1 of *Wiley Series in Probability and Statistics*, chapter 3, pp. 49–102. John Wiley & Sons, Ltd.

Rotila, D.-M., M. Onofrei, & A. M. Andries (2015): "The relation between stock returns, trading volume and return volatility of the CEE banks." *Transformations in Business and Economics* **14**.

Roulet, C. (2018): "Basel III: Effects of Capital and Liquidity Regulations on European Bank Lending." *Journal of Economics and Business* **95**: pp. 26–46.

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

Saatcioglu, K. & L. T. Starks (1998): "The stock price-volume relationship in emerging stock markets: the case of Latin America." *International Journal of Forecasting* **14(2)**: pp. 215–225.

Saikkonen, L. (2015): "Estimation of substitution and transformation elasticities for South African trade." *UNU-WIDER Working Paper 104/2015*, United Nations University, World Institute for Development, Helsinki, Finland.

SAITO, M. (2004): "Armington Elasticities in Intermediate Inputs Trade: A Problem in Using Multilateral Trade Data." *The Canadian Journal of Economics* **37(4)**: pp. 1097–1107.

SANA HSIEH, H.-C. (2014): "The causal relationships between stock returns, trading volume, and volatility: Empirical evidence from Asian listed real estate companies." *International Journal of Managerial Finance* **10(2)**: pp. 218–240.

SAUQUET, A., F. LECOCQ, P. DELACOTE, S. CAURLA, A. BARKAOUI, & S. GARCIA (2011): "Estimating Armington elasticities for sawnwood and application to the French Forest Sector Model." *Resource and Energy Economics* **33(4)**: pp. 771–781.

SCHULARICK, M. & A. TAYLOR (2012): "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008." *American Economic Review* **102(2)**: pp. 1029–1061.

SCHURENBERG-FROSCH, H. (2015): "We Could Not Care Less About Armington Elasticities- But Should We? A Meta-Sensitivity Analysis of the Influence of Armington Elasticity Misspecification on Simulation Results." *Ruhr Economic Papers 594/2015*, Ruhr-Universitat Bochum, Germany.

SCHWERT, G. W. (2003): "Anomalies and market efficiency." *Handbook of the Economics of Finance* **1**: pp. 939–974.

SCHWERT, M. (2018): "Bank Capital and Lending Relationships." *The Journal of Finance* **73(2)**: pp. 787–830.

SHEU, H.-J., S. WU, & K.-P. KU (1998): "Cross-sectional relationships between stock returns and market beta, trading volume, and sales-to-price in Taiwan." *International Review of Financial Analysis* **7(1)**: pp. 1–18.

SHIELLS, C. R. & K. A. REINERT (1993): "Armington models and the terms-of-trade effects: Some econometric evidence for North America." *The Canadian Journal of Economics* **26(2)**: pp. 299–316.

SHIELLS, C. R., R. M. STERN, & A. V. DEARDORFF (1986): "Estimates of the elasticities of substitution between imports and home goods for the United States." *Weltwirtschaftliches Archiv* **122(3)**: pp. 497–519.

SHU, P.-G., Y.-H. YEH, & Y.-C. HUANG (2004): "Stock price and trading volume effects associated with changes in the MSCI free indices: evidence from Taiwanese firms added to and deleted from the indices." *Review of Pacific Basin Financial Markets and Policies* **7(04)**: pp. 471–491.

STANLEY, T. & H. DOUCOULIAGOS (2010): "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economic Surveys* **24(1)**: pp. 170–191.

STANLEY, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

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

STANLEY, T. D. (2013): "Meta-analysis of economics research reporting guidelines." *Journal of Economic Surveys* **27(2)**: pp. 390–394.

STANLEY, T. D. & H. DOUCOULIAGOS (2012): *Meta-regression analysis in economics and business.* New York, USA: Routledge.

STANLEY, T. D. & H. DOUCOULIAGOS (2014): "Meta-Regression Approximations to Reduce Publication Selection Bias." *Research Synthesis Methods* **5(1)**: pp. 60–78.

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

STATMAN, M., S. THORLEY, & K. VORKINK (2006): "Investor overconfidence and trading volume." *The Review of Financial Studies* **19(4)**: pp. 1531–1565.

STEEL, M. F. J. (2020): "Model Averaging and its Use in Economics." *Journal of Economic Literature* **(forthcoming)**.

TAHIR, S. H., F. ALL, N. GHAFFAR, & H. M. SABIR (2016): "Causal relationship among trading volume, returns and stock volatility: evidence from an emerging market." *Technical report*, University Faisalabad.

TAPA, A. & M. HUSSEIN (2016): "The relationship between stock return and trading volume in Malaysian ACE market." *International Journal of Economics and Financial Issues* **6(75)**: pp. 271–278.

TAUCHEN, G. E. & M. PITTS (1983): "The price variability-volume relationship on speculative markets." *Econometrica: Journal of the Econometric Society* pp. 485–505.

THAKOR, A. V. (2014): "Bank Capital and Financial Stability: An Economic Trade-Off or a Faustian Bargain?" *Annual Review of Financial Economics* **6(1)**: pp. 185–223.

TOURINHO, O., A. FONTES, H. KUME, & A. C. DE SOUZA PEDROSO (2003): "Armington Elasticities for Brazil – 1986-2002: New Estimates." *IPEA Discussion Papers 974/2003*, Institute for Applied Economic Research, Rio de Janeiro, Brazil.

TOURINHO, O., A. FONTES, H. KUME, & A. C. DE SOUZA PEDROSO (2010): "Armington Elasticities for Brazil." In C. DE MIGUEL, J. D. LIMA, P. GIORDANO, J. GUZMĂˇN, A. SCHUSCHNY, & M. WATANUKI (editors), "Modeling Public Policies in Latin America and the Caribbean," chapter 9, pp. 285–316. United Nations Publication, Santiago, Chile.

TREFLER, D. (1995): "The case of the missing trade and other mysteries." *American Economic Review* **85(5)**: pp. 1029–46.

TRIPATHY, N. (2011): "The relation between price changes and trading volume: A study in Indian stock market." *Interdisciplinary Journal of Research in Business* **1(7)**: pp. 81–95.

TURNER, K., S. Y. HA, G. HEWINGS, P. MCGREGOR, & K. SWALES (2012): "Econometric estimation of Armington import elasticities for a regional CGE model of the Illinois economy." *Economic Systems Research* **24(1)**: pp. 1–19.

VALICKOVA, P., T. HAVRANEK, & R. HORVATH (2015a): "Financial development and economic growth: A meta-analysis." *Journal of Economic Surveys* **29(3)**: pp. 506–526.

VALICKOVA, P., T. HAVRANEK, & R. HORVATH (2015b): "Financial Development And Economic Growth: A Meta-Analysis." *Journal of Economic Surveys* **29(3)**: pp. 506–526.

VANDENBUSSCHE, J., U. VOGEL, & E. DETRAGIACHE (2015): "Macroprudential Policies and Housing Prices: A New Database and Empirical Evidence for Central, Eastern, and Southeastern Europe." *Journal of Money, Credit and Banking* **47(S1)**: pp. 343–377.

VanHoose, D. (2007): "Theories of Bank Behavior Under Capital Regulation." *Journal of Banking & Finance* **31(12)**: pp. 3680–3697.

Wang, M. B. & T. Sun (2013): "How Effective are Macroprudential Policies in China?" *Imf working paper no. wp/13/75*, International Monetary Fund.

Warr, P. G. & E. A. Lapiz (1994): "Estimation of Elasticities of Substitution/Transformation between Domestic and Imported/Exported Commodities." *Technical report*, mimeo.

Watanabe, W. (2006): "Does a Large Loss of Bank Capital Cause Evergreening? Evidence from Japan." *Forthcoming in the Journal of the Japanese and International Economies* .

Watanabe, W. (2010): "Does a Large Loss of Bank Capital Cause Evergreening? Evidence from Japan." *Journal of the Japanese and International Economies* **24(1)**: pp. 116–136.

Watkins, B. D. (2007): "The economic and predictive value of trading volume growth: A tale of three moments." *Applied Financial Economics* **17**: pp. 1489–1509.

WB (2018): "World Development Indicators (WDI): Public database." *Database*, The World Bank: Washington, DC. Online at https://data.worldbank.org/ [Accessed Sept-5, 2018].

Weale, M. & T. Wieladek (2022): *Fifty Shades of QE Revisited.* Centre for Economic Policy Research.

Welsch, H. (2006): "Armington elasticities and induced intra-industry specialization: The case of France, 1970-1997." *Economic Modelling* **23(3)**: pp. 556–567.

Welsch, H. (2008): "Armington elasticities for energy policy modeling: Evidence from four European countries." *Energy Economics* **30(5)**: pp. 2252–2264.

Wolf, H. (2000): "Intranational home bias in trade." *Review of Economics and Statistics* **82**: pp. 555–563.

Wooldridge, J. M. (2015): *Introductory econometrics: A modern approach.* Scarborough, Canada: Nelson Education.

Xue, X., W. R. Reed, R. C. van Aert *et al.* (2022): "Social Capital and Economic Growth: A Meta-Analysis." *Working Paper 20/22*, University of Canterbury, Department of Economics and Finance.

Yin, Y. & Y. Liu (2018): "Information of unusual trading volume." *Emerging Markets Finance and Trade* **54(11)**: pp. 2409–2432.

Ying, C. C. (1966): "Market prices and volumes of sales." *Econometrica* **34**: pp. 676–685.

Yonis, M. (2014): "Trading volume and stock return: Empirical evidence for Asian tiger economies." *Technical report*, UMEA Universitet.

Zeugner, S. & M. Feldkircher (2015): "Bayesian model averaging employing fixed and flexible priors: The BMS package for R." *Journal of Statistical Software* **68(4)**: pp. 1–37.

Zhang, X.-g. & G. Verikios (2006): "Armington Parameter Estimation for a Computable General Equilibrium Model: A Database Consistent Approach." *Economics Discussion/Working Papers 10/2006*, The University of Western Australia, Department of Economics.

Zhong, A., D. Chai, B. Li, & M. Chiah (2018): "Volume shocks and stock returns: An alternative test." *Pacific-Basin Finance Journal* **48**: pp. 1–16.

Zigraiova, D., T. Havranek, Z. Irsova, & J. Novak (2021): "How Puzzling is the Forward Premium Puzzle? A Meta-Analysis." *European Economic Review* **134**: p. 103714.

Appendix A

Appendix to the Chapter 2

This appendix provides robustness checks and diagnostics of the BMA exercise included in the main body of the paper.

Table A.1: Description and summary statistics of the regression variables

Variable	Description	Mean	SD	WM
Armington elasticity	The reported long-run estimate of the Armington elasticity.	1.56	1.91	1.67
Standard error (SE)	The reported standard error of the long-run Armington elasticity estimate.	0.82	1.26	0.65
Data characteristics				
Data aggregation	The level of data aggregation according to SIC classification (min = 1 if fully aggregated, max = 8 if disaggregated).	6.46	1.56	6.01
Results aggregation	The level of results aggregation according to SIC classification (min = 1 if fully aggregated, max = 8 if disaggregated).	5.00	1.21	5.17
Monthly data	=1 if the data are in monthly frequency.	0.06	0.24	0.08
Quarterly data	=1 if the data are in quarterly frequency (reference category for the group of dummy variables describing data frequency).	0.24	0.43	0.30
Annual data	=1 if the data are in yearly frequency.	0.70	0.46	0.62
Panel data	=1 if panel data are used (reference category for the group of dummy variables describing time and cross-sectional dimension of data).	0.39	0.49	0.28
Time series	=1 if time-series data are used.	0.51	0.50	0.62
Cross-section	=1 if cross-sectional data are used.	0.09	0.29	0.10
Data period	The length of time period in years.	13.38	7.85	14.76
Data size	The logarithm of the total number of observations used to estimate the elasticity.	4.72	2.06	4.40

Continued on next page

Table A.1: Description and summary statistics of the regression variables (continued)

Variable	Description	Mean	SD	WM
Midyear	The median year of the time period of the data used to estimate the elasticity.	23.42	12.43	22.27
Spatial Variation				
Primary sector	=1 if the estimate is for the primary sector (agriculture and raw materials; reference category for the group of dummy variables describing sectors).	0.11	0.32	0.23
Secondary sector	=1 if the estimate is for the secondary sector (manufacturing).	0.86	0.35	0.70
Tertiary sector	=1 if the estimate is for the tertiary sector (services).	0.01	0.11	0.02
Developing countries	=1 if the estimate is for a developing country (reference category for the group of dummy variables describing the level of development).	0.22	0.42	0.31
Developed countries	=1 if the estimate is for developed country.	0.82	0.39	0.72
Market size	The logarithm of the market size of the home country (GDP in billions of USD, 2015 prices).	6.28	1.81	6.21
Tariffs	The tariff rate of the home country (weighted mean, all products, %).	6.87	7.63	6.64
Non-tariff barriers	Additional cost to import of the home country (USD per container).	0.94	0.25	0.93
FX volatility	The volatility of the exchange rate using the DEC alternative conversion factor (home country currency unit per USD).	0.64	0.56	0.66
National pride	Home bias captured by the percentage of "I am very proud of my country" answers from the World Values Survey.	0.49	0.22	0.53
Internet usage	The number of fixed broadband subscriptions of the home country (per 100 people).	3.44	5.31	1.25
Estimation technique				
Static model	=1 if a static model is used for estimation.	0.26	0.44	0.35
Distributed lag and trend model	=1 if a distributed lag and trend model of is used.	0.12	0.32	0.16
Partial adjustment model	=1 if a partial adjustment model is used for estimation.	0.14	0.35	0.14
Error-correction model	=1 if an error-correction model is used.	0.03	0.18	0.08
Non-linear model	=1 if a non-linear model is used.	0.33	0.47	0.13
Other models	=1 if another model is used (reference category for the group of dummy variables describing models used).	0.11	0.32	0.13
OLS	=1 if the OLS or GLS estimation method is used.	0.41	0.49	0.61

Continued on next page

Table A.1: Description and summary statistics of the regression variables (continued)

Variable	Description	Mean	SD	WM
CORC	=1 if the Cochrane-Orcutt or FGLS estimation method is used.	0.18	0.38	0.17
TSLS	=1 if two-stage least squares are used.	0.08	0.28	0.07
GMM	=1 if the GMM estimation method is used.	0.29	0.45	0.08
Other methods	=1 if other types of estimation are used (reference category for the group of dummy variables describing the estimation method used).	0.04	0.19	0.07
Import constraint	=1 if the study controls for import restrictions.	0.04	0.19	0.10
Seasonality	=1 if the study controls for seasonality.	0.13	0.34	0.17
Publication characteristics				
Impact factor	The recursive discounted impact factor from RePEc.	0.13	0.26	0.22
Citations	The logarithm of the number of Google Scholar citations normalized by the number of years since the first draft of the paper appeared in Google Scholar.	1.17	0.97	1.14
Published	=1 if a study is published in a peer-reviewed journal.	0.34	0.47	0.56

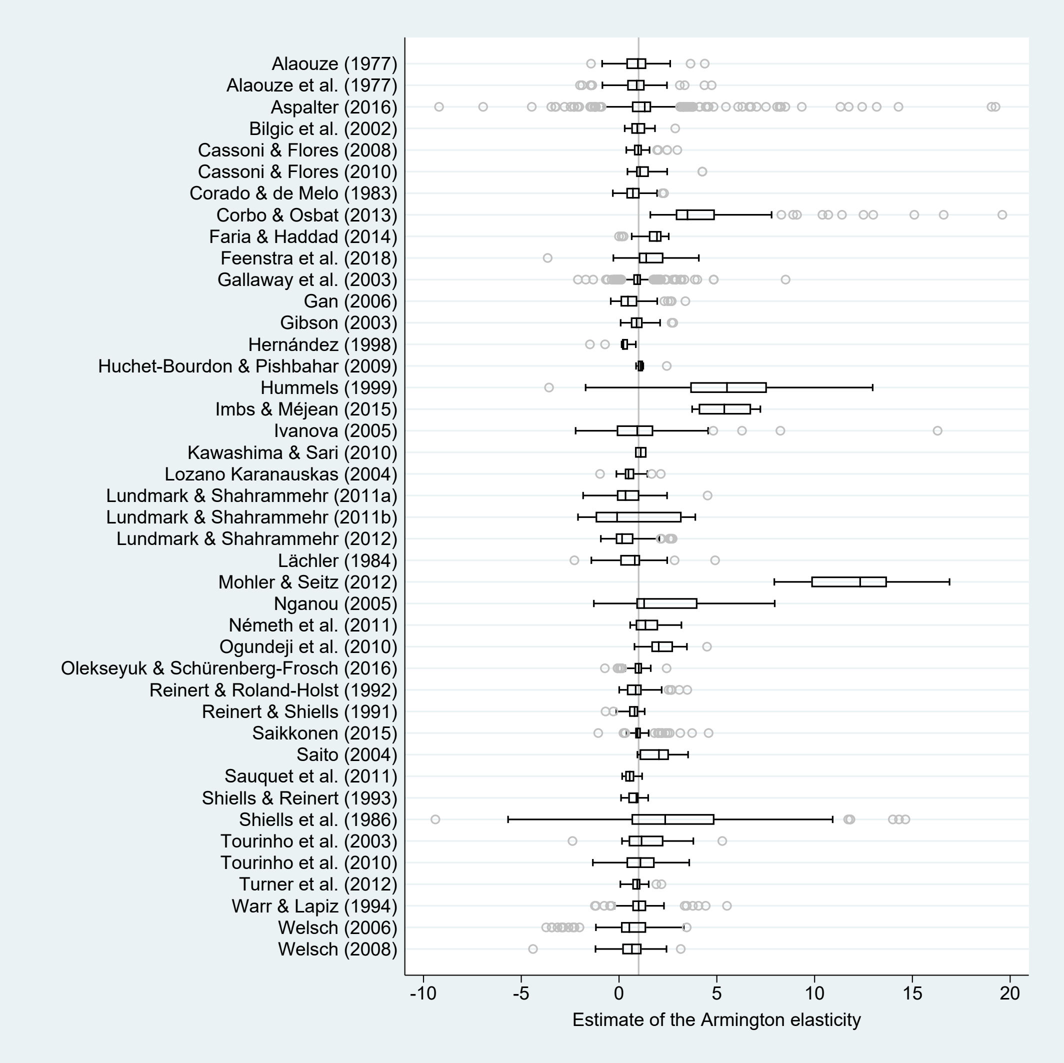
Notes: SD = standard deviation, WM = mean weighted by the inverse of the number of estimates reported per study, SIC = Standard Industrial Classification system for classifying industries by a four-digit code. Market size, tariff and non-tariff barriers, FX volatility, and internet usage have been collected from the World Bank database (WB 2018), data on national pride from the World Values Survey (Inglehart *et al.* 2014). The impact factor is downloaded from RePEc and the number of citations from Google Scholar. The rest of the variables are collected from studies estimating the Armington elasticity.

Table A.2: Potential mediating factors of publication bias

	Imp	Cit	Pub	IF+Cit+Pub	Disagg	All
Constant	0.737***	0.321	0.972***	0.431	1.386**	1.748***
	(0.109)	(0.326)	(0.253)	(0.298)	(0.622)	(0.535)
SE	0.809***	0.685***	0.757***	0.651***	0.761	-0.470
	(0.0888)	(0.138)	(0.0709)	(0.144)	(0.467)	(0.398)
SE * Impact f.	-0.257			-0.0436		-0.196
	(0.178)			(0.198)		(0.217)
SE * Citations		0.0138		-0.0130		0.0435
		(0.0645)		(0.0612)		(0.0738)
SE * Published			0.163	0.192		0.264
			(0.168)	(0.191)		(0.200)
SE * Data dis..					0.00810	0.151***
					(0.0624)	(0.0409)
Observations	2,968	2,968	2,968	2,968	2,968	2,968

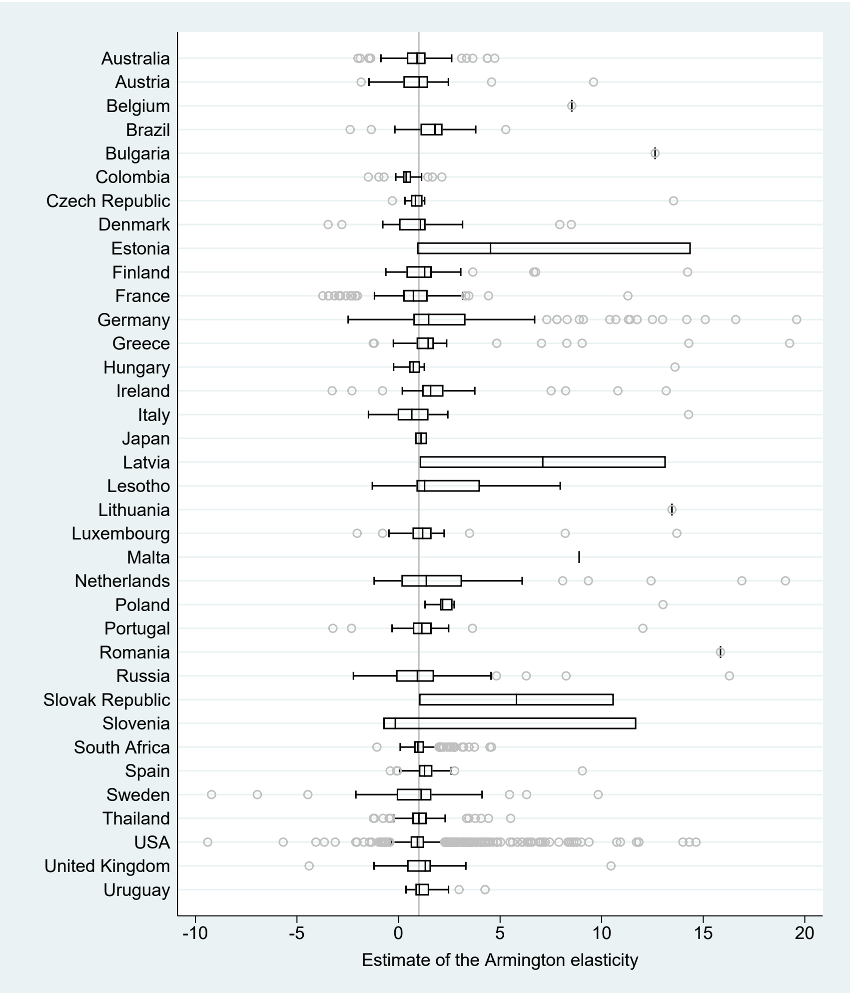
Notes: The response variable is an estimate of the long-run Armington elasticity. Standard errors are clustered at the study and country level. Variables interacted with the standard error (Impact factor, Citations, Published, Data disaggregation) are also included separately but the coefficients are not reported.

Figure A.1: Estimates vary both within and across studies



Notes: The figure shows a box plot of the estimates of the Armington elasticity reported in individual studies. The length of each box represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. Outliers are excluded from the figure. The solid vertical line denotes unity.

Figure A.2: Estimates vary both within and across countries



Notes: The figure shows a box plot of the estimates of the Armington elasticity reported for individual countries. The length of each box represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. Outliers are excluded from the figure. The solid vertical line denotes unity.

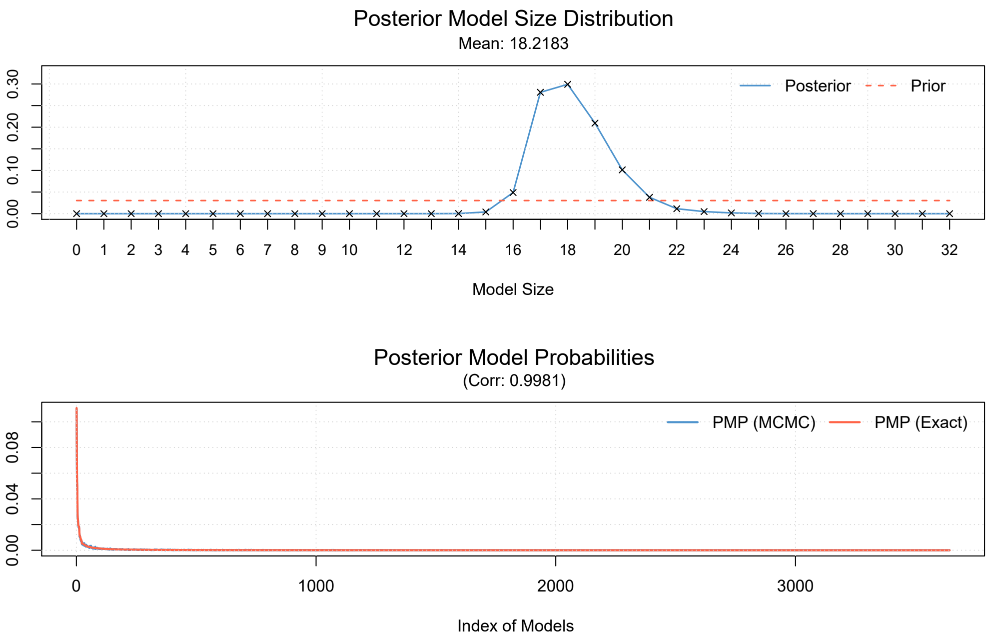
A.1 Diagnostics of BMA and Robustness Checks

Table B1: Diagnostics of the baseline BMA estimation

Mean no. regressors	*Draws*	*Burn-ins*	*Time*	*No. models visited*
18.2183	$3 \cdot 10^5$	$1 \cdot 10^5$	28.42869 secs	42,041
Modelspace	*Visited*	*Topmodels*	*Corr PMP*	*No. obs.*
$4.3 \cdot 10^9$	0.00098%	100%	0.9981	2,968
Model prior	*g-prior*	*Shrinkage-stats*		
Random / 16	UIP	Av = 0.9997		

Notes: We employ the g-prior suggested by Eicher *et al.* (2011) and model dilution prior suggested by George (2010). The results of this BMA exercise are reported in Table 3.

Figure B1: Model size and convergence of the baseline BMA estimation



Notes: The figure depicts the posterior model size distribution and the posterior model probabilities of the BMA exercise reported in Table 3.

Table B2: Why elasticities vary (alternative priors and weights)

Response variable:	Alternative BMA prior			Study-weighted BMA			Precision-weighted BMA		
Armington elasticity	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP
Constant	-1.58	NA	1.00	0.00	NA	1.00	1.16	NA	1.00
Standard error	0.75	0.03	1.00	0.70	0.03	1.00	2.86	0.26	1.00
Data characteristics									
Data disaggregation	0.22	0.05	1.00	0.24	0.03	1.00	-0.29	0.04	1.00
Results disaggregation	-0.24	0.04	1.00	-0.25	0.04	1.00	0.14	0.04	1.00
Monthly data	-0.41	0.19	0.89	-0.30	0.16	0.87	-0.01	0.04	0.07
Annual data	-1.07	0.15	1.00	-0.69	0.09	1.00	0.00	0.02	0.04
Time series	0.59	0.14	1.00	0.02	0.07	0.14	0.04	0.11	0.20
Cross-section	1.99	0.24	1.00	1.51	0.14	1.00	0.00	0.03	0.04
Data period	0.03	0.01	1.00	0.02	0.00	1.00	0.00	0.00	0.40
Data size	0.33	0.02	1.00	0.44	0.02	1.00	0.10	0.01	1.00
Midyear	0.00	0.00	0.03	-0.04	0.01	1.00	0.00	0.00	0.08
Structural Variation									
Secondary sector	0.00	0.01	0.02	-0.24	0.08	0.98	0.25	0.03	1.00
Tertiary sector	-0.01	0.07	0.04	-0.91	0.20	1.00	0.35	0.04	1.00
Developed countries	0.00	0.02	0.03	-1.02	0.12	1.00	-1.05	0.14	1.00
Market size	-0.10	0.02	1.00	-0.14	0.02	1.00	-0.07	0.02	1.00
Tariffs	0.03	0.01	1.00	0.00	0.01	0.21	-0.04	0.00	1.00
Non-tariff barriers	0.09	0.16	0.30	0.60	0.11	1.00	0.03	0.07	0.26
FX volatility	0.32	0.08	1.00	0.00	0.01	0.05	0.02	0.05	0.13
National pride	0.00	0.03	0.02	0.00	0.05	0.06	0.00	0.03	0.04
Internet usage	0.00	0.01	0.10	0.01	0.02	0.37	0.00	0.00	0.19
Estimation technique									
Static model	-0.01	0.04	0.04	-0.32	0.07	1.00	-0.81	0.08	1.00
Distributed lag and trend model	-0.01	0.06	0.05	-0.41	0.09	1.00	-1.04	0.10	1.00
Partial adjustment model	-0.05	0.12	0.22	0.00	0.03	0.05	-0.26	0.07	0.99
Nonlinear model	-0.01	0.07	0.05	1.40	0.21	1.00	0.28	0.25	0.63
OLS	-0.05	0.12	0.21	0.01	0.06	0.09	0.04	0.07	0.27
CORC	-0.01	0.06	0.06	0.02	0.07	0.11	-0.10	0.13	0.43
TSLS	0.24	0.23	0.57	-0.70	0.16	0.99	0.01	0.06	0.04
GMM	-0.08	0.16	0.22	-1.48	0.16	1.00	0.39	0.08	1.00
Import constraint	0.52	0.19	0.95	0.77	0.11	1.00	-0.01	0.04	0.05
Seasonality	-0.64	0.12	1.00	0.00	0.03	0.06	1.17	0.14	1.00
Publication characteristics									
Impact factor	0.13	0.20	0.351	0.00	0.03	0.057	1.04	0.14	1.000
Citations	0.60	0.05	1.000	0.48	0.05	1.000	-0.29	0.03	1.000
Published	0.57	0.11	1.000	1.17	0.09	1.000	0.21	0.16	0.691
Studies	39			39			39		
Observations	2,968			2,968			2,968		

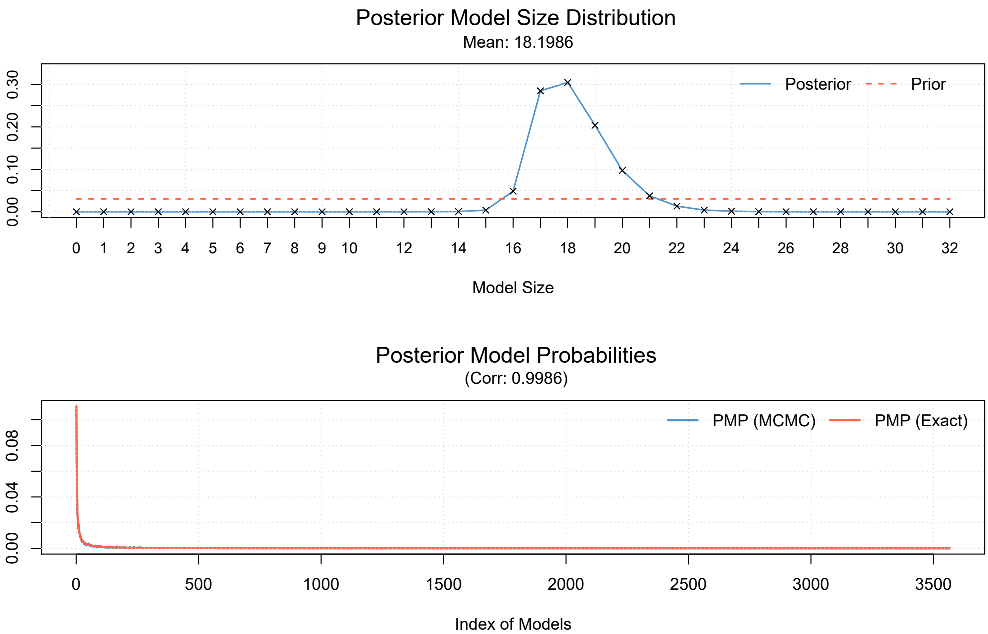
Notes: SD = standard deviation. PIP = posterior inclusion probability. All three panels represent results of Bayesian model averaging (BMA). The first panel employs an alternative model priors; Zellner's g prior is set according to Fernandez *et al.* (2001) and a dilution prior as suggested by George (2010). The two weighted specifications employ a uniform model prior suggested by Eicher *et al.* (2011) and a dilution prior suggested by George (2010). The study-weighted specification represents BMA applied to data weighted by the inverse of the number of observations reported per study, precision-weighted specification represents BMA applied to data weighted by the inverse of the standard error. All variables are described in Table A.1.

Table B3: Diagnostics of the alternative BMA prior

Mean no. regressors	*Draws*	*Burn-ins*	*Time*	*No. models visited*
18.1986	$3 \cdot 10^5$	$1 \cdot 10^5$	28.40841 secs	41,665
Modelspace	*Visited*	*Topmodels*	*Corr PMP*	*No. obs.*
$4.3 \cdot 10^9$	0.00097%	100%	0.9986	2,968
Model prior	*g-prior*	*Shrinkage-stats*		
Random / 16	BRIC	Av = 0.9997		

Notes: We employ the g-prior suggested by Fernandez *et al.* (2001) and model dilution prior suggested by George (2010). The results of this BMA exercise are reported in Table B2.

Figure B2: Model size and convergence of the alternative BMA prior



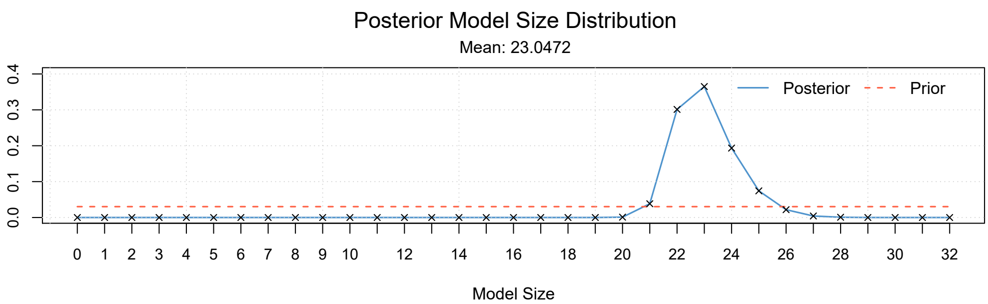
Notes: The figure depicts the posterior model size distribution and the posterior model probabilities of the BMA exercise with alternative g-prior according to Fernandez *et al.* (2001) reported in Table B2.

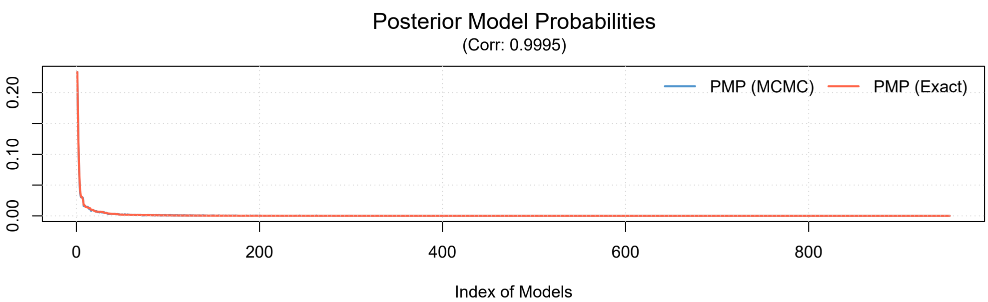
Table B4: Diagnostics of the study-weighted BMA estimation

Mean no. regressors	*Draws*	*Burn-ins*	*Time*	*No. models visited*
23.0472	$3 \cdot 10^5$	$1 \cdot 10^5$	23.9105 secs	33,872
Modelspace	*Visited*	*Topmodels*	*Corr PMP*	*No. obs.*
$4.3 \cdot 10^9$	0.00079%	100%	0.9995	2,968
Model prior	*g-prior*	*Shrinkage-stats*		
Random / 16	UIP	Av = 0.9997		

Notes: We employ the g-prior suggested by Eicher *et al.* (2011) and model dilution prior suggested by George (2010). The results of this BMA exercise are reported in Table B2.

Figure B3: Model size and convergence of the study-weighted BMA estimation





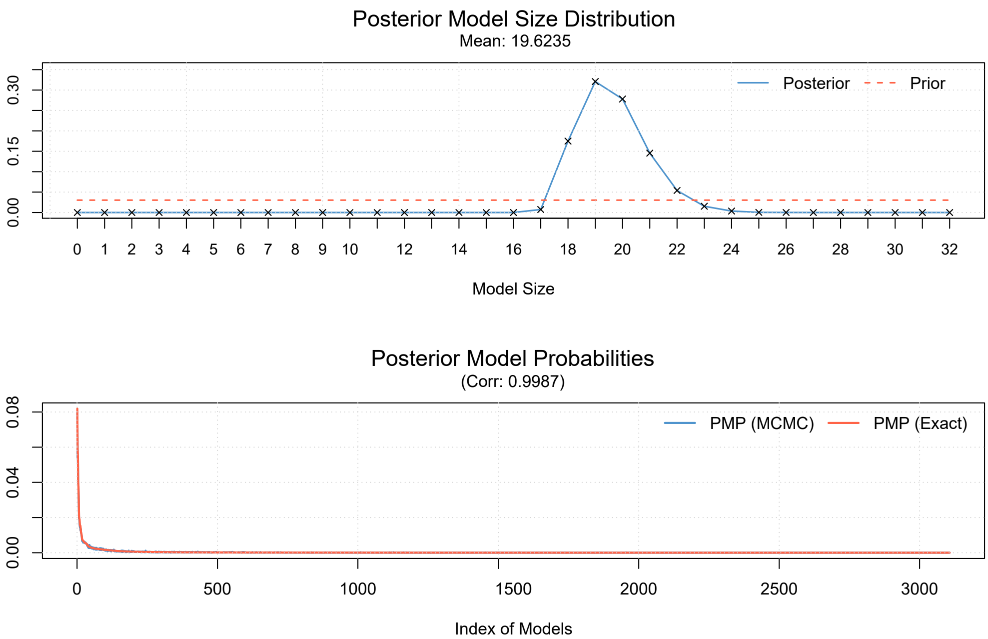
Notes: The figure depicts the posterior model size distribution and the posterior model probabilities of the study-weighted BMA exercise reported in Table B2.

Table B5: Diagnostics of the precision-weighted BMA estimation

Mean no. regressors	*Draws*	*Burn-ins*	*Time*	*No. models visited*
19.6235	$3 \cdot 10^5$	$1 \cdot 10^5$	29.30259 secs	44,536
Modelspace	*Visited*	*Topmodels*	*Corr PMP*	*No. obs.*
$4.3 \cdot 10^9$	0.001%	100%	0.9987	2,968
Model prior	*g-prior*	*Shrinkage-stats*		
Random / 16	UIP	Av = 0.9997		

Notes: We employ the g-prior suggested by Eicher *et al.* (2011) and model dilution prior suggested by George (2010). The results of this BMA exercise are reported in Table B2.

Figure B4: Model size and convergence of the precision-weighted BMA estimation



Notes: The figure depicts the posterior model size distribution and the posterior model probabilities of the precision-weighted BMA exercise reported in Table B2.

Appendix B

Appendix to Chapter 3

B.1 Data Collection and Fragmentation

B.1.1 Paper Selection Procedure

Similarly to other meta-analyses in economics, we searched for the most relevant primary studies via Google Scholar using the following query:

> "bank capital regulation" OR "capital requirements" OR "bank capital" OR "capital surplus" OR "capital ratio" OR "macroprudential regulation" OR "macroprudential policyâ€ť AND "lending" OR "credit" OR "loans"

Our search is limited to studies published in 2010 and later. We scan the first 300 papers returned in the search results. After initial screening, we expand our search by investigating the references from the relevant studies.[1] The most recent study was added in November 2020 when we concluded our search.

We collected 1,639 estimates from 46 studies[2], encompassing both articles published in refereed journals and working papers (see Table B1).[3] We iden-

[1]We screen all the references from the selected studies to identify additional relevant studies ("snowballing" method).

[2]Our sample size is comparable to similar meta-analyses. For example, Fidrmuc & Lind (2020) employed 312 estimates out of 48 primary studies on the effect of capital regulation on macroeconomic activity while Araujo *et al.* (2020) worked with 58 primary studies and more than 6,000 estimates in their extensive study of the effects of macroprudential policies on credit, household credit, and house prices.

[3]Besides the estimates and corresponding standard errors, t-statistics, p-value or confidence intervals, we collected about 40 other primary study characteristics. First, the data were collected and cross-examined by two of the co-authors of this paper. Afterwards, the two other co-authors cross-checked the whole dataset in several rounds to identify systematic or idiosyncratic errors and to ensure the consistency of the whole dataset.

tified several studies whose journal and working paper version differ in some aspects. Therefore, we collected both versions for each article when available and indicated the reason for the difference (data coverage, model specification or estimation technique). Only unique estimates then entered our final analysis.

Table B1: Journal Articles and Working Papers Included in the Meta-analysis

	Journal articles		Working papers	Differ?
1	Aiyar *et al.* (2014a)		-	-
2	Aiyar *et al.* (2016)	1	Aiyar *et al.* (2014b)	Y (M)
3	Akram (2014)	2	Akram (2012)	Y (M, P)
4	Auer *et al.* (2018)		-	-
5	Berrospide & Edge (2010b)	3	Berrospide & Edge (2010a)	Y (M)
	-	4	Berrospide *et al.* (2016)	-
	-	5	Bridges *et al.* (2014)	-
6	Brei *et al.* (2013)	6	Brei *et al.* (2011)	Y (A)
7	Buch & Prieto (2014)	7	Buch & Prieto (2012)	N
8	Carlson *et al.* (2013)	8	Carlson *et al.* (2011)	Y (M, P)
9	Cohen & Scatigna (2016)	9	Cohen (2013)	Y (C)
10	De Jonghe *et al.* (2020)	10	De Jonghe *et al.* (2016)	Y (M)
11	Deli & Hasan (2017)	11	Deli *et al.* (2017)	N
	-	12	De Nicolò (2015)	-
12	Drehmann & Gambacorta (2012)		-	-
	-	13	Galac *et al.* (2010)	-
13	Gambacorta & Marques-Ibanez (2011a)	14	Gambacorta & Marques-Ibanez (2011b)	N
14	Gambacorta & Shin (2018)	15	Gambacorta & Shin (2016)	N
15	Huang & Xiong (2015)		-	-
16	Imbierowicz *et al.* (2018)	16	Kragh & Rangvid (2016)	Y (M, P)
	-	17	Joyce & Spaltro (2014)	-
	-	18	Kanngiesser *et al.* (2017)	-
	-	19	Kolcunová & Malovaná (2019)	-
17	Kim & Sohn (2017)		-	-
18	Košak *et al.* (2015)	20	Košak *et al.* (2014)	N
	-	21	Labonne & Lamé (2014)	-
	-	22	Lambertini & Mukherjee (2016)	-
19	Malovaná & Frait (2017)		-	-
20	Meeks (2017)		-	-
21	Mésonnier & Stevanovic (2017)	23	Mésonnier & Stevanovic (2012)	Y (M, P)
22	Mora & Logan (2012)	24	Mora & Logan (2010)	Y (M)
23	Naceur *et al.* (2018)	25	Naceur *et al.* (2017)	Y (A)
24	Noss & Toffano (2016)	26	Noss & Toffano (2014)	N
	-	27	Olszak *et al.* (2014)	-
25	Roulet (2018)		-	-
	-	28	Wang & Sun (2013)	-
26	Watanabe (2010)	29	Watanabe (2006)	N

Note: Y/N – journal version and working paper do/do not differ; M – journal version and working paper use different model or methodology; P – the versions differ in time period examined; C – different number of countries is studied. Estimates that differ between journal article and working paper enter the meta-analysis; A – additional estimates that are not reported in the working paper; in this case only journal articles enter the analysis. If the estimates are the same, they enter the meta-analysis only once. Hence, the final set of studies comprises 26 journal articles and 20 working papers (29 working papers minus 7 that do not differ from the journal version, minus 2 that include fewer estimates than the journal version).

Each primary study included in our final data set meets two other criteria. First, the study reports some measure of uncertainty of its estimates (standard error, t-statistics, p-value or confidence intervals). Second, changes to bank capital ratios are measured continuously. For example, we do not consider studies using categorical or discrete variables to capture changes to bank capital ratios stemming from a regulatory shock. We are interested in the sensitivity of the credit dynamics to changes in the capital ratio in real terms, not in the frequency of use of the regulatory (prudential) policy (Cerutti *et al.* 2017a) or its direction (Cerutti *et al.* 2017b; Akinci & Olmstead-Rumsey 2018).[4] This is one of the main distinctions between our paper and Araujo *et al.* (2020), who perform meta-analysis relying on studies using various dummy-coded macroprudential indices. The dummy-coding of policy actions does not allow for an estimation of the quantitative effects of policies which is generally a key issue for policymakers (Alam *et al.* 2019). The study selection path is captured in Figure B1.

There are ten articles in our sample which report impulse response functions instead of regression semi-elasticities. In these cases we recover the numerical estimates and their confidence intervals via pixel coordinates. Specifically, we collect the immediate response, the effect after one period of time and the maximum response to the capital ratio shock. Following Fidrmuc & Korhonen (2006), we then treat each response collected as an estimate and differentiate between the contemporaneous, lagged and maximal effect by employing the corresponding dummies.

We need to make a few more adjustments to render the estimated semi-elasticities and corresponding standard errors comparable. First, we calculate the standard errors when t-statistic, p-value or confidence intervals are provided. Second, we adjust the semi-elasticities and corresponding standard errors to reflect annual changes. For example, if the semi-elasticity refers to a non-annualized quarterly change in credit, we multiply it by four. Likewise, we multiply the standard error. Third, we divide the cumulative effects by the respective number of periods. Fourth, some model specifications contain interaction terms with the respective capital ratio or additional lags of the capital ratio. We approach these semi-elasticities in the same way as the other, i.e. "stand-alone" semi-elasticities. We define respective dummy variables and analyze their significance in the model averaging exercise to control for the potential heterogeneity introduced by employing interaction terms or additional

[4]Existing cross-country databases use an almost predominantly dummy-coding approach which lacks information on the intensity of the change. A 2 percentage point change of say the counter-cyclical buffer is effectively treated at par with a 0.25 pp change. Notable exceptions of databases which capture changes to the regulatory policy continuously include Vandenbussche *et al.* (2015) and Richter *et al.* (2019); Alam *et al.* (2019) for loan-to-value (LTV) limits.

lags.[5] All semi-elasticities are expressed in percentage points.[6] Lastly, a few extreme outliers appear in the dataset, and we thus winsorize the estimates at the 2.5% level from each side.

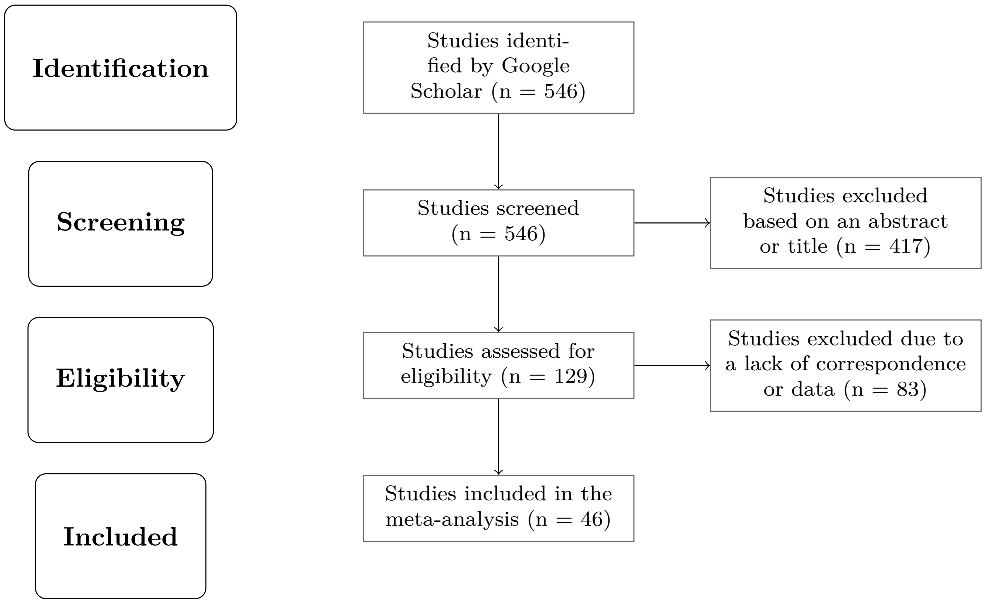
B.1.2 PRISMA Diagram

Figure B1 depicts the overall process employed in the selection of primary studies. In the *identification* phase, we scanned the first 300 research articles returned by Google Scholar using a tailor-made search query, limiting our search to papers published in or after 2010 (see B.1.1). We then went through all the citations in each of the relevant studies and identified an additional 246 articles, bringing us to a total of 546 articles for *screening*. In the next step, we reviewed all the titles and abstracts with the aim of effectively identifying studies that were not acceptable, even from a high-level perspective. In doing so, we eliminated 417 studies and assessed the remaining 129 for *eligibility*. During this step, we went through each article in more detail and filtered out 83 studies due to a lack of correspondence or data. The main elimination criteria were: (1) the study must report numerical results; (2) estimated semi-elasticities must be presented together with the corresponding test statistic – standard error, t-statistic, p-value or exact confidence interval; (3) the effect is not a cross-boarder effect; (4) the measure of lending cannot be expressed as a ratio to some other continuous variable such as total loans or total bank assets; and finally, (5) the variable capturing capital cannot be expressed by a dummy-coded index (such as in Cerutti *et al.* 2017b). All in all we ended up with 46 primary studies *included* in the meta-analysis.

[5]We consider only interaction terms with dummy variables where the effect can be easily separated. We do not include semi-elasticities linked to interaction with continuous variables since we would not be able to separate the effect of changes in capital ratio from changes in the continuous variable.

[6]We collected information on units in which variables were expressed from the primary studies which we then used to correctly transform the semi-elasticities.

Figure B1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Diagram



B.1.3 Additional Summary Statistics According to Different Characteristics

Table B2: Which Capital Requirements Were Explored in Primary Studies

Article	Capital requirements	Country	Period	Mean	(95% CI)
Bridges *et al.* (2014)	Req. cap. ("trigger") ratio	UK	1990–2014	-3.98	(-11.56, 0.26)
De Jonghe *et al.* (2016)	Pillar 2 capital requirements	BE	2013–2015	-1.05	(-3.15, 0.20)
De Jonghe *et al.* (2020)	Pillar 2 capital requirements	BE	2013–2015	-1.69	(-4.85, -0.16)
Kolcunová & Malovaná (2019)	Overall capital requirements	CZ	2013–2017	-0.57	(-1.60, 0.27)
Meeks (2017)	Req. cap. ("trigger") ratio	UK	1975–2008	-0.80	(-3.48, 1.10)

Note: The required capital ("trigger") ratio is an analogy to Pillar 2 capital requirements, i.e. bank-specific capital requirements set by the Bank of England (before 2001) and the FSA. In 2001, the trigger ratio was renamed Individual Capital Guidance (ICG) and became part of the Pillar 2 process under Basel II (Bridges *et al.* 2014, see). Overall, regulatory capital requirements refer to a combination of minimum capital requirements, Pillar 2 capital add-ons and additional capital buffers (the systemic risk buffer and countercyclical capital buffer).

Table B3: Partial Correlation Coefficients for Different Variable Transformations and Capital Ratios

	Obs.	Articles	Mean	Median	5%	95%	Skew.
Total	*1,639*	*46*	*0.004*	*0.005*	*-0.106*	*0.114*	*-0.20*
Variable transformation							
Credit growth $\sim \beta \times$ capital ratio	1,395	32	0.008	0.010	-0.098	0.090	-0.03
Other transformations	244	15	-0.024	-0.045	-0.283	0.213	0.28
Different capital ratios (all transformations)							
Capital-to-asset ratio	559	20	0.023	0.016	-0.083	0.153	0.69
Regulatory capital ratio	710	22	0.000	0.016	-0.112	0.088	-1.59
Capital requirements	337	9	-0.029	-0.009	-0.114	0.012	-1.34

Note: The partial correlation coefficient (PCC) from i^{th} estimate of the j^{th} study can be derived from the t-statistics of the reported estimates and residual degrees of freedom: $PCC_{ij} = t_{ij}/\sqrt{(t_{ij}^2 - df_{ij})}$. The regulatory capital ratio represents the ratio between regulatory capital (Common Equity Tier 1, Additional Tier 1 and Tier 2 items) and risk-weighted exposures. Capital requirements represent the ratio between various categories of capital requirements (minimum, Pillar 2 add-ons and capital buffers) and risk-weighted exposures. Some articles include multiple different variable transformations and capital variables; therefore, the sum of articles across different categories reported in the third column exceeds the number of primary studies included. In addition, some articles use the level of capital (not expressed in relation to banksâ€™ assets or risk-weighted exposures); therefore, the sum of observations across different capital ratios in the second column is lower than the total number of collected estimates.

B.2 Extensions to the Publication Bias

In this section, we present result of two additional methods for estimating publication bias – the random effects model and the three-level model. The random effects model assumes that effect sizes vary both within and between studies, and uses estimates of within-study and between-study variance to obtain an overall effect size estimate:

$$\beta_{i,j} = \alpha + \delta_i + \gamma SE_{i,j} + \epsilon_{i,j} \tag{B.1}$$

where $\beta_{i,j}$ is the ith estimated semi-elasticity and $SE_{i,j}$ its standard error for each study j; α is the effect beyond bias and γ is the intensity of publication bias. The new term δ_i is a normally distributed random effect $\delta_i \sim N(0, \tau^2)$ with τ^2 being the between-study variance in "true" effects.

The three-level model extends the random effects model by adding a third level of analysis to account for variation in effect sizes due to differences in primary study characteristics, such as differences in data or methods:

$$\beta_{i,j} = \alpha + \delta_i + \kappa_{i,j} + \gamma SE_{i,j} + \epsilon_{i,j} \tag{B.2}$$

The new term $\kappa_{i,j}$ accounts for within-cluster (within-study j) heterogeneity and is also normally distributed.

The quadratic models of PEESE are extended with the same two terms δ_i and $\kappa_{i,j}$. All models are estimated with weights equal to the inverse of the estimate's variance to control for heteroskedasticity. Table B1 reports the regression results.

Table B1: Estimation of Publication Bias – Additional Method

	Cap.-to-Asset Ratio	Regulatory Cap. Ratio	Capital Req.
Panel A: FAT-PET			
Meta-analysis random effects			
Effect beyond bias (1/SE)	0.394**	0.294***	-0.620**
	(0.167)	(0.091)	(0.145)
Publication bias (constant)	-0.071	-0.306	-0.743***
	(0.390)	(0.376)	(0.091)
I^2 (%)	67	93	53
Meta-analysis three-level model			
Effect beyond bias (1/SE)	1.014***	0.279***	-0.384***
	(0.175)	(0.053)	(0.081)
Publication bias (constant)	-0.668***	-0.439***	-0.885***
	(0.116)	(0.118)	(0.119)
I^2 level 1 (%)	3	4	40
I^2 level 2 (%)	2	30	25
I^2 level 3 (%)	95	66	35
Panel B: PEESE			
Meta-analysis random effects			
Effect beyond bias (1/SE)	0.372***	0.230***	-0.857***
	(0.083)	(0.054)	(0.132)
Publication bias (SE)	-0.001	-0.052	-0.111
	(0.011)	(0.081)	(0.065)
I^2 (%)	66	93	57
Meta-analysis three-level model			
Effect beyond bias (1/SE)	0.535***	0.184***	-0.755***
	(0.115)	(0.043)	(0.062)
Publication bias (SE)	-0.010	-0.057	-0.112***
	(0.008)	(0.042)	(0.030)
I^2 level 1 (%)	5	4	38
I^2 level 2 (%)	4	32	34
I^2 level 3 (%)	91	64	28
Observations	514	652	229
Studies	16	18	5
Observations per study (mean)	32	36	46

Note: Standard errors, clustered at the study level, are reported in parentheses. I^2 measures the share of the effect heterogeneity as a percentage of total variance. For the random effects model, I^2 measures between-study variance in true effect. For the three-level model, I^2 measures the amount of heterogeneity variance within studies (level 2) and between studies (level 3). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Extensions to the Analysis of Heterogeneity

B.3.1 Variables Used in the Meta-analysis

Characteristics of the Primary Study

Data characteristics. The literature provides little prior knowledge regarding the impact of different data characteristics on the relationship between bank capital and lending. Nevertheless, the existing meta-analytic studies on monetary policy transmission identify significant discrepancies caused by different data frequency and length of the data sample (Havranek & Rusnak 2013; Ehrenbergerova *et al.* 2021). Given the character of the relationship studied, we extended this set of control variables significantly. Specifically, we account for the type of credit used as a dependent variable, data frequency, the number of observations, the midpoint of the data, the region of the analysis, and data confidentiality. We also distinguish between different panel-data structures: bank-level vs macro-level and multi-country vs single-country.

Model specification and estimation. As a next step, we control for different aspects of the econometric approach used in the primary study. A number of meta-analytic studies proved that these factors play a significant role in the direction and size of the estimated semi-elasticities (Zigraiova *et al.* 2021). First, we are interested in the model specification. We distinguish between a static and dynamic model[7], models with different lag structures or models missing some key control variables. We also search for more specific factors, such as the presence of additional capital variables[8] or interaction term[9].

[7]A dynamic model contains a lagged dependent variable, in our case credit growth.

[8]Additional capital variables included in the same estimation equation, on top of the capital ratio, may distort the relationship between bank capital and lending studied. Specifically, the presence of two related variables may result in the effects going in opposite direction, which makes it difficult to distill the correct sign and size of the relationship between capital and lending. For example, the same estimation equation may include, among other combinations, both a simple capital-to-asset ratio and regulatory capital ratio or regulatory capital ratio and capital requirements. In each case, we consider the ratio more related to the capital regulation to be our main semi-elasticity (i.e. the regulatory capital ratio in the first example and capital requirements in the second example).

[9]By imposing interaction terms, the researcher explores heterogeneity in the effect analyzed. Distinguishing between crisis and non-crisis periods is among the most frequent interactions implemented. So far, the results in the literature on the impact of the crisis period on the relationship between bank capital and lending have been mixed. Some studies find the relationship to be strong and statistically significant only in crisis periods (Gambacorta & Marques-Ibanez 2011a; Carlson *et al.* 2013; Kim & Sohn 2017), others show quite the opposite in the same equation (Bridges *et al.* 2014; Naceur *et al.* 2018).

Second, we explore the impact of different estimation techniques. As a part of that, we distinguish between specifications including time and unit fixed effects.

Publication characteristics. This group of characteristics is expected to correlate with unobserved features related to the relevance and quality of the primary study. These include the year of publication, an indication of whether the study was published in a journal, the discounted recursive impact factor, the number of citations and an indication of whether this or some other version of the study was published by a central bank. The reasons to control for these characteristics are supported by the literature. For instance, Araujo *et al.* (2020) find that journal articles show some signs of publication bias while non-published studies do not. Moreover, the lower the impact factor, the higher the publication bias. In other words, the effect may be overestimated in the series with lower quality. Further, Fabo *et al.* (2021) show that reported findings may be affected by the institution which publishes them.

Reference model. We have defined our control variables against a reference model which is based on the predominant characteristics of the primary studies. For example, we define three dummy variables capturing the estimation technique (GMM, FE and other techniques) while we treat OLS, the predominant technique, as the reference group. We take the liberty of choosing the baseline characteristics more freely when the difference between the groups is not large or when we want to make the interpretation more intuitive. The predominant characteristics can be found in Table B1.

Structural (External) Variables

On top of the primary study characteristics, we consider nine external variables capturing cross-country or cross-regional differences.[10] Some of these factors may not have been accounted for in the primary study, and therefore, may play a significant role in the heterogeneity observed among reported estimates.

First, we include four variables closely connected with the **monetary policy** stance: the three-month interbank rate, the spread between the ten-year government bond yield and the three-month interbank rate, a variable measuring the number of consecutive years during which interest rates are very low, and an index of central bank independence. In general, studies show that the monetary policy stance may significantly affect the relationship between bank

[10]Some other potentially relevant external variables (for example, GDP growth, ratio of exports and imports to GDP, credit growth, financial development index, or measure of regulatory quality) were excluded due to high multicollinearity.

capital and lending (Gambacorta & Marques-Ibanez 2011a; De Jonghe *et al.* 2020) and that monetary and macroprudential policy are not independent, as they affect both the monetary and credit conditions via their effect on credit growth (Malovaná & Frait 2017; Akinci & Olmstead-Rumsey 2018; Gambacorta & Murcia 2020). Furthermore, a prolonged period of low interest rates may lead to a build-up of financial vulnerabilities (Malovaná *et al.* 2022) to which the macroprudential policy has to react by tightening its stance while, at the same time, monetary policy may become less effective in stimulating credit growth (Borio & Gambacorta 2017). In this respect, the independence of central banks promotes financial stability (Klomp & De Haan 2009) which, in its core, underlines the overall demand for macroprudential measures and might affect their effectiveness.

Second, we employ some **macroeconomic and macro-financial variables**: unemployment rate, deviation of consumer inflation from its long-run trend[11] and growth in house prices. We include the first two variables to control for the overall economic conditions and the third one to capture macro-financial linkages and the average position in the financial cycle. Multiple studies document that the reaction of bank lending to changes in bank capitalization differs over the course of the business and financial cycle (Gambacorta & Marques-Ibanez 2011a; Brei *et al.* 2013). Moreover, (Fitzpatrick & McQuinn 2007; Anundsen *et al.* 2016) document a mutually reinforcing relationship between credit growth and house prices.

Third, we explore the impact of two additional variables linked to the character of the **domestic financial system**: the ratio of bank assets to GDP (as a proxy for the size of the banking sector) and an index of financial openness. De Jonghe *et al.* (2020) show that the larger the banking sector in terms of its assets, the more detrimental the effect of increasing bank capital on credit growth. Also, several studies find differences between the impact of capital regulation on credit between advanced and emerging market economies (Cerutti *et al.* 2017a; Akinci & Olmstead-Rumsey 2018; Alam *et al.* 2019). They offer various explanations but one may suspect that the relative size of the financial sector could also be important. The degree of financial development and openness also affects the relationship studied. On the one hand, Deli & Hasan (2017) show that a countryâ€™s financial development and openness reinforces the link between the capitalization of banks and credit growth as these countries

[11]We approximate the long-run trend of inflation using its mean value and we set the band around which it deviates to +1/-1 pp.

are less constrained in raising the capital. On the other, Cerutti *et al.* (2017a) find evidence of weaker associations between capital regulation and credit in financially more open or developed economies.

Most external variables enter the analysis as a simple average calculated for the same time period as was employed in the primary study.[12] There are two exceptions: the low for long variable and the deviation of consumer inflation from its long-run trend. These two variables represent a number of consecutive periods in a given time frame that meet a certain criteria (see Table B1).

[12]For example, if the semi-elasticity of the relationship between bank capital and lending comes from a primary study employing a time period between 2000 and 2014, we calculate the simple average of the external variables for the time period between 2000 and 2014 as well. For multi-country panel data, we either used an aggregate provided by a respective database or we calculated one using single-country data series.

Table B1: Description and Summary Statistics of MR Variables

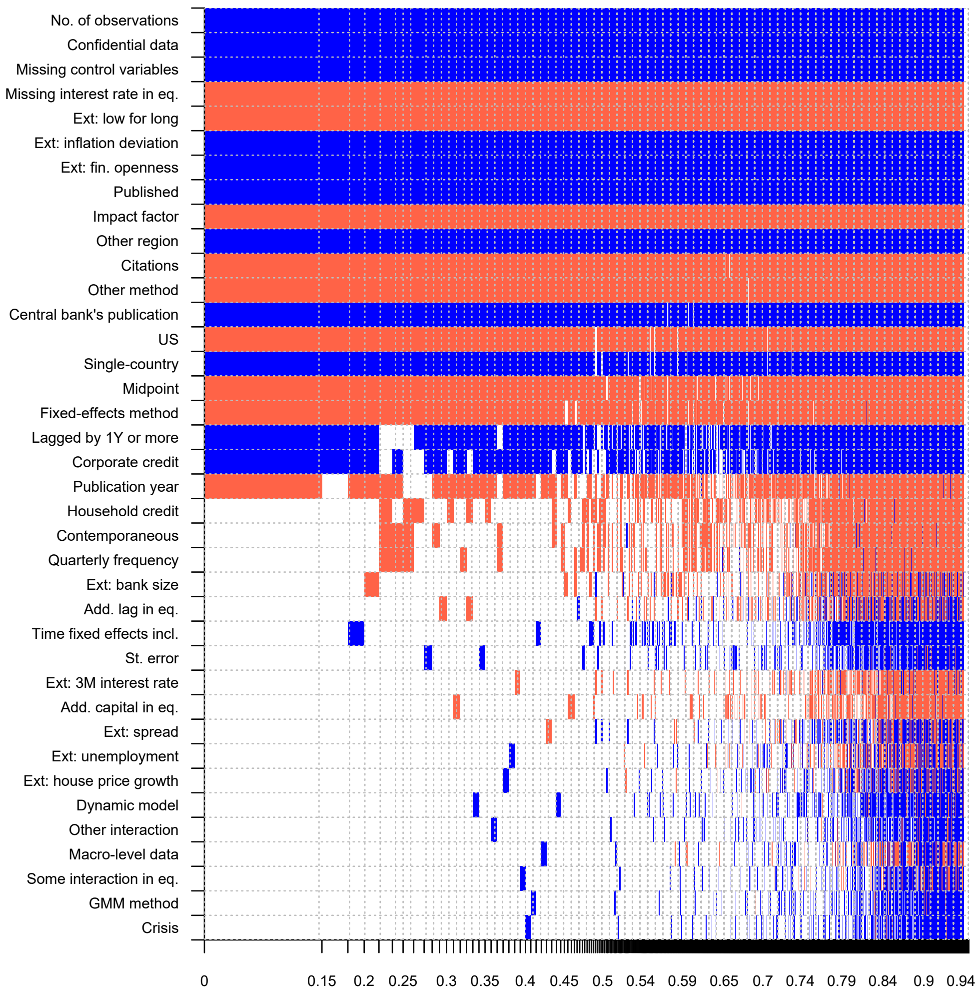
Variable	Description	Mean (St. Dev.)					
		Capital-to-Asset Ratio		Reg. Capital Ratio		Capital Req.	
Estimate	The reported estimate of the beta coefficient.	0.34	(1.38)	0.14	(0.80)	-1.74	(2.59)
St. error	The reported standard error of the beta coefficient.	1.22	(4.58)	0.49	(1.24)	1.52	(2.37)
Data characteristics							
Total credit	= 1 if total credit is used as a dependent variable.	0.38	(0.49)	0.56	(0.50)	0.10	(0.30)
Corporate credit	= 1 if corporate credit is used as a dependent variable.	0.35	(0.48)	0.21	(0.41)	0.79	(0.41)
Household credit	= 1 if household credit is used as a dependent variable.	0.27	(0.45)	0.22	(0.42)	0.12	(0.32)
Midpoint	The logarithm of the midpoint of the data sample.	3.27	(0.17)	3.28	(0.16)	3.22	(0.79)
No. of observations	The logarithm of the total number of observations.	7.53	(1.63)	7.67	(1.60)	10.56	(3.66)
Quarterly frequency	= 1 if data frequency is quarterly.	0.11	(0.31)	0.21	(0.41)	1.00	(0.00)
Confidential data	= 1 if confidential (supervisory) data are used (as opposed to publicly available data).	0.07	(0.26)	0.19	(0.39)	0.33	(0.47)
Macro-level data	= 1 if macro-level data are used (as opposed to bank-level data).	0.39	(0.49)	0.31	(0.46)	0.00	(0.00)
Single-country	= 1 if the study covers a single country (as opposed to a cross-country study).	0.45	(0.50)	0.54	(0.50)	1.00	(0.00)
US	= 1 if the study covers only US.	0.34	(0.47)	0.44	(0.50)	0.00	(0.00)
Other region	= 1 if the study covers countries outside the US and Europe or the sample includes a mix of countries from different regions.	0.15	(0.36)	0.25	(0.44)	0.00	(0.00)
Model specification and estimation							
GMM method	= 1 if the general method of moments (GMM) is used.	0.33	(0.47)	0.13	(0.34)	0.17	(0.38)
Fixed-effects method	= 1 if the fixed-effects (FE) regression method is used.	0.17	(0.37)	0.35	(0.48)	0.68	(0.47)
Other method	= 1 if a method other than the OLS, GMM or FE method is used.	0.04	(0.19)	0.15	(0.35)	0.15	(0.36)
Time fixed effects incl.	= 1 if time fixed effects are included.	0.42	(0.49)	0.23	(0.42)	0.85	(0.36)
Dynamic model	= 1 if the model is dynamic, i.e., contains a lagged dependent variable.	0.46	(0.50)	0.44	(0.50)	0.33	(0.47)
Lagged by 1Y or more	= 1 if the estimate is lagged by a year (4 quarters) or more.	0.69	(0.46)	0.77	(0.42)	0.38	(0.49)
Contemporaneous	= 1 if the estimate is contemporaneous (not lagged at all).	0.21	(0.40)	0.03	(0.18)	0.03	(0.16)
Missing control variables	= 1 if the model is missing either supply-side (banks-specific) or demand-side (macroeconomic) control variables.	0.18	(0.39)	0.21	(0.40)	0.85	(0.36)
Missing interest rate in eq.	= 1 if the model is missing interest rate as a control variable.	0.42	(0.49)	0.21	(0.41)	0.85	(0.36)
Crisis	= 1 if the estimate is interacted with a crisis dummy variable.	0.10	(0.30)	0.14	(0.35)	0.02	(0.13)
Other interaction	= 1 if the estimate is interacted with another dummy variable.	0.18	(0.38)	0.23	(0.42)	0.23	(0.42)
Add. capital in eq.	= 1 if the model contains an additional capital variable on top of the studied capital ratio.	0.27	(0.44)	0.38	(0.49)	0.98	(0.13)
Add. lag in eq.	= 1 if the model contains additional lag(s) of the studied capital ratio.	0.01	(0.11)	0.01	(0.08)	0.62	(0.49)
Some interaction in eq.	= 1 if the model contains some interaction term (discrete or continuous) with the studied capital ratio.	0.33	(0.47)	0.39	(0.49)	0.52	(0.50)

Continued Table B1.

Variable	Description	Mean (St. Dev.)					
		Capital-to-Asset Ratio		Regulatory Capital Ratio		Capital Requirements	
Publication characteristics							
Publication year	The logarithm of the publication year of the primary study minus the earliest publication year in our dataset plus one.	1.85	(0.56)	1.88	(0.44)	2.10	(0.29)
Impact factor	The recursive impact factor.	0.68	(0.63)	0.67	(0.57)	1.07	(0.84)
Citations	The logarithm of the number of citations divided by the number of years from its publication until 2021.	2.38	(0.81)	2.79	(0.56)	2.77	(0.51)
Published	= 1 if the primary study was published in a journal with an impact factor.	0.75	(0.43)	0.84	(0.37)	0.43	(0.50)
Central bank publication	= 1 if the primary study was published by a central bank.	0.26	(0.44)	0.25	(0.44)	1.00	(0.00)
External variables							
Ext: 3M interest rate	The average 3-month interest rate in percent.	2.71	(2.04)	2.52	(1.65)	1.81	(3.16)
Ext: spread	The average spread (difference between 10-year government bond yield and 3-month interest rate) in percentage points.	-1.80	(0.59)	-1.69	(0.65)	-1.05	(0.77)
Ext: low for long	The number of consecutive quarters during which the 3M interest rate is below its first quartile.	9.80	(4.59)	9.53	(4.62)	6.73	(7.08)
Ext: inflation deviation	The number of consecutive quarters during which CPI inflation is outside the +/- 1 percentage point band around its long-term mean.	21.11	(14.23)	14.81	(6.77)	12.04	(19.50)
Ext: house price growth	The average annual house price growth in percent.	1.90	(4.38)	1.98	(4.93)	2.70	(3.28)
Ext: unemployment	The average unemployment rate in percent.	8.10	(1.56)	7.50	(1.58)	7.94	(1.22)
Ext: bank size	The average ratio between banking sector assets and GDP in percent.	89.00	(33.21)	78.53	(19.48)	84.14	(20.94)
Ext: fin. openness	The average financial openness index (Chinn & Ito 2008).	2.00	(0.37)	1.99	(0.60)	2.28	(0.23)

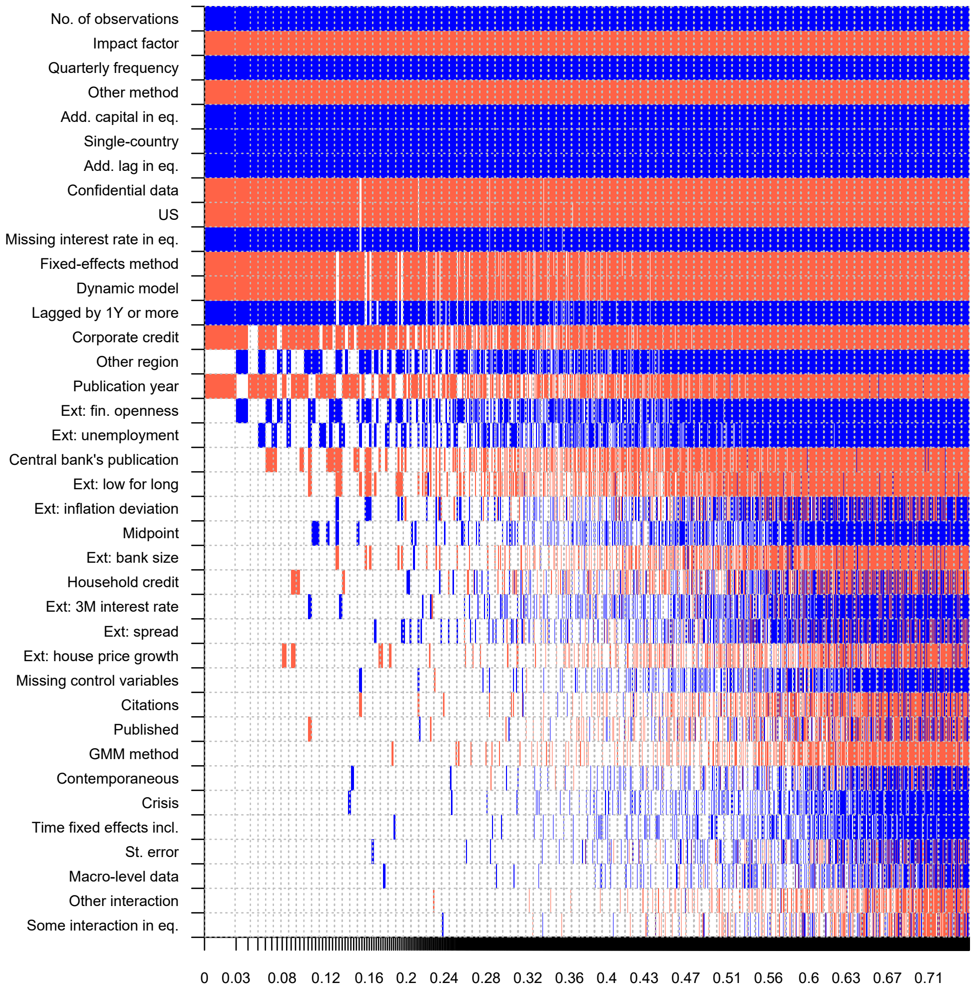
B.3.2 Bayesian Model Averaging

Figure B1: BMA Results – Capital-to-Asset Ratio



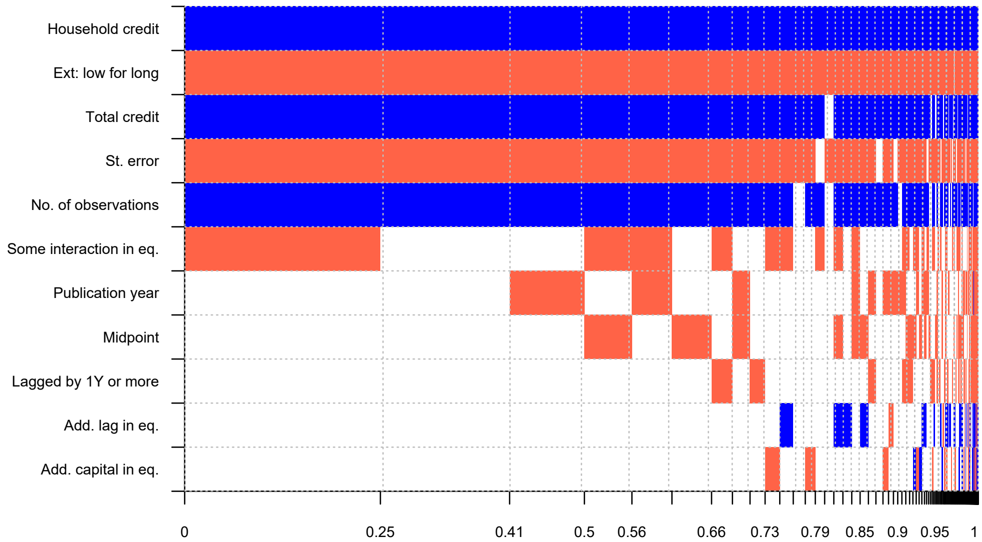
Note: The response variable is the estimated effect of a 1 percentage point change in capital-to-asset ratio on credit growth. The columns denote the individual models; the variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes the cumulative posterior model probabilities; the 10,000 best models are shown. To ensure convergence we employ 3 million iterations and 1 million burn-ins. Blue (darker in the grayscale) indicates that the variable is included and the estimated sign is positive, i.e., the transmission is stronger, given that the mean effect is positive. Red (lighter in the grayscale) indicates that the variable is included and the estimated sign is negative, i.e., the transmission is weaker, given that the mean effect is positive. No color indicates that the variable is not included in the model.

Figure B2: BMA Results – Regulatory Capital Ratio



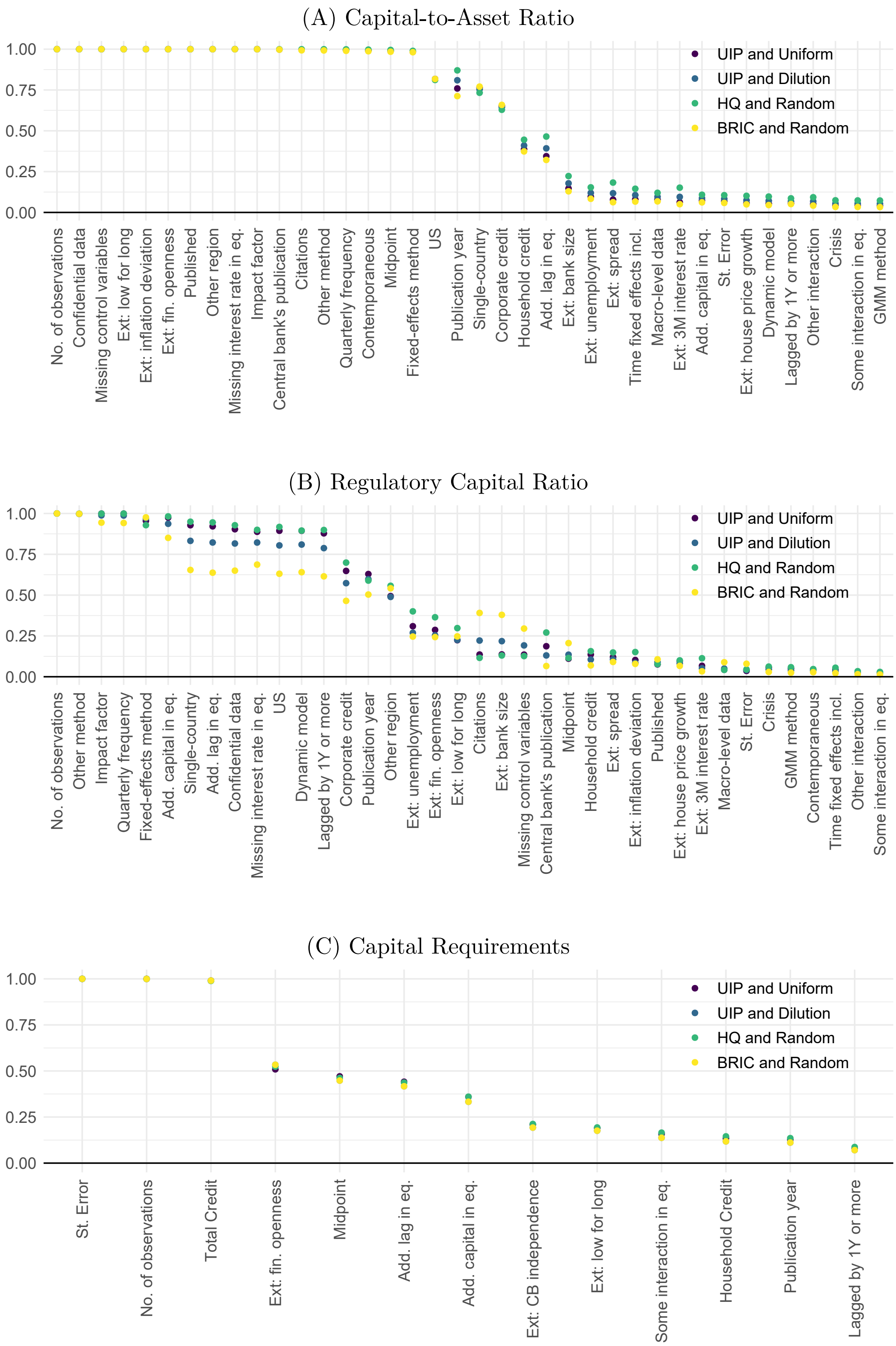
Note: The response variable is the estimated effect of a 1 percentage point change in the regulatory capital ratio on credit growth. The columns denote the individual models; the variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes the cumulative posterior model probabilities; the 10,000 best models are shown. We employ 3 million iterations and 1 million burn-ins to ensure convergence. Blue (darker in the grayscale) indicates that the variable is included and the estimated sign is positive, i.e., the transmission is weaker, given that the mean effect is negative. Red (lighter in the grayscale) indicates that the variable is included and the estimated sign is negative, i.e., the transmission is stronger, given that the mean effect is negative. No color indicates that the variable is not included in the model.

Figure B3: BMA Results – Capital Requirements



Note: The response variable is the estimated effect of a 1 percentage point change in capital requirements on credit growth. The columns denote the individual models; the variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes the cumulative posterior model probabilities; the 10,000 best models are shown. We employ 3 million iterations and 1 million burn-ins to ensure convergence. Blue (darker in the grayscale) indicates that the variable is included and the estimated sign is positive, i.e., the transmission is weaker, given that the mean effect is negative. Red (lighter in the grayscale) indicates that the variable is included and the estimated sign is negative, i.e., the transmission is stronger, given that the mean effect is negative. No color indicates that the variable is not included in the model.

Figure B4: Bayesian Model Averaging – Prior Sensitivity



Note: The figures show posterior inclusion probability for different prior combinations. In our baseline, we use a unit information g-prior (UIP) and a uniform model prior which reflects our lack of prior knowledge. The uniform model prior gives each model the same prior probability, and the unit information g-prior provides the same information as one observation from the data. As a robustness check, we use a dilution model prior, as proposed by (George 2010), to account for potential collinearity between explanatory variables. Next, we also employ a combination of the Hannan-Quinn (HQ) g-prior and random model prior and a combination of the BRIC g-prior and random model prior. The HQ g-prior adjusts data quality while the BRIC g-prior minimizes the prior effect on the results. The random model prior gives equal prior probability to every model size (Gechert *et al.* 2022).

B.3.3 Frequentist Model Averaging and Frequentist Check

Table B2: FMA and OLS Results – Capital-to-Asset Ratio

	Frequentist model averaging			Frequentist check (OLS)		
	Coef.	SE	p-value	Coef.	SE	p-value
Constant	21.918	15.240	0.150	-3.880	6.142	0.536
St. Error	0.006	0.019	0.750			
Data characteristics						
No. of observations	0.264	0.049	0.000	0.259	0.054	0.000
Confidential data	6.940	1.434	0.000	6.624	0.774	0.000
Quarterly frequency	-2.107	1.002	0.035			
Other region	3.780	1.802	0.036	5.515	0.998	0.000
Midpoint	-9.750	5.209	0.061	-3.673	2.324	0.134
Corporate credit	0.473	0.430	0.271	0.707	0.258	0.015
Macro-level data	-0.188	0.214	0.382			
US	-1.465	2.512	0.560	-2.352	0.672	0.003
Household credit	-0.237	0.422	0.575			
Single-country	0.452	2.275	0.842	3.450	0.848	0.001
Model specification and estimation						
Missing control variables	2.980	0.638	0.000	3.243	0.757	0.001
Missing interest rate in eq.	-2.398	0.646	0.000	-2.625	0.945	0.013
Other method	-11.121	4.258	0.009	-8.861	1.935	0.000
Fixed-effects method	-3.854	1.574	0.014	-1.929	0.446	0.001
Contemporaneous	-1.843	1.067	0.084			
Lagged by 1Y or more	1.351	0.828	0.103	2.859	0.796	0.002
Add. lag in eq.	-4.631	3.043	0.128			
Time fixed effects incl.	0.197	0.134	0.142			
Crisis	0.400	0.476	0.400			
Other interaction	0.364	0.448	0.417			
Dynamic model	0.283	0.355	0.426			
GMM method	0.147	0.202	0.469			
Some interaction in eq.	-0.178	0.323	0.581			
Add. capital in eq.	0.183	0.442	0.679			
Publication characteristics						
Impact factor	-1.440	0.417	0.001	-1.513	0.259	0.000
Published	3.539	1.069	0.001	3.579	0.639	0.000
Citations	-2.504	0.836	0.003	-1.912	0.397	0.000
Central bank publication	3.427	1.622	0.035	2.564	0.341	0.000
Publication year	-1.660	0.813	0.041	-1.259	0.545	0.035
External variables						
Ext: fin. openness	7.943	2.413	0.001	6.593	0.889	0.000
Ext: inflation deviation	0.230	0.100	0.022	0.332	0.057	0.000
Ext: low for long	-0.284	0.142	0.045	-0.447	0.094	0.000
Ext: 3M interest rate	-0.745	0.531	0.160			
Ext: spread	1.035	0.834	0.215			
Ext: house price growth	-0.022	0.030	0.453			
Ext: unemployment	0.129	0.270	0.634			
Ext: bank size	-0.001	0.026	0.960			
			Adj. R^2		0.690	

Note: The table presents the estimation results of the collected estimate of the beta coefficient on the primary study characteristics and external (structural) variables, searching for potential sources of heterogeneity. The estimation is based on 514 observations from 17 studies. Frequentist model averaging applies Mallowâ€™s weights (Hansen 2007) using the orthogonalization of covariate space suggested by Amini & Parmeter (2012) to reduce the number of estimated models. The frequentist check (OLS) includes the variables with posterior inclusion probability (PIP) estimated by Bayesian model averaging (BMA) above 0.5 and is estimated using standard errors clustered at the study level. A description and summary of all the variables is provided in Table B1.

Table B3: FMA and OLS Results – Regulatory Capital Ratio

	Frequentist model averaging			Frequentist check (OLS)		
	Coef.	SE	p-value	Coef.	SE	p-value
Constant	-13.024	5.862	0.026	-2.046	0.623	0.004
St. Error	-0.054	0.040	0.177			
Data characteristics						
No. of observations	0.242	0.026	0.000	0.236	0.040	0.000
Quarterly frequency	3.541	0.657	0.000	3.138	0.457	0.000
Confidential data	-2.340	0.715	0.001	-2.194	0.318	0.000
Single-country	3.709	1.077	0.001	1.993	0.182	0.000
US	-2.090	0.677	0.002	-1.187	0.217	0.000
Other region	0.872	0.327	0.008	0.398	0.214	0.081
Corporate credit	-0.361	0.172	0.035	-0.175	0.068	0.020
Midpoint	2.452	1.371	0.074			
Household credit	-0.200	0.171	0.240			
Macro-level data	-0.072	0.116	0.536			
Model specification and estimation						
Add. capital in eq.	0.835	0.227	0.000	1.009	0.166	0.000
Lagged by 1Y or more	0.808	0.254	0.001	0.513	0.211	0.026
Add. lag in eq.	3.094	0.924	0.001	1.794	0.186	0.000
Other method	-0.790	0.311	0.011	-1.337	0.163	0.000
Contemporaneous	0.599	0.377	0.112			
Fixed-effects method	-0.418	0.286	0.143	-0.884	0.193	0.000
Time fixed effects incl.	0.099	0.085	0.247			
Crisis	0.253	0.220	0.251			
Missing interest rate in eq.	0.230	0.205	0.263	0.562	0.152	0.002
Dynamic model	-0.259	0.278	0.353	-1.068	0.122	0.000
GMM method	-0.095	0.116	0.412			
Other interaction	0.164	0.215	0.446			
Some interaction in eq.	-0.146	0.225	0.516			
Missing control variables	0.132	0.261	0.614			
Publication characteristics						
Impact factor	-0.428	0.115	0.000	-0.555	0.095	0.000
Central bank publication	-1.712	0.634	0.007			
Publication year	-0.812	0.440	0.065	-0.252	0.175	0.168
Citations	-0.319	0.270	0.238			
Published	0.321	0.340	0.345			
External variables						
Ext: fin. openness	0.968	0.366	0.008	0.210	0.164	0.218
Ext: low for long	-0.102	0.047	0.030			
Ext: 3M interest rate	0.340	0.192	0.077			
Ext: spread	-0.360	0.278	0.196			
Ext: inflation deviation	0.024	0.021	0.252			
Ext: unemployment	0.097	0.109	0.371			
Ext: house price growth	0.007	0.011	0.526			
Ext: bank size	0.006	0.011	0.579			
			Adj. R^2		0.539	

Note: The table presents the estimation results of the collected estimate of the beta coefficient on the primary study characteristics and external (structural) variables, searching for potential sources of heterogeneity. The estimation is based on 652 observations from 18 studies. Frequentist model averaging applies Mallowâ€™s weights (Hansen 2007) using the orthogonalization of covariate space suggested by Amini & Parmeter (2012) to reduce the number of estimated models. The frequentist check (OLS) includes the variables with posterior inclusion probability (PIP) estimated by Bayesian model averaging (BMA) above 0.5 and is estimated using standard errors clustered at the study level. A description and summary of all the variables is provided in Table B1.

Table B4: FMA and OLS Results – Capital Requirements

	Frequentist model averaging			Frequentist check (OLS)		
	Coef.	SE	p-value	Coef.	SE	p-value
Constant	-1.430	1.451	0.324	-2.715	1.382	0.121
St. error	-0.366	0.158	0.020	-0.245	0.318	0.485
Data characteristics						
Total credit	2.726	1.956	0.163	2.301	0.717	0.033
No. of observations	0.153	0.121	0.203	0.175	0.088	0.118
Household credit	1.963	1.621	0.226	3.743	0.857	0.012
Midpoint	-0.475	0.441	0.281			
Model specification and estimation						
Add. lag in eq.	0.937	0.965	0.332			
Some interaction in eq.	-0.413	0.453	0.362			
Lagged by 1Y or more	0.060	0.281	0.831			
Add. capital in eq.	0.120	0.906	0.895			
Publication characteristics						
Publication year	-0.248	0.741	0.738			
External variables						
Ext: low for long	-0.048	0.114	0.673	-0.172	0.041	0.014
			Adj. R^2		0.619	

Note: The table presents the estimation results of the collected estimate of the beta coefficient on primary study characteristics and external (structural) variables, searching for potential sources of heterogeneity. The estimation is based on 229 observations from 5 studies. Frequentist model averaging applies Mallowâ€™s weights (Hansen 2007) using the orthogonalization of covariate space suggested by Amini & Parmeter (2012) to reduce the number of estimated models. The frequentist check (OLS) includes the variables with posterior inclusion probability (PIP) estimated by Bayesian model averaging (BMA) above 0.5 and is estimated using standard errors clustered at the study level. A description and summary of all the variables is provided in Table B1.

Appendix C

Appendix to Chapter 4

Figure B1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram

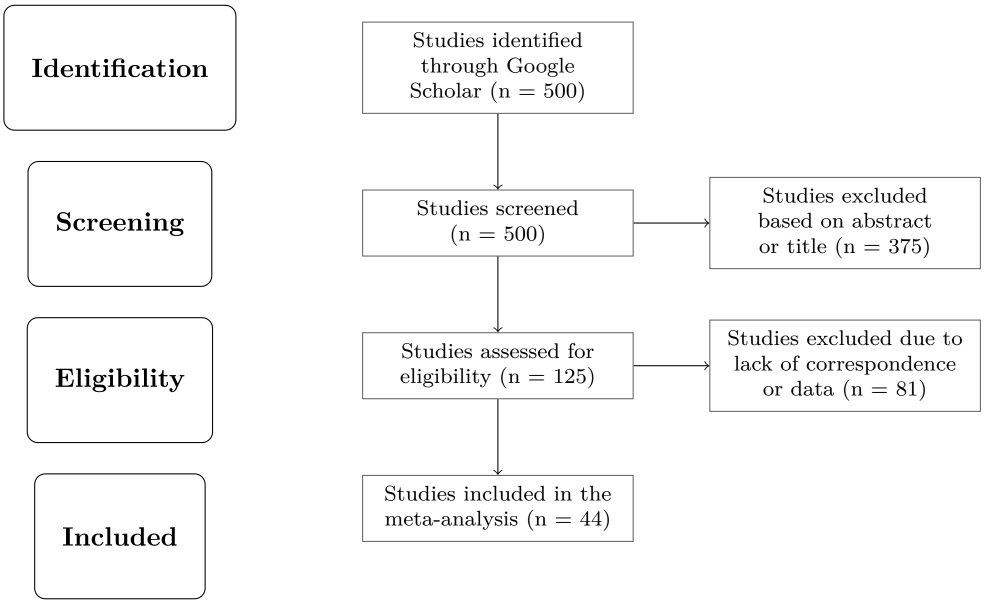


Table B1: Indication of publication bias

	Log-log cases	Log-level cases
PANEL A: Unweighted estimations		
OLS		
SE (publication bias)	0.186	1.467***
	(0.190)	(0.240)
Constant (effect beyond bias)	-0.036	-0.089
	(0.065)	(0.158)
Between effects		
SE (publication bias)	0.463***	0.737***
	(0.155)	(0.271)
Constant (effect beyond bias)	-0.032	0.177
	(0.046)	(0.164)
PANEL B: Weighted OLS estimations		
Weighted by the inverse of the number of estimates reported per study		
SE (publication bias)	0.312	0.999***
	(0.237)	(0.285)
Constant (effect beyond bias)	-0.011	0.103
	(0.038)	(0.104)
Weighted by the inverse of the standard error		
SE (publication bias)	-0.222	1.206***
	(0.429)	(0.257)
Constant (effect beyond bias)	-0.001***	-0.000
	(0.000)	(0.000)
Observations	217	231

Notes: The table above displays the results of the regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \epsilon_{it}$, where S_{it} denotes the ith estimate of the effect size in study j and $SE(SE_{it})$ stands for the respective standard error. Specification (1) uses OLS. Specification (2) employs a panel data regression with between effects. The estimates in specification (3) use WLS with precision weights. Similarly, specification (4) uses the reciprocal of the number of estimates reported per study as the weights. Standard errors are in parentheses and clustered at the country and study level (except between effects; the usage of two-way clustering follows Cameron *et al.* 2012). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B2: Why the estimates vary – various robustness checks

Response variable:	BMA – UIP prior			BMA – BRIC prior			Frequentist model averaging			BMA – dilution – Top		
	Post. mean	Post. SD	PIP	Coef.	SE	PIP	Coef.	SE	p-value	Coef.	SE	PIP
Constant	0.175	NA	1.000	0.151	NA	1.000	0.209	0.101	0.039	0.150	NA	1.000
Winsorized SE	-0.004	0.067	0.057	0.000	0.028	0.013	-0.304	0.395	0.441	-0.001	0.030	0.016
Data characteristics												
Contemporaneous	0.001	0.006	0.078	0.000	0.002	0.014	0.027	0.019	0.154	0.000	0.002	0.018
Price change	0.000	0.007	0.050	0.000	0.003	0.012	0.028	0.040	0.480	0.000	0.005	0.018
Absolute returns	0.029	0.034	0.484	0.045	0.037	0.639	0.026	0.030	0.393	0.046	0.037	0.674
Abnormal returns	0.207	0.044	1.000	0.194	0.037	1.000	0.207	0.045	0.000	0.195	0.041	1.000
Excess returns	-0.062	0.038	0.804	-0.076	0.027	0.941	-0.068	0.037	0.065	-0.075	0.031	0.914
Dollar volume	-0.014	0.024	0.312	-0.002	0.011	0.062	-0.044	0.026	0.088	-0.008	0.019	0.174
Shares traded	-0.009	0.020	0.221	-0.003	0.013	0.082	-0.034	0.030	0.266	-0.007	0.019	0.153
Detrended series	-0.035	0.034	0.584	-0.016	0.028	0.281	-0.058	0.031	0.059	-0.032	0.036	0.519
Data period	0.000	0.000	0.052	0.000	1.000	0.011	0.001	0.001	0.437	0.000	0.000	0.015
Data size	0.020	0.003	1.000	0.020	0.003	1.000	0.018	0.005	0.000	0.021	0.003	1.000
Midyear	-0.057	0.013	1.000	-0.053	0.011	1.000	-0.058	0.018	0.001	-0.050	0.013	0.989
Daily data	0.001	0.010	0.075	0.000	0.005	0.018	0.030	0.040	0.449	0.000	0.005	0.023
Weekly data	0.022	0.037	0.315	0.015	0.033	0.198	0.085	0.055	0.125	0.015	0.033	0.207
Monthly data	-0.171	0.034	1.000	-0.166	0.026	1.000	-0.184	0.054	0.001	-0.175	0.032	1.000
Time series	0.042	0.042	0.592	0.015	0.031	0.229	0.057	0.037	0.121	0.018	0.031	0.293
Cross section	0.052	0.070	0.432	0.012	0.040	0.111	0.118	0.056	0.033	0.025	0.054	0.214
Structural variation												
Index	-0.013	0.027	0.235	-0.003	0.015	0.057	-0.042	0.034	0.223	-0.004	0.017	0.083
NASDAQ	0.001	0.009	0.053	0.000	0.005	0.014	0.025	0.037	0.500	0.000	0.005	0.014
Banks	-0.005	0.020	0.104	-0.001	0.009	0.024	-0.051	0.047	0.283	-0.002	0.014	0.047
Firms	0.054	0.029	0.838	0.072	0.022	0.960	0.021	0.037	0.561	0.064	0.024	0.930
Developing	0.001	0.006	0.069	0.000	0.004	0.021	-0.035	0.027	0.196	0.000	0.004	0.024
Market size	-0.003	0.006	0.264	-0.001	0.003	0.069	-0.008	0.008	0.300	-0.002	0.006	0.152
Asia	0.001	0.006	0.089	0.000	0.003	0.025	0.034	0.033	0.309	0.006	0.018	0.140
Europe	-0.001	0.009	0.068	0.000	0.005	0.019	-0.005	0.034	0.879	0.000	0.006	0.025
Australia	0.001	0.011	0.059	0.000	0.005	0.014	-0.012	0.051	0.808	0.002	0.014	0.037
Other Continents	0.263	0.048	1.000	0.277	0.043	1.000	0.324	0.064	0.000	0.275	0.044	1.000
Observations	468			468			468			468		
Studies	44			44			44			44		

Response variable = partial correlation between trading volume and stock returns. SD = standard deviation. PIP = posterior inclusion probability. SE = standard error. In BMA uniform g-prior with uniform model prior and BRIC g-prior with random model prior are deployed (Eicher *et al.* 2011, Fernandez *et al.* (2001)). FMA employs Mallow's criterion for model averaging using the orthogonalization of the parameter space (Hansen 2007, Amini & Parmeter 2012). The BMA – dilution – Top model deploys UIP g-prior and dilution model prior as suggested George (2010). Besides publication characteristics are captured by variable *Top* in the last model. All the variables describes Table 4.6.

Table B2: Why the estimates vary – various robustness checks (continued)

Response variable:	BMA – UIP prior			BMA – BRIC prior			Frequentist model averaging			BMA – dilution – Top		
	Post. mean	Post. SD	PIP	Coef.	SE	PIP	Coef.	SE	p-value	Coef.	SE	PIP
Estimation technique												
VAR	-0.110	0.030	0.988	-0.120	0.026	0.997	-0.155	0.075	0.040	-0.116	0.027	0.996
Simple model	0.001	0.005	0.059	0.000	0.003	0.016	0.019	0.024	0.411	0.001	0.006	0.032
GARCH	0.001	0.013	0.062	0.000	0.006	0.016	-0.070	0.074	0.344	0.000	0.006	0.018
Monday	0.002	0.013	0.074	0.000	0.005	0.015	0.017	0.041	0.681	0.001	0.006	0.020
Trimmed	0.021	0.031	0.383	0.007	0.019	0.131	0.053	0.027	0.052	0.009	0.022	0.187
January Excluded	0.003	0.012	0.116	0.001	0.006	0.028	0.040	0.023	0.077	0.002	0.009	0.054
MLE	0.005	0.017	0.116	0.001	0.008	0.029	0.050	0.031	0.111	0.002	0.010	0.042
GMM	0.001	0.012	0.067	0.000	0.005	0.018	0.065	0.061	0.286	0.000	0.005	0.021
Other methods	-0.081	0.023	0.986	-0.094	0.018	0.993	-0.058	0.022	0.009	-0.092	0.017	1.000
Publication characteristics												
Impact factor	-0.001	0.005	0.102	0.000	0.003	0.026	0.017	0.022	0.438	-	-	-
Citations	-0.003	0.007	0.194	-0.001	0.004	0.045	-0.010	0.012	0.419	-	-	-
Published	0.000	0.005	0.062	0.000	0.002	0.015	0.019	0.028	0.499	-	-	-
Top	-	-	-	-	-	-	-	-	-	0.010	0.024	0.188
Observations	468			468			468			468		
Studies	44			44			44			44		

Response variable = partial correlation between trading volume and stock returns. SD = standard deviation. PIP = posterior inclusion probability. SE = standard error. In BMA uniform g-prior with uniform model prior and BRIC g-prior with random model prior are deployed (Eicher *et al.* 2011, Fernandez *et al.* (2001)). FMA employs Mallow's criterion for model averaging using the orthogonalization of the parameter space (Hansen 2007, Amini & Parmeter 2012). The BMA – dilution – Top model deploys UIP g-prior and dilution model prior as suggested George (2010). Besides publication characteristics are captured by variable *Top* in the last model. All the variables describes Table 4.6.

Appendix D

Referee Reports

Firstly, I would like to thank my supervisor, Prof. PhDr. Tomáš Havránek, Ph.D., together with all three referees for the time they spent reading my dissertation and writing the reports. I am very grateful for their valuable comments and suggestions. As all the referees, including the supervisor, suggest just points to the discussion and recommended the thesis for defense without substantial changes, the committee decided that all three articles might remain without changes, and that the suggestions and discussions should be included in the form of one or two pages long discussion in the newly written Chapter 5. Finally, to clarify the raised points in detail, I extend the Chapter 5 up to six pages. Moreover, the points raised toward the introduction and formal part of the thesis were incorporated also. The reports are included below and typed in Roman. The responses of the author are in *italics*.

D.1 Supervisor's report

Prof. PhDr. Tomáš Havránek Ph. D., Institute of Economic Studies, Charles University.

a) Can you recognize an original contribution of the author? **YES**

b) Is the thesis based on relevant references? **YES**

c) Is the thesis defendable at your home institution or another respected institution where you gave lectures? **YES**

d) Do the results of the thesis allow their publication in a respected economic journal? **YES, already published**

e) Are there any additional major comments on what should be improved? **NO**

f) What is your overall assessment of the thesis? **(a) I recommend the thesis for defense without substantial changes**

Josef Bajzík selected three papers for his dissertation, but has co-authored at least eight finished papers that could have been included. His publications are impressive: Journal of International Economics, IMF Economic Review, Journal of Financial Stability, Journal of Economic Surveys (3x), International Review of Financial Analysis, and others. Some of these publications are solo-authored. Such a publication record would warrant tenure at most Czech universities. Moreover, the first two publications (JIE and IMF ER) would be considered a success for a PhD student at any of the world's leading economics departments. All papers included in the dissertation have already been published in good journals – which means that they went through several rounds of revisions in the peer-review process. Therefore I raise no requirements for further revisions. Overall, I have no choice but to recommend Josef Bajzík's dissertation for defense without major changes. I will use the rest of the space allocated to me in this report to reflect upon Josef's work and studies.

Josef is capable of good independent research work, which is documented by his solo publications, one of which is included in the dissertation. I appreciate his assertiveness: when he was a BA student, he skillfully lobbied me to provide him with a trainee position at the Czech National Bank. There we started to work on a long-term project that was eventually published in the Journal of International Economics. Josef has stayed at the Czech National Bank, in the Research Department, and has been very productive. He is also very efficient with his time and has a good drive in terms of publishing his papers. If he stays in the research industry (as I hope he will, either at the national bank or in academia), he will prosper. His assertiveness is perhaps sometimes a bit excessive, which at times has created tensions with co-authors. But I appreciate how Josef has been able to handle and remedy these issues. He has been stress-tested as a researcher and a colleague and is well prepared for further professional work.

It is also appropriate to mention the awards that Josef has received for his work. In 2020, he was honored with the Karel Englis Prize by the Czech

Economic Society for the best paper related to the Czech economic policy. Also in 2020, he received the Czech National Bank research award. According to RePEc, Josef ranks 34th among Czech economists (when only publications from the last 10 years are considered). According to Google Scholar, his work has received more than 200 citations. Josef has also received several grants from the Grant Agency of Charles University. He has also refereed for Journal of Economic Surveys (repeatedly), International Review of Financial Analysis, Czech Journal of Economics and Finance, and the National Bank of Slovakia. All these awards and statistics are impressive for a PhD student.

The dissertation itself consists of three papers. The first paper, on the Armington elasticity, was published in the Journal of International Economics. The paper shows that, after correction for publication bias and misspecification in some studies, the trade elasticity is on average likely to be around 3, but depends on the context (the paper provides many estimates for various contexts). The second paper, published in the Journal of Economic Surveys, focuses on the nexus between bank capital and lending. My understanding of the paper is that there is no robust relationship between capital and lending. Finally, the third, solo-authored paper provides a meta-analysis of the literature on trading volume and stock returns. The paper focuses on publication bias and heterogeneity. One thing I do not understand is the conclusion chapter (Chapter 5), which does not contain any information. I think it can be deleted. Also, the footnote to each chapter (such as on page 8) should clearly list all co-authors of the chapter. Apart from that, I see no problems and think that the dissertation can be defended as it is.

Author's response: *Amended. I list clearly the names of all respective authors at the beginning of each chapter. The conclusion part is finally not deleted, but expanded due to the comments of other referees and decision of the committee of the incorporation of the raised discussion there.*

D.2 Report from referee 1

Dr.-Ing. Jerome Geyer-Klingeberg (University of Augsburg, Germany)

a) Can you recognize an original contribution of the author?
With his dissertation, Josef Bajzík has made a commendable, original contribution to international and financial economics by applying meta-analysis as a method for systematic quantitative reviews to summarize

and compare the existing but ambiguous empirical evidence in these two research fields. His findings provide a benchmark for what is known and guide new research on the Armington elasticity, the effect of changes in capital-based measures on lending, and the predictability of the stock returns based on trading volume performances. His original contribution is especially reflected by the fact that all of his three thesis chapters are published in leading international research journals: Journal of International Economics, Journal of Economic Surveys, and International Review of Financial Analysis.

b) Is the thesis based on relevant references?
The thesis is well-grounded in relevant references. The author demonstrates a thorough understanding of the existing literature, utilizing key papers from the relevant research fields to support his arguments. The referenced articles are in line with the research focus of the thesis and reflect a comprehensive review of the relevant scholarly studies in the field.

c) Is the thesis defendable at your home institution or another respected institution where you gave lectures?
Mr. Josef Bajzík's thesis would without any doubt also be a defendable and excellent thesis at my university (University of Augsburg, Applied Data Analysis Group). His research is methodologically sound, and the findings are well-supported by the data. He effectively addresses potential challenges and demonstrates a strong command of the advanced metaanalytical methods he is applying. Overall, the thesis meets high standards of academic rigor and quality.

d) Do the results of the thesis allow their publication in a respected economic journal?
The three main chapters of the thesis are already published in well-respected economic journals: Journal of International Economics, Journal of Economic Surveys, and International Review of Financial Analysis.

e) Are there any additional major comments on what should be improved?
No major comments for improvement. The thesis is well written and shows a high degree of scientific quality. The research design, the methodology, the presentation and discussion of the results are all well-crafted and reflect a high level of scientific expertise in the field of meta-analysis.

Josef Bajzík has also effectively addressed potential concerns and limitations.

f) What is your overall assessment of the thesis?
I recommend the thesis for defense without substantial changes.

The scientific community, especially also in economics and finance, has witnessed a substantial surge in empirical research outputs over the past decades, often resulting in a deluge of empirical findings on the same or similar research questions. In light of this remarkable increase in research results, a key challenge emerges: How do we discern valuable and accurate information amidst potential exaggerations, biases, or even misinformation? Meta-analysis is one of the promising and well-established approaches to tackle these challenges. It is a statistical investigation of research findings related to a specific hypothesis, empirical effect, phenomenon, or policy intervention. It constitutes a systematic and quantitative review encompassing all available scientific knowledge on a given topic, which plays a vital role in evidence-based practices across disciplines such as medicine, education, and the social sciences. Meta-analysis stands out as an objective and statistically rigorous approach to systematic reviews, thereby furnishing the evidential basis for the broader evidence-based practice movement. In disciplines like social science, particularly economics, there exists considerable variation in reported estimates for a given parameter or scientific phenomenon. Given that economics hinges on human behavior, the empirical phenomena studied will inherently exhibit substantial variation and academic pressure to publish novel findings can intensify conflicts among empirical results. Therefore, without an intelligent synthesis of economics research, informed policy actions would be impractical.

Chapter 2 of Mr. Bajzík's thesis "Estimating the Armington Elasticity: The Importance of Data Choice and Publication Bias" is a study in international economics that was published in the highly reputable Journal of International Economics. In this chapter, he applies meta-analysis on a very large sample of 3,524 reported estimates of the Armington elasticity. This chapter examines three important aspects via a state-of-the-art meta-analysis: (a) the overrepresentation of large and statistically significant elasticities, and (b) the drivers of the substantial heterogeneity in the literature, and c) the mean elasticity implied by the literate after correcting for publication bias and potential misspecifications.

The insights presented in this chapter are novel, shedding new light on the ongoing research discourse regarding the welfare effects of globalization, particularly in examining the responsiveness of demand for domestic versus foreign goods to changes in relative prices. Mr. Bajzík's application of the latest meta-regression methods to this pivotal research question in international economics is characterized by competence, comprehensiveness, and rigor. Mr. Bajzík effectively disentangles the intricate theoretical and empirical dimensions surrounding the elasticity of substitution between domestic and foreign goods. Notably, he masters the application of an advanced meta-regression toolkit, encompassing Bayesian Model Averaging, non-linear publication bias testing, and best practice estimation. In this regard, Mr. Bajzík excels, demonstrating a nuanced understanding of how his sophisticated meta-regression holds important implications for both the research community and policymakers. Each meta-method is meticulously applied, resulting in clear and well-described outcomes. The discussion successfully guides the reader in comprehending the added value delivered by Mr. Bajzík's study. I have personally conducted several meta-regression analyses and reviewed multiple meta-studies in economics and finance and I can say that Chapter 2 compares favorably with the best.

I find Chapter 3 "Bank Capital, Lending and Regulation: A Meta-analysis", published in the Journal of Economic Surveys, another excellent application of a high quality meta-regression analysis. The data set is rich with 1,600 estimates on the relationship between bank capital and lending, as well as 40 variables of heterogeneity. Mr. Bajzík's second application of metaregression is, again, quite comprehensive and well-researched. He shows a great deal of judgement and deep knowledge of the relevant literature in the choices of this moderator variables. The applied meta-methods are well selected, including Bayesian Model Averaging and extensive publication bias testing using a wide range of advanced linear and non-linear testing methods. I am very impressed how Mr. Bajzík's applies the most sophisticated and recently developed tests for publication bias and excels in the interpretation of the results. Chapter 3 present the third meta-analysis of Mr. Bajzík's thesis conducted on the relationship between trading volume and stock return. It provides a comprehensive and rigorous investigation of 468 estimates from 44 studies. The chapter addresses publication bias and employs Bayesian and frequentist model averaging to explain the heterogeneity in the reported estimates. The research critically evaluates the historical evolution of studies on trading volume and stock returns, highlighting the increasing variance in estimates over time. The meta-analysis not only corrects for publication bias but also explores the impact of data characteristics, structural variations, and methodological approaches, providing valuable insights for both researchers and practitioners in the field of finance. The findings emphasize the importance of considering the specifics of each stock when forming portfolios, calibrating models, devising trading strategies, or conducting further research. Overall, Mr. Bajzík's work contributes significantly to the understanding of this complex relationship and serves as a valuable resource of knowledge accumulation for scholars and market practitioners alike.

In summary, I am very impressed by Josef Bajzík's doctoral thesis. His thesis represents metaregression analysis at the highest level, which is reflected by his publications in top-ranked economics journals. It would without any doubt also be a defendable and excellent thesis at my home university (University of Augsburg, Applied Data Analysis Group). Therefore, I recommend the thesis for defense without substantial change.

Author's response: *Amended. Dr.-Ing. Jerome Geyer-Klingeberg did not ask for any changes or improvements.*

D.3 Report from referee 2

doc. Ing. Karel Brůna, Ph.D. (Faculty of Finance and Accounting, Prague University of Economics)

The presented text is a review of the dissertation thesis titled "Essays on International and Financial Economics," authored by Mgr. Josef Bajzík and supervised by Prof. PhDr. Tomá‡s Havránek, Ph.D. This thesis was prepared for the evaluation of the dissertation's quality for the defense within the Economics and Finance doctoral program at the Institute of Economic Studies, Faculty of Social Sciences, Charles University.

The dissertation comprises three articles published in prestigious international journals indexed in the Web of Science with high quartiles (two articles indexed in Q1(D2) and one in Q2 in the respective year according to the AIS factor). Mgr. Josef Bajzík is the sole author of one of the articles in the collection and a co-author of two articles included in the dissertation. Each article constitutes a main chapter of the dissertation, accompanied by a unifying introduction, a concise conclusion, and a literature review. Unlike the usual focus on a specific set of interconnected economic issues, the unifying theme of the dissertation is the application of empirical analysis methodology in the form of meta-analysis. This allows the examination and evaluation of the properties of empirical estimates conducted by other authors through existing studies on a common theme.

Thematically, meta-analysis in the dissertation addresses three distinct economic problems, which, at first glance, may seem challenging to directly connect. The first article, titled "Estimating the Armington elasticity: The importance of study design and publication bias" (published in the Journal of International Economics in 2020, co-authors T. Havránek, Z. Irsová, J. Schwarz), explores the elasticity of demand for domestic and foreign goods (Armington elasticity). The second article, "Bank Capital, Lending, and Regulation: A Meta-analysis" (published in the Journal of Economic Surveys in 2023, co-authors S. Malovaná, M. Hodula, and Z. Gric), focuses on the impact estimates of increasing bank capital or capital requirements on the volume of bank loans. The third article, exclusively authored by the doctoral candidate, on "Trading volume and stock returns: A meta-analysis" (published in the International Review of Financial Analysis in 2021), examines the frequently studied problem of the influence of trading activity on stock returns. Each article faces the challenge of publication bias and investigates various data and methodological characteristics influencing the reported results.

I consider Mr. Josef Bajzík's dissertation to be of exceptionally high quality and competitive at numerous reputable scientific institutions, as evidenced by its composition of three articles published in highly ranked international scientific journals indexed in WoS in high quartiles. The work is meticulously crafted, with due attention given to each step of the research, and all procedures are adequately explained and referenced. The dissertation's text is logical and highly comprehensible, with no apparent weaknesses in the articles. While the exclusive focus on the empirical side of analyzed problems is a limitation, the work successfully incorporates theoretical discussions into the context of these problems.

I appreciate the extensive coverage of relevant empirical studies, as the author utilizes a dataset comprising over 3000 published estimates found in articles and working papers. The use of studies published in languages other than English is also noteworthy. The analysis of heterogeneity in results is supported by more than 40 different criteria. Processing such extensive data is labor-intensive, and I find it appropriate to positively evaluate this aspect of the author's work.

The presented work is valuable and contributory, especially in the comprehensive approach of applying meta-analysis methodology to selected problems. Mr. Josef Bajzík demonstrates a perfect understanding of this methodological group throughout the dissertation, utilizing a wide range of sub-approaches to calculating parameter mean values, estimating publication bias size, and correcting publication bias. The detailed analyses are particularly noteworthy for explaining the heterogeneity of parameter estimates based on various data characteristics, estimation methods, endogeneity omission, and the type of model used.

Given that the articles presented in the dissertation have undergone rigorous peer review in high-quality scientific journals, it is unnecessary to focus on minor details that may occur to the reviewer during their reading. The author no longer has the opportunity to incorporate these comments into the articles, and as mentioned earlier, the quality of the work is unquestionable. I would only mention one specific comment on the second article, where the results indicating the heterogeneous impact of capital or capital requirements on provided bank loans (e.g., higher sensitivity of corporate loans) could be significantly influenced not only by banks (loan supply) but also by the behavior of loan demand. Unfortunately, the authors did not find room to test factors influencing the heterogeneity of parameter estimates in this regard. In other words, the impact of capital on the volume of provided loans may be significantly strengthened or weakened by the extent to which, simultaneously with changes in loan supply, the interest of bank clients in a bank loan changes concerning the current phase of the economic (or financial) cycle (see, e.g., the countercyclicality of the economic cycle as a factor influencing the demand (especially of non-financial firms) for a loan and the cycle of risk aversion of banks influencing the loan supply cycle (see the theory of the financial accelerator)).

Before the comprehensive defense of the dissertation, it would be advisable to refine the conclusion of the work. Here, the author could elaborate more on the broader significance of the findings made within the analyzed problems. The time interval since the publication of the articles could provide the author with greater perspective, allowing him to express his view on the significance of his own research, which is undeniable, and share his opinion on the usability of the results in the broader context of the current economy. Simultaneously, it would be possible to expand on the context from which the author's interest in the given issue arose and highlight key theoretical or empirical publications that inspired the author to work on these topics. It would also be possible to anchor the analyzed problems in a broader overview of theories related to these problems. I would also recommend conducting another formal check of the work (e.g., the repeated paragraph in the introduction on page 6 or the missing figure on page 38).

Based on the above, I can unequivocally recommend Mr. Josef Bajzík's dissertation for defense after incorporating my comments on the conclusion and, if necessary, the introduction to the dissertation.

Author's response: *Amended. The discussion about loan demand is newly accomodated in the newly written conclusion part. I tried to capture the supply and demand sides of the respective regions and periods as close as possible. Thus, I together with mz co-authors include in the analysis external factors such as spread, housing price growth, monetary policy rate or unemployment. I believe we captured the surroundings as closely as possible, but there is still the question of whether we cannot capture the demand side in a better way. The discussion of broader significance of the findings and the broader context of the interconnectedness of the articles with the economy and economic research are also described in the concluding part of the thesis. Further formal check was conducted also and redundant paragraph was deleted.*

D.4 Report from referee 3

Ing. Mária Širáňová, MA. PhD. (Institute of Ecnomocic Research SAS, Slovakia

a) Can you recognize an original contribution of the author?

b) Is the thesis based on relevant references?

c) Is the thesis defendable at your home institution or another respected institution where you gave lectures?

d) Do the results of the thesis allow their publication in a respected economic journal?

e) Are there any additional major comments on what should be improved?

f) What is your overall assessment of the thesis? (a) I recommend the thesis for defense without substantial changes, (b) the thesis can be defended after revision indicated in my comments, (c) not-defendable in this form.

a) **Can you recognize an original contribution of the author?**
All three essays in the Dissertation thesis make use of the latest advances in the meta-analysis approach, which addresses a recent replication crisis in the field of economics and finance. From this perspective, all three essays bring significant new knowledge to the field, both in terms of the methodological approach of meta-analysis, and in terms of thematic topics (e.g. the need to calibrate more precisely the values of the Armington elasticity in theoretical models).

b) **Is the thesis based on relevant references?**
Indeed, all of the references are up-to-date and published in respected journals.

c) **Is the thesis defendable at your home institution or another respected institution where you gave lectures?**
This thesis can easily be defended at any economics department in Slovakia or at the Institute of Economic Research of the Slovak Academy of Sciences.

d) **Do the results of the thesis allow their publication in a respected economic journal?**
All three essays are already published, in the top field journals (first in Journal of International Economics, the second in Journal of Economic Surveys, the third in International Review of Financial Analysis).

e) **Are there any additional major comments on what should be improved?**
I don't have any major comments, just a few open questions that might be the subject of further discussion during the thesis defense. In my comment I focus specifically on open questions related to the use of meta-analysis as an analytical tool, which was applied in all three essays.

1. inclusion/focus on author characteristics
The meta-study by Fabo et al. (2021, 2024) argues that author's affiliation can be a source of bias in reported estimates. Other publications include a set of author's characteristics, such as information on the most cited author or an author's origin (e.g., Fidrmuc and Daniskova, 2019). Looking closely at the list of studies included in all three essays, apart from the characteristics mentioned above, I

think two additional avenues for research could be considered:
Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. I studied recent seminal articles about the central bank bias Fabo et al. (2021). As two of my articles were already published during such a discussion, we try to include at least to our macroprudential policy paper the information, whether the original working paper was central bank one or not (as substitute for the authors' origin), but we do not find it significant in our study. I believe that it is sufficient substitute to the set of author's characteristics.*

a) one author publishes several papers in a very short time window Observations reported by the same author for studies conducted in a similar context over a very short period of time are more likely to be characterised by the use of the same method, driven by the author's individual subjective preferences (i.e. unobserved author fixed effects). For example, see Essay 1 and publications by Lundmark & Shahrammehr (2011a, 2011b, 2012) for estimates of the Armington elasticity for forest biomass, roundwood, etc. These cases may or may not be labelled as a form of replication study, depending on the context. From this reason, one could argue for the inclusion of an author fixed effect, or, if there are enough instances of the same author effect in a sample (e.g., in Essay 1 - Welsch 2006 and 2008), a subjective publication strategy of authors in general could be further investigated.
Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. I found out, that some authors published several articles within short time period Lundmark & Shahrammehr (2011a, 2011b, 2012). Of course, the objective of their studies differ a bit (e.g. Armington elasticity for roundwood or biomass), but still it was pointed out that there is a big danger, that the attitude and the study design would remain very similar. I understand the point, but on the contrary, the question for me was whether it is right to include them all, so their approaches might receive higher weight incorrectly, or whether should I discard some of them; and thus some discretionary bias would appear. I ultimately decided to include all papers, as the meta-analytic paradigm is to work with*

"all relevant articles". To mitigate the issue I rely on a dilution prior in the Bayesian model averaging to give less weight to such potentially collinear studies (as Bayesian statistics does not use fixed effects).

a) both the working paper and the subsequent published article are included in the final sample.
As an example, see Essay 2 and publications by De Jonghe et al. (2016 and 2020) or Carlson et al. (2011 and 2013). While this may not be a problem per se, as published papers are coded as a separate category in the list of covariates, we see that the reported estimates in a journal version are strongly conditioned by the working paper version (see the almost negligible change in variation in Figure 3.1 for both examples). From a more general perspective, one might therefore ask whether these publications should be treated in the same way as other working papers in a dataset. Furthermore, by exploiting this feature and focusing on the differences between two versions of the same paper, one could potentially discuss the value of a review process itself (is there a substantial improvement due to the review process per se), selection bias by authors (i.e. authors of working papers are already distinguished editors/researchers), among others.
Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. Doucouliagos & Stanley (2013) suggest that results of meta-analyses working with published articles only do not statistically differ from those including both journals and working papers. Nevertheless, we tried to mine this issue for data in the second article, where we differentiated whether the working paper version of a later published article differed in the model used, the number of countries or time span, or in neither of those. Then we included all original estimates regardless of whether the journal version was included or not. Finally, we found no statistical difference, allowing us to conclude that in the area of banks, capital, and lending the results are not biased by editor or referee preference for a particular methodology or country and period framework. On the contrary, I believe that similar research should be included and expanded in other meta-analyses as well.*

2. endogeneity concerns in original papers
   In Essay 3, a dummy variable 'Lagged by 1Y or more' is listed among the covariates, interpreted as a dynamic estimate of a CAR effect (Table B1, 'the estimate is lagged by a year (4 quarters) or more'), and this interpretation is also supported in the discussion of the empirical findings on p. 84 ('We also find that studies which consider a short-term relationship â€¦ report lower semi-elasticity than studies considering a capital ratio lagged by one year of more. This may imply that changes to bank capital affect lending â€¦ more positively over a longer horizon').
   In my view, what captures this variable may be a more complex issue, as the use of a one-year lag in the underlying studies can be justified to address the endogeneity concerns stemming from the reverse causality between the change in credit and CAR (i.e. credit aggregates reported on the asset side and by being part of the denominator in any CAR measure). Interestingly, one of the main contributions of Essay 2 (the use of continuous CAR measures) is precisely what makes this issue of greater concern compared to the meta-analysis on the similar topic by Araujo et al. (2020) or Fidrmuc Lind (2020).
   A study included in the meta-sample (Roulet, 2018, p. 30) states: "bank-specific variables are lagged once (t-1) in order to mitigate possible endogeneity problems." To further complicate matters, the use of the Sys-GMM technique allows the flexibility to specifically control for potential endogeneity (Gambacorta and Shin, p. 27). One study included in the meta-sample (Gambacorta and Shin, 2018, p. 22, Table 2) lists all capital-related covariates as fully endogenous, in addition to other bank-specific determinants.
   In light of this discussion, one can consider the role of endogeneity concerns raised by the authors themselves in their original papers as a potential extension of the meta-analytic approach. The endogeneity concerns (more on the side of a problem of reverse causality) have been frequently discussed in the more recent papers, not least due to the more widespread use of advanced econometric techniques.
   Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. Endogeneity should not be issue in the meta-analytical estimation itself, but might appear in*

the primary articles (Gambacorta and Shin, 2018; Roulet, 2018). It is thus possible that the final data for meta-analysis included observations affected by endogeneity. To my knowledge, the best approach to deal with endogeneity in the meta-analysis is to capture the respective estimation methodologies from primary articles with dummies. This helps the researcher to discern whether methodologies that tackle endogeneity directly (e.g. GMM) provide systematically different estimates than those which do not. In this way, one can determine whether and to what extent endogeneity affects the primary estimates. An alternative way to deal with endogeneity would be to capture via dummy whether the primary study dealt with endogeneity or not. However, I find it redundant to include this treatment when I already control for different methodologies.

3. higher granularity in period coverage
   In line with meta-analysis standards, all three essays include variables describing the time dimension of a sample (Essay 1 - 'mid-year', 'data period', Essay 2 - 'the midpoint of the data', Essay 3 - 'mid-year', 'data period'). However, in Essay 2, the authors argue that the gradual change in regulatory requirements from the introduction of Basel I to the latest requirements of Basel III may affect the estimates in the underlying studies, assuming the adoption of different risk optimisation techniques (IRB approach) as well as banks' response to regulatory changes. In Essay 3, the motivational part of the study discusses the existence of a possible downward trend in reported estimates.
   Given these stories (and the possible presence of structural breaks in the underlying data samples), one can argue that a more granular specification of separate time periods in the data (e.g., 5-year windows, decades included in the sample) could provide a better fit to some of the (unanswered) hypotheses. In particular, a more granular time coverage (e.g. Basel II period from 2004/2008 to 2010/2015/2022) could uncover more detailed dynamics in Essay 2, as it only discusses the midpoint as a control variable (the same midpoint '2008' can conceal data from 1980 as well as starting only from 2004).
   Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. I am aware of the fact*

that each of my studies only uses the "midpoint and "length" of the primary data concerning the time span and length of the primary articles, suggesting the relationship of the data midyear is linear. In reality, one might suggest that, for instance, credit reactions under Basel I, II, and III might be different and hence, separate dummies for each decade might fit the data better. I considered this issue deeply and made the change in several of my studies, but with results not significantly different from those published in the thesis. Thus, I ultimately decided not to include this change in my articles.

Minor comments:

- Essay 1 compares estimates of short run and long run Armington elasticities:
  - is there not some sort of inherent multi-collinearity issue present in the list of covariates, assuming that the short run estimates are likely to be reported for the AR/ECM type of models - does the BMA alleviate this concern?

    Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. Our findings indicate that study characteristics are systematically associated with reported results. These characteristics were examined only on long-run estimates of the sample, so there is no danger that the short-run and long-run estimates are correlated in this part of the analysis. Besides we use the dilution prior in the BMA alleviate the possible multi-collinearity problem greatly, as it gives less weight to more collinear variables.*
- Motivational Figure 4.2. discusses the historical downward trend in reported estimates associated with an increasing variance:
  - The problem with using such illustrative figures is that, at first sight, removing an outlier (2005, estimate below -0.25) changes the dynamics completely - hence the use of winsorisation,
  - However, this raises the issue of 'significant outliers' - see the publication by Fabo et al. (2021), the follow-up by Weale and Wieladek (2022) and the consequent response by Fabo et al. (2024) and the removal of the most 'respected' estimates,

Author's response: *Amended. The discussion is newly accommodated in the newly written conclusion part. I have studied Fabo et al. (2021, 2024), Weale and Wieladek (2022) articles regarding their discussion about the inclusion of the outliers into the meta-analysis. Their discussion culminated after all my research included in the dissertation was published, so I did not incoporate their findings into my thesis. With respect to the proposed research question, I used meta-analytical approaches that represented the state of the art at the time for this type of application. For example, I used winsorization instead of trimming. Next, whenever I referred to a study for illustration, I used medians (e.g. per study, or per year) instead of means, to avoid possible suspicion of excessive visual manipulation of the data after winsorization. Still, from my understanding of the discussion between Fabo et al. (2021, 2024), Weale and Wieladek (2022) I agree that the topic of the outliers should be given more attention in future meta-analyses.*

- A careful reading would be beneficial, for example there are few formal errors (which is understandable given the size):
  - meta-analysis (p. 2), MalovanĂˇ & Frait 2017 (p.XXI)
    Author's response: *Amended. I went through the whole thesis and I repaired all the incriminated words.*

Literature

Araujo, J., M. Patnam, A. Popescu, F. Valencia, & W. Yao (2020): "Effects of Macroprudential Policy: Evidence from over 6,000 Estimates." IMF working paper wp/20/67, International Monetary Fund.

Fabo, B., Jančoková, M., Kempf, E., & Pástor, Ľ. (2021): Fifty shades of QE: Comparing findings of central bankers and academics, Journal of Monetary Economics, Volume 120, p. 1-20.

Fabo, B., Jančoková, M., Kempf, E., & Pástor, Ľ. (2024): Fifty shades of QE: Robust evidence, Journal of Banking & Finance, Volume 159, February, 107065.

Fidrmuc, J., & Danišková, K. (2019): Meta-Analysis of the New Keynesian Phillips Curve in Developed and Emerging Economies, Emerging Markets Finance and Trade, 1-22.

Fidrmuc, J., & R. Lind (2020): Macroeconomic Impact of Basel III: Evidence from a Meta-Analysis. Journal of Banking & Finance, 112: p. 105359.

Gambacorta, L., & H. S. Shin (2018): Why Bank Capital Matters for Monetary Policy. Journal of Financial Intermediation 35: pp. 17-29.

Roulet, C. (2018): Basel III: Effects of Capital and Liquidity Regulations on European Bank Lending. Journal of Economics and Business, 95, p. 26-46.

Weale, M. & T Wieladek (2022): DP17700 Fifty Shades of QE Revisited, CEPR Discussion Paper No. 17700. CEPR Press, Paris & London.

f) **What is your overall assessment of the thesis? (a) I recommend the thesis for defense without substantial changes, (b) the thesis can be defended after revision indicated in my comments, (c) not-defendable in this form.**

All comments that I have raised are just in a form of the suggestions, perhaps suitable for a further discussion, but it is not essential to implement them. My overall assessment therefore is:

"I recommend the thesis for defense without substantial changes."