Working Paper Series —— 8/2022

Borrower-Based Macroprudential Measures and Credit Growth: How Biased is the Existing Literature?

Simona Malovaná, Martin Hodula, Zuzana Gric, Josef Bajzík

Czech National Bank —— Working Paper Series —— 8/2022

CZECH NATIONAL BANK

www.cnb.cz

Reviewed by:	Ján Klacso (National Bank of Slovakia)
	Jiří Gregor (Czech National Bank)
Project Coordinator:	Dominika Ehrenbergerová

Borrower-Based Macroprudential Measures and Credit Growth: How Biased is the Existing Literature?

Simona Malovaná, Martin Hodula, Zuzana Gric, and Josef Bajzík *

Abstract

The ever-increasing use of borrower-based measures such as loan-to-value, debt-to-income, and debt service-to-income limits has created a demand to better understand the transmission and effectiveness of such policy. In this paper, we collect more than 700 estimates from 34 studies on the effect of borrower-based measures on bank loan provision. A birds-eye view of our dataset points to significant fragmentation of the literature in terms of the estimated coefficients. On average, the introduction or tightening of borrower-based measures reduces annual credit growth by 1.6 pp. Using a battery of empirical tests, we verify the presence of a strong publication bias, especially against positive and statistically non-significant estimates. The bias-corrected coefficient is about half the size of the uncorrected mean of the collected estimates but remains safely negative. Further, we explore the context in which researchers obtain such estimates and we find that differences in the literature are best explained by model specification, estimation method, and underlying data characteristics.

Abstrakt

Stále se rozšiřující využívání úvěrových ukazatelů, jako jsou limity pro poměr výše úvěru k hodnotě zajištění, poměr výše dluhu k příjmům a poměr dluhové služby k příjmům, vytvořilo poptávku po lepším pochopení transmise a efektivnosti této politiky. V tomto článku shromažďujeme více než 700 odhadů z 34 studií o vlivu úvěrových ukazatelů na poskytování bankovních úvěrů. Celkový pohled na náš datový soubor ukazuje významnou fragmentaci odborné literatury, pokud jde o odhadované koeficienty. Zavedení nebo zpřísnění úvěrových ukazatelů v průměru snižuje meziroční růst úvěrů o 1,6 p. b. S využitím řady empirických testů potvrzujeme značnou publikační selektivitu, a to zejména vůči kladným a statisticky nevýznamným odhadům. Koeficient očištěný o publikační selektivitu činí zhruba polovinu neočištěného průměru shromážděných odhadů, ale zůstává bezpečně v záporných hodnotách. Dále zkoumáme kontext, v němž výzkumníci své odhady získávají, a zjišťujeme, že rozdíly v literatuře lze nejlépe vysvětlit specifikací modelu, metodou odhadu a charakteristikami použitých dat.

JEL Codes: C83, E58, G21, G28, G51.

Keywords: Bayesian model averaging, borrower-based measures, macroprudential policy, meta-analysis, publication bias.

* Simona Malovaná, Czech National Bank, simona.malovana@cnb.cz

Martin Hodula, Czech National Bank and Technical University of Ostrava, martin.hodula@cnb.cz

Zuzana Gric, Czech National Bank and Masaryk University in Brno, zuzana.gric@cnb.cz

Josef Bajzík, Czech National Bank and Charles University in Prague, josef.bajzik@cnb.cz

The authors note that the paper represents their own views and not necessarily those of the Czech National Bank. We would like to thank Ján Klacso, Jiří Gregor, Dominika Ehrenbergerová, and participants at the Czech National Bank's seminar for useful comments. All errors and omissions remain the fault of the authors.

1. Introduction

The use of macroprudential policy measures has increased significantly since the Global Financial Crisis of 2008–2009, with borrower-based measures taking their place among the most popular policy tools in both advanced and emerging market economies (Cerutti et al., 2017a; Alam et al., 2019). The IMF iMaPP database covers 134 countries and records 87 borrower-based policy actions over the 1990–2007 period and 316 over the 2008–2020 period.[1] Borrower-based measures, such as loan-to-value, debt-to-income, and debt service-to-income limits, are applied at the level of individual loans and can therefore directly restrict the amount of credit available to the private sector. The limits are meant to increase borrowers' and lenders' resilience and lower potential bank losses during economic and financial downturns.

Despite the growing use of borrower-based measures, there is a limited consensus on how well the toolkit works in practice. In fact, the empirical literature has so far not fully succeeded in rigorously quantifying the effects of various borrower-based measures on bank lending and has displayed significant fragmentation in terms of the estimated coefficients (Figure 1). The focus on the effect of borrower-based (loan-targeted) measures on bank lending is natural given that such regulation directly targets bank credit. In fact, domestic bank credit growth is often seen as the intermediate target of macroprudential policy because of its direct link to boom-bust financial cycles. Numerous studies find that credit booms typically precede crises (Mendoza and Terrones, 2008; Jordà et al., 2011; Schularick and Taylor, 2012).[2] Placing bank credit at the center of policy and academic analyses is further supported by the rich evolving literature on the interaction of monetary and macroprudential policy. While the policies have different aims, they both have direct or indirect impacts on bank credit (Galati and Moessner, 2018), impacts which can, in certain situations, be at odds (Malovaná and Frait, 2017).

The fragmentation is due in part to a reliance on dummy-type policy action indices, which precludes estimation of the quantitative effects of the policy. The indices assign a value of one when a macroprudential policy tool is implemented or changed and zero otherwise. In addition, one might suspect that significant endogeneity problems hamper a proper assessment of macroprudential policy effects. Endogeneity bias may stem from the fact that macroprudential measures are usually taken in response to developments in credit and asset prices, making it tricky to properly identify policy shocks and estimate their effects on the economy. Furthermore, there is limited evidence on the effects of using borrower-based measures individually and in combination with one another. For instance, there is an emerging literature that provides evidence of different impacts of imposing *value-based* loan-to-value (LTV) limits compared to *income-based* debt-to-income (DTI) or debt-service-to-income (DSTI) limits (Grodecka, 2020; Hodula et al., 2021).

In this paper, we assign a pattern to the observed differences in the estimated effect of imposing borrower-based measures on bank loan provision. To this end, we collect more than 700 estimates of the elasticity from 34 articles. To explain the differences, we collect an additional 24 variables that reflect the context in which the estimates were produced. With this newly created database,

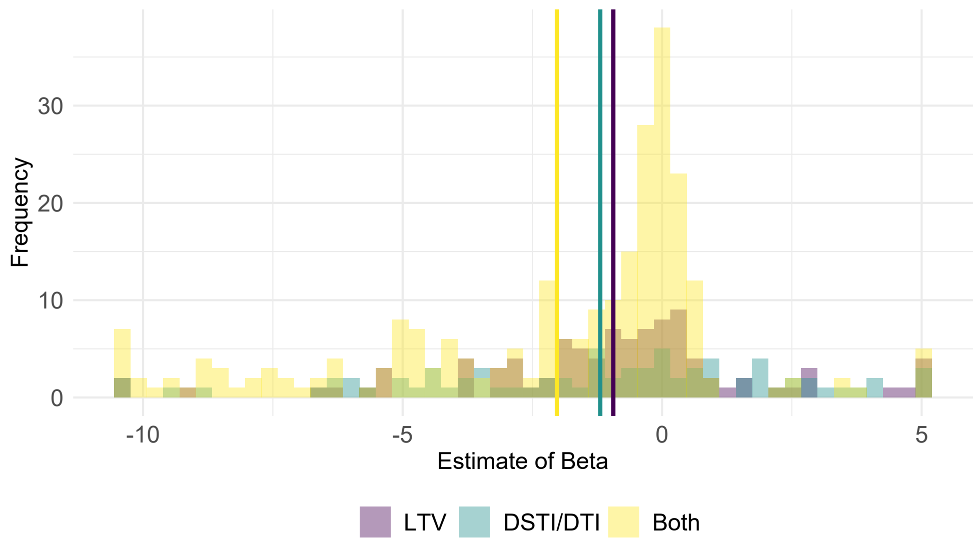
[1] The capital requirements for banks, which include risk weights, systemic risk buffers, and minimum capital requirements, saw a similar increase in their frequency of use but were used far less often (51 during the 1990–2007 period vs. 257 between 2008 and 2020). The countercyclical capital buffer was used sporadically (1 vs. 68).

[2] See also Fuster et al. (2010), Geanakoplos (2010), Boz and Mendoza (2014), Gennaioli et al. (2015), Boissay et al. (2016), and Bordalo et al. (2018) for stylized models that try to capture the boom-bust credit cycle associated with financial crises.

we first derive the "average" effect of LTV, DTI, and DSTI limits on bank lending when used individually and in combination with one another. Since borrower-based measures are becoming a standard policy tool for addressing imbalances in the residential mortgage loan market, the estimated average (general) effect of individual tools could be of value to policy practitioners when deciding on regulatory policy placement. However, beyond the average effects, policymakers, lenders, and the public at large are keen to better understand the heterogeneity in the effects. Therefore, we next explain why the estimates from the literature vary across different studies and describe what the most commonly employed empirical strategy is. We then use state-of-the-art meta-analytical techniques to estimate the true effect of imposing borrower-based measures on bank lending, as well as model-averaging methods to identify significant drivers of the heterogeneity in the observed estimates.

The meta-analysis employed in the paper is well suited to tackling the heterogeneity of the estimates surveyed. If the subject of a meta-analysis is a deep/structural parameter in a correct model (i.e., a model that correctly describes the data-generating process), then the heterogeneity of the estimates should be fairly low and given by econometric issues only (estimation method, variable transformation, etc.). If the subject of a meta-analysis is a reduced-form parameter of a correct model, then higher heterogeneity of the estimates should be expected, because different policy regimes imply different values of the parameter. So, the heterogeneity of the estimates surveyed should also be explained by data characteristics (time period of data set, country, etc.). Finally, if the model does not correctly follow the data-generating process, which is exactly the case for modeling the impact of borrower-based measures on bank credit, the heterogeneity of the estimates will be even greater and closely related to the model specification. All the above-mentioned reasons for heterogeneity apply to the exercises presented in this paper.

Figure 1: The Mean Effect of Borrower-Based Measures on Annual Credit Growth is Negative



Note: The figure shows a histogram of the estimated beta for all the collected estimates on the relationship between bank annual credit growth and borrower-based measures expressed as an index or a dummy variable (see equation 1). The beta is the percentage point change in annual credit growth in response to a one-unit increase in the borrower-based index or dummy. LTV stands for loan-to-value ratio, DSTI/DTI for debt service-to-income/debt-to-income ratio, and both for measures encompassing multiple borrower-based measures. The solid vertical lines indicate the means of the winsorized estimates; the collected estimates are winsorized at 2.5% from each side.

In taking a panoramic view of the collected estimates, we expose three stylized facts. First, the average estimated effect of borrower-based measures on annual bank credit growth is negative, with a mean value of -1.6 pp, but can vary widely, as the estimates range from -8.8 pp to 3.1 pp at the 5% level of significance. Since most studies work with dummy-type indices to capture regulatory change, the identified effect refers only to the introduction of, or changes to, the measure, not to its intensity or binding nature. Second, the mean effect of imposing a value-based limit (LTV) individually and in combination with income-based (DTI, DSTI) limits on bank lending is observably different in terms of the size of the reported coefficient. The reported estimates of the joint impact of LTV, DTI, and DSTI limits are found to be more than two times higher than those of applying just an LTV limit. We find that DTI and DSTI limits reduce credit growth more than an LTV limit does. Third, more recent studies report more negative estimates of the impact of borrower-based measures on bank lending, but it is not clear whether the observed trend reflects change in the strictness of regulation or improved data and techniques used by more recent studies.

Meta-analysis also allows us to correct for the presence of publication bias, another potential problem associated with the average estimated effect. Publication bias arises when researchers do not publish all their estimates, but only those that are significant or have the "correct" sign (Stanley, 2005). Left unaddressed, such selectivity can lead to biased estimates and misleading confidence sets in published studies (Andrews and Petroulakis, 2019). In the context of the effect of borrower-based measures on bank lending, it is safe to assume that the expected sign is negative. Given numerous possibilities in both study design and the choice of proper estimation approach, one can always "try harder" to find significant or "correctly signed" estimates (Card and Krueger, 1995; Stanley et al., 2013). To test and correct for publication bias, we use a battery of state-of-the-art statistical tests. We find that researchers strongly prefer negative estimates and discard positive ones. The bias-corrected coefficient is almost half the size of the uncorrected mean of the collected estimates. This shows that the literature generally overestimates the negative effect of borrower-based measures on bank lending. The documented publication bias is found to be driven by elasticities that have the "correct" negative sign and are "just" significant at the 10% level. In other words, researchers over-report negative significant estimates and under-report positive significant ones. Interestingly, we also identify over-reporting of positive non-significant estimates, suggesting that researchers are "just fine" about reporting a positive effect as long as it is not statistically significant.

Of course, the mean effect may conceal important differences in the context in which the coefficients are obtained. Using Bayesian and frequentist model-averaging methods, we indeed show that certain study characteristics are systematically associated with the results reported. The most important are those related to model specification, estimation methods, and data characteristics. Among other things, our results point to the presence of a strong endogeneity bias related to the estimation of the effects of borrower-based measures on bank lending. Endogeneity of credit and macroprudential policy biases the estimates toward zero. We find that studies that took measures to ensure exogeneity of the policy shock report more negative estimates. This may be linked to another finding that studies using confidential data also report a more negative effect. Confidential data usually capture developments at the level of a single entity or product (e.g., a bank, a firm, or a loan) and studies using such data may in theory be more successful than others in correctly identifying exogenous regulatory shocks. We further find that contemporaneously specified models deliver lower negative elasticity estimates than models specified with a lag, suggesting that the impact of the measures takes time to materialize (at least in the data).

This paper contributes to the rich empirical literature on the effects of macroprudential policy measures. The most widely cited papers in the relevant literature build on cross-country data and capture the occurrence of macroprudential policy measures (including borrower-based ones) using dummy-type indices (Lim et al., 2011; Kuttner and Shim, 2016; Cerutti et al., 2016; Akinci and Olmstead-Rumsey, 2018). Studies typically find that the application of macroprudential measures lowers the real credit growth rate and slows down house price growth. While the dummy approach allows one to compare the effects of regulation across countries, it falls short on capturing the intensity of a measure. For instance, reducing the LTV limit from 100% to 80% is treated in the same way as lowering it from 80% to 70%. This may cast doubt on the ability of the literature to correctly estimate the effects of macroprudential policy. However, even intensity-adjusted indices are not without flaws, as they are unavoidably subjective and fail to capture the binding nature of the policy shock. This is where a synthesis of the literature, such as ours, provides an important service to the field – by putting together estimates from different studies, meta-analysis can estimate the average effect of regulation as well as examine the systematic dependencies of the reported results on study design and filter out the effects of misspecifications. It thus informs policymakers and the wider economic community about the effects of macroprudential policy measures. The cross-country studies are complemented by a second group of papers using micro-level evidence, mostly based on the use of one or a few policy tools in a single country. While the micro-founded evidence offers more precise estimates of the effects, it provides a single-country perspective at the very best (Ahuja and Nabar, 2011; Igan and Kang, 2011; Acharya et al., 2020b). The meta-analytical summary can serve as a benchmark against which country-specific studies can assess their estimates.

There are several other studies that sum up the effects of macroprudential policy measures on bank credit. Early examples of such studies include Galati and Moessner (2013) and Galati and Moessner (2018), who conduct a narrative review of the literature. Gambacorta and Murcia (2020) performs a meta-analysis focused on estimates from seven Latin American countries. Araujo et al. (2020) summarizes estimates of the effects of macroprudential policy on bank credit, house prices, and the real economy. Malovaná et al. (2021) perform a meta-analysis of the literature estimating the effects of continuous changes to (regulatory) bank capital on bank lending. Among other things, they highlight that the reported elasticity estimates are significantly affected by the researchers' choice of empirical approach.

The remainder of this paper is structured as follows. Section 2 describes how we collect the data from the primary studies. Section 3 tests for publication bias and estimates the effect beyond bias. Section 4 explores the heterogeneity of the estimated elasticities, and Section 5 concludes.

2. Collection Process and Formation of the Dataset

The focus of this paper is on the estimated effect of borrower-based measures on bank lending. Borrower-based policy tools can be grouped into three categories. The first category contains the loan-to-value (LTV) limit, a cap that directly restricts the amount of a loan to a certain fraction of the total value of the residential real estate. The second category includes income-based limits such as those on the debt-to-income (DTI) and the debt-service-to-income (DSTI) ratios. Here, the aim is to prevent the provision of subprime loans to borrowers whose total debt or debt service payments relative to their income exceed a certain limit set by the regulator. The third category consists of other borrower-based measures, which are usually country-specific. These include, but are not limited to, taxes on real estate gains, direct quantitative limits on mortgage lending, and changes in regulatory risk weights.

To estimate the effect in question, researchers regress credit on various expressions of macroprudential borrower-based measures. In our selection procedure, we therefore consider all empirical studies with some form of borrower-based measures on the right-hand side and bank credit on the left-hand side of the regression (equation 1). A dominant proportion of studies (31 out of 34) focus on modeling the impact of borrower-based measures on the stock of credit. This somewhat contrasts with the fact that borrower-based measures are applied to new business, i.e., on flows, while authors study the impact on the stock of loans. One can therefore expect the measures to have a more gradual impact on the change in the stock of loans. About 60% of the estimates collected are based on the same variable transformation – credit growth and a borrower-based index or dummy (see Table 2). Having a single prevailing variable transformation allows us to provide an empirical summary of the effect of changes to borrower-based limits on bank lending, which provides a significant benefit over a unitless standardized effect approach (Doucouliagos and Laroche, 2003; Ahmadov, 2014; Valickova et al., 2015; Bajzik, 2021).[3] Our findings will let us draw more convincing conclusions not only on the true direction of the analyzed effect, but also on the determinants of the observable heterogeneity in the estimates.

The estimated elasticities $\hat{\beta}$ entering the analysis in sections 3–4 refer to the following equation:

$$\%\Delta L_{it} = \hat{\beta} BBM_{it} + \gamma X_{it} + \varepsilon_{it} \tag{1}$$

where $\%\Delta L_{it}$ is annual credit growth, BBM_{it} is a borrower-based measure (dummy or index), and X_{it} is a vector of control variables for time t (year, quarter, month) and unit i (country or bank). A dummy variable takes the discrete value of 1 if a borrower-based measure is active and 0 if it is inactive. An index is a dummy-type indicator which is calculated as the sum of the active borrower-based measures across the period analyzed. The literature uses two basic indices – a frequency index and a direction index. The frequency index, as the name suggests, adds up all the policy actions regardless of the direction of the borrower-based measures (whether the policy is tightening or loosening). The direction index, on the other hand, takes negative values (-1) for loosening actions and positive values (+1) for tightening actions. Given the similar construction of these three measures, we can interpret the estimated elasticities $\hat{\beta}$ as the percentage point change in annual credit growth in response to a one-unit increase in the borrower-based index or dummy.

In the absence of a "unified macroprudential policy function", the literature includes several variations of equation 1 involving different types and definitions of the credit variable or borrower-based measures, different sets of control variables, and different estimation approaches. For instance, our sample studies consider two different types of borrower-based measures – value-based (limits on the LTV ratio) and income-based (limits on the DSTI and DTI ratios). We use state-of-the-art meta-analytical techniques to construct summaries of the estimated elasticities, aiming to verify the presence of publication bias as well to explain why the estimates may vary.

2.1 Paper Selection Procedure

We started the data collection process by building a list of primary studies relevant to our topic. We employed Google Scholar as the main database for our search and screened the first 350 articles returned by the following query:

> LTV OR LTI OR DSTI OR DTI OR borrower-based OR loan-to-value OR loan-to-income OR debt-service-to-income OR debt-to-income AND lending OR credit OR loans

[3] Estimate heterogeneity in terms of variable transformation or different measurement can be handled by transforming the collected estimates to standardized effects, i.e., partial correlation coefficients. We explore this option and analyze the significance of these and other heterogeneity drivers later in the paper.

We limited our search to studies published in 2010 or later. This is to account for the fact that macroprudential measures have been used more extensively since the GFC, with borrower-based measures being the most popular tool in advanced economies (Alam et al., 2019). Furthermore, many empirical studies examining the effects of borrower-based measures have been published with a delay. Naturally, there is a lag between activating a policy and collecting a sufficient number of observations for a rigorous analysis of its effects. Limiting our search also allows us to better compare the results of this study with our sister study examining the impact of capital-based measures (Malovaná et al., 2021).

After reviewing the first set of articles, we continued with the "snowballing" method. This involved going through all the citations in each of the relevant studies and identifying an additional 139 articles for screening. We closed the identification process in November 2020. As a next step we assessed the eligibility of each article and discarded 389 of them based on the abstract and an additional 66 due to a lack of correspondence or data (see Figure A1). We ended up with 34 primary studies, from which we collected 722 point estimates.[4] In the process of data collection, we closely followed the guidelines established in Havránek et al. (2020).[5]

The size of our final sample was influenced by several other aspects. First, the set of primary studies consists of both journal articles and working papers. While the inclusion of working papers might be frowned upon, we need to keep in mind that working papers are one of the main communication tools used in the central banking industry. As such, they go through a rigorous peer-review process, much like journal articles (Malovaná et al., 2020). During the screening process, we searched for a working paper (journal) version of each journal article (working paper) and assessed whether and how these two versions differ (see Table 1). In the end, we employed only those estimates from journal and working paper versions which were unique. Second, we disregarded all the elasticities that were reported without standard errors, p-values, t-statistics or confidence intervals. Third, we harvested elasticities for both the borrower-based measure and its interaction with the dummy variable[6] where the specification employed one. We then treated those elasticities as separate observations in our data set and distinguished between the two using an indicator variable. We used a similar approach for specifications with additional lags of the borrower-based measure. That is, we collected all the elasticities for multiple lags and employed each of them in the analysis together with an indicator variable to examine the potential heterogeneity introduced by this setting. Last but not least, two studies in our sample report impulse-response functions (IRFs) instead of regression estimates. In these cases, we harvested the elasticities for the immediate, after-one-period, and maximum responses from each IRF. Again, we then used all three observations as "stand-alone" elasticities and differentiated between them in the model-averaging exercise (see Section 4) by employing corresponding dummies.[7]

[4] Ioannidis et al. (2017) examined the statistical significance and bias of estimates reported in 159 meta-analytical studies, which altogether employed more than 64,000 estimates of economic parameters. The mean sample size in the underlying meta-analyses was 402 observations.

[5] The data set was prepared by two of the authors of the present paper and then cross-checked by the other two in several rounds to limit inconsistencies and systematic mistakes.

[6] The distinction between the effect of a borrower-based measure and the effect of an interacted variable if it is continuous is not straightforward. As a result, we decided to filter such cases out of our data set.

[7] Fidrmuc and Korhonen (2006) used a similar framework.

Table 1: Journal Articles and Working Papers Included in the Meta-Analysis

	Journal articles		**Working papers**	**Do they differ?**
	-	1	Acharya et al. (2020a)	-
	-	2	Afanasieff et al. (2015)	-
	-	3	Ahuja and Nabar (2011)	-
1	Akinci and Olmstead-Rumsey (2018)	4	Akinci and Olmstead-Rumsey (2015)	Y (M)
	-	5	Alam et al. (2019)	-
	-	6	Ayyagari et al. (2017)	-
	-	7	Ayyagari et al. (2018)	-
2	Bachmann and Rüth (2020)	8	Bachmann and Rüth (2017)	Y (O)
	-	9	Budnik (2020)	-
3	Carreras et al. (2018)		-	-
4	Cerutti et al. (2017a)	10	Cerutti et al. (2016)	N
5	de Araujo et al. (2020)	11	de Araujo et al. (2017)	Y (M)
	-	12	Epure et al. (2018)	-
6	Gadatsch et al. (2018)	13	Gadatsch et al. (2017)	N
7	Geršl and Jašová (2014)	14	Gersl and Jasova (2012)	Y(O)
	-	15	Igan and Kang (2011)	-
	-	16	Jácome and Mitra (2015)	-
	-	17	Kronick (2015)	-
	-	18	Krznar and Morsink (2014)	-
8	Kuttner and Shim (2016)	19	Kuttner and Shim (2013)	Y (M,P)
	-	20	Lim et al. (2011)	-
	-	21	McDonald (2015)	-
9	Morgan et al. (2019)	22	Morgan et al. (2015)	Y (P,C)
	-	23	Neagu et al. (2015)	-
10	Poghosyan (2020)	24	Poghosyan (2019)	N
	-	25	Richter et al. (2018)	-
	-	26	Tantasith et al. (2020)	-
	-	27	Wang and Sun (2013)	-
11	Zhang and Zoli (2016b)	28	Zhang and Zoli (2016a)	N
12	Zhang and Tressel (2017)		-	-

Note: Y/N – journal version and working paper do/do not differ; M – journal version and working paper use different model or methodology; P – versions differ in time period examined; C – different number of countries is studied; O – journal and working paper version differ in the sense that only one version contains elasticities of interest. Estimates that differ between the journal article and the working paper enter the meta-analysis. If the estimates are the same, they enter the meta-analysis only once. Hence, the final set of studies comprises 11 journal articles (Bachmann and Rüth, 2020, is excluded, as only the working paper version contains elasticities of interest) and 23 working papers (28 working papers minus 4 that do not differ from the journal version, minus 1 that does not contain elasticities of interest).

In the next step, we adjusted the collected estimates to prepare them for the analysis and to achieve overall comparability. First, we manually computed standard errors for all the observations for which they were not reported, using information on t-statistics, p-values or confidence intervals. Second, we unified all the elasticities that capture the effect on credit growth to yearly effects (for example, by multiplying non-annualized quarterly estimates by four). Third, we converted the cumulative elasticities to average one-period estimates by dividing them by the equivalent number of periods. Finally, we transformed all the elasticities to percentage-point effects. We harvested all the necessary information for this step during the data collection process. The same modifications were applied to the corresponding standard errors.

As discussed above, when selecting the articles and collecting the data, we focused on estimates that capture the effect of some borrower-based measure on bank lending, regardless of the transformation of the variables. This is because we do not have any prior knowledge of the most common model specifications for the relationship studied. The first part of Table 2 provides information on the most frequent combinations of left-hand and right-hand side variables in our initial data set. 59% of all the estimates concern the effect of a one-unit change in the borrower-based index or dummy on annual credit growth. All the other pairs of variable transformations are rather minor, with the runner-up combination accounting for just over 12% of the estimates. All in all, the majority of primary studies worked with the same typology of borrower-based measure – either a borrower-based index or a dummy (BBM index/dummy). We consider these two classes of variables in one category, because the effect of both when employed in the estimation is measured as the effect of a one-unit change.

Table 2: Partial Correlation Coefficients for Different Variable Transformations and Measures

	Obs.	Articles	Mean	5%	95%	Skewness
Total	722	34	-0.003	-0.131	0.124	0.050
Credit growth $\sim \beta \times$ BBM index/dummy	422	23	-0.013	-0.110	0.080	0.430
Log credit $\sim \beta \times$ BBM index/dummy	91	7	-0.027	-0.173	0.123	0.120
Other transformations and measures	209	7	0.026	-0.109	0.140	-0.540

Note: The table presents summary statistics for all the collected estimates converted to partial correlation coefficients (PCCs), which is a standardization method commonly used in meta-analyses (Doucouliagos, 2005; Havranek and Irsova, 2010; Havranek et al., 2016). PCCs allow one to compare estimates with different units of measurement which are not directly comparable. However, by making this transformation, we cannot draw any conclusions about the true size of the estimate, only about its direction. The PCC from the i^{th} estimate of the j^{th} study can be derived from the t-statistics of the reported estimates and the residual degrees of freedom: $PCC_{ij} = t_{ij}/\sqrt{(t_{ij}^2 - df_{ij})}$. BBM index/dummy refers to borrower-based measures reported in the form of a dummy variable or index. The category of other transformations and measures includes different combinations of variable transformation and different ways of capturing borrower-based actions (e.g., LTV value or change, distance from LTV limit or LTV tightening reported by banks in Bank Lending Survey).

Even though it is essential for the meta-analysis to maintain some degree of heterogeneity in the data, large differences in variable transformation make it impossible to compare the resulting elasticities. The partial correlation coefficient (PCC) method is able to address a part of this issue, because it uses the estimates' t-statistics and residual degrees of freedom to form partial correlation coefficients that are in fact comparable across different transformations. Nevertheless, it disregards information about the estimated elasticities themselves, hence the PCCs reflect only the direction, not the true size of the effect. The second part of Table 2 reports the results of the PCC method. The two most common transformations as well as the total sample consistently lean toward a negative correlation between borrower-based measures and credit. Although the other transformations exhibit a positive mean PCC, they are negatively skewed.

This exercise suggests that the effect of borrower-based measures on bank lending is largely negative. We consider this result to be vital, but we want to examine the relationship more thoroughly. To be able to assess the direction and also quantify the *size* of the effect, we focus only on the most frequent variable transformation – Credit growth $\sim \beta \times$ BBM index/dummy, which encompasses 424 observations drawn from 23 studies. In this case, instead of PCCs we can directly employ the collected estimates in the analysis. The elasticities corresponding to the selected transformation represent the percentage-point effect of a one-unit increase in the borrower-based index or dummy on annual credit growth.

2.2 Early View of the Collected Elasticities

An early view of the collected elasticities suggests five stylized facts. First, the average estimated effect of introducing or tightening borrower-based measures on bank lending is highly negative:

a unit increase in a borrower-based index or dummy is associated with a decrease in annual bank lending growth of about 1.6 pp (Table 3). Borrower-based measures can directly restrict the amount of credit available to the private sector. Thus, a negative sign is largely expected when such a measure is introduced or tightened. However, the mean value is surrounded by wide confidence intervals, ranging from -8.8 to 3.1 pp at the 5% level of significance. This points to a large degree of heterogeneity that warrants exploration, despite the clear negative sign on the relationship. The reported estimates vary both within and across studies (see Figure A2).

Table 3: Breakdown into Categories of Different Borrower-Based Measures

	Obs.	Articles	Mean	5%	95%	Skewness
Total	422	23	-1.63	-8.8	3.05	-0.75
Limit set by the authority						
LTV	95	15	-0.94	-5.83	4.54	-0.68
DSTI/DTI	76	12	-1.19	-7.21	4.01	-0.50
Both limits	251	12	-2.03	-8.97	0.76	-0.88
Measure used in the analysis						
Dummy	102	10	-0.89	-9.22	4.45	-0.95
Index (direction)	191	11	-1.07	-5.76	0.56	-1.47
Index (frequency)	129	8	-3.06	-9.72	3.64	0.03
Direction of the proposed limit						
Tightening	57	8	-0.71	-5.29	4.26	-0.03
Loosening	41	4	0.84	-4.74	6.42	0.57
Both directions	324	19	-1.89	-9.22	2.66	-0.80
Short-term vs long-term effect*						
Short-term effect	224	15	-1.15	-7.59	3.44	-0.90
Long-term effect	224	15	-1.53	-11.70	3.58	-1.51

Note: The table presents summary statistics for the narrow sample of estimates related to primary studies using either a dummy variable or an index to capture the effect of borrower-based measures in the estimation. If a "loosening" dummy or index takes on a positive value in the primary study, we multiply the collected estimate by -1 to make the direction of the effects comparable to the whole sample (i.e., the overall direction index assigns -1 to periods of loosening). In other words, we interpret loosening as a decrease in tightening (assuming a linear effect). This is not true if we explore the loosening actions separately, as for example, in the third part of this table. *We further narrowed down the sample to include only estimates obtained based on a dynamic model specification. In this case, we are able to calculate the long-term effect and compare it to the short-term effect. The long-term effect is calculated as the short-term effect β divided by 1 minus the autoregressive coefficient α in a dynamic model specification such as the following: $CreditGrowth_t = \alpha \cdot CreditGrowth_{t-1} + \beta \cdot BBM_t + \gamma \cdot X_t + \varepsilon_t$.

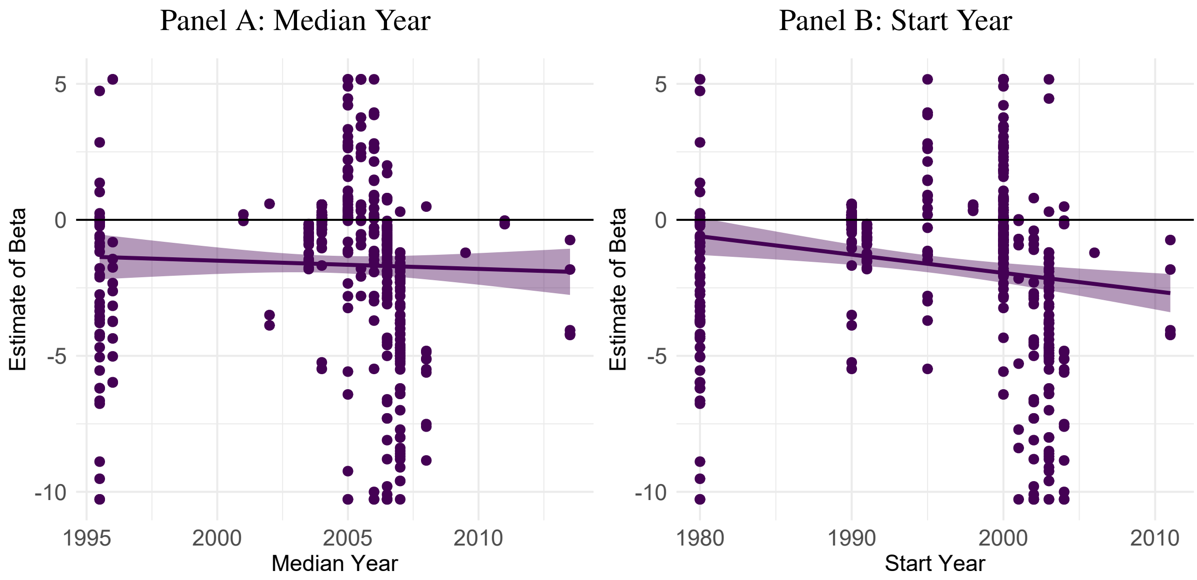
Second, the mean estimated effect of value-based (LTV) limits on bank lending is observably different from that of income-based (DTI, DSTI) limits in terms of magnitude. Specifically, the application of DTI and DSTI limits seems to be more effective in reducing the excess credit growth than the LTV limit. The mean effect of introducing and/or tightening the LTV cap comes in at -0.9 pp, while the mean for the application of DTI and DSTI limits is -1.2 pp. This is in line with Claessens et al. (2013), Cerutti et al. (2017a), and Kuttner and Shim (2016), who estimate the effect of the DSTI limit on credit growth to be negative while finding a much smaller effect for the LTV limit. Additionally, this result echoes the literature showing that while LTV limits can be successful in lowering demand for and the granting of larger loans (Richter et al., 2019; Armstrong et al., 2019), their de-risking effect can be partially offset by borrowers pledging properties of a higher value (Van Bekkum et al., 2019) or using various strategies for increasing the value of collateral accepted by banks (Hodula et al., 2021). The application of both groups of borrower-based limits (LTV and DSTI/DTI) together leads to the strongest decline in annual bank credit growth, averaging -2 pp. The "belt and braces" approach of imposing value-based LTV limits combined with income-based DTI and DSTI limits thus appears to have the biggest negative effect on bank lending. This is in line with the estimates of a structural model in Grodecka (2020) for the Swedish economy and the estimates from a reduced-form model in Hodula et al. (2021) for the Czech housing market. Evidence that LTV, DTI, and DSTI limits complement each other has also been found in the case of Slovakia (Jurča et al., 2020; Cesnak et al., 2021).

Third, researchers' choice of how to capture borrower-based policies in their model matters for the size of the estimated coefficient. Employing a simple binary dummy (1 for the occurrence of a policy tool and 0 otherwise) delivers a smaller effect than considering a dummy-based index that sums all the borrower-based actions during the period analyzed. While all dummy-based indicators fail to account for the intensity of the change in borrower-based measures, a binary dummy fails to account for the concurrence of more than one policy event. For this purpose, many studies use a cumulative policy index which sums policy actions over time and thus captures not only the immediate effect of the introduction of one policy tool, but also the overall tightness of the regulatory environment. Studies generally consider two types of cumulative index – one that marks with 1 the application of any policy measure (regardless of the tightening/easing nature of the measure) and thus captures the general use of macroprudential policies (frequency index) and another that assigns a value of 1 to tightening and -1 to easing actions and thus better captures the overall stringency of macroprudential policy (direction index). The mean effect in studies that consider the frequency index is -3.2 pp, while the mean for the direction index is -1.1 pp. Moreover, we find that the mean in studies considering tightening measures only (-0.71 pp) is fairly close to that in studies considering loosening measures only (0.84 pp), suggesting that borrower-based measures have a symmetric effect. This insight from our empirical summary of the literature is somewhat at odds with single studies, which generally arrive at rather conflicting conclusions regarding the symmetry. Using 99 episodes of changes to borrower-based limits, Poghosyan (2020) finds that the impact of loosening measures is stronger than that of tightening ones. McDonald (2015) reaches a different conclusion, as he shows that loosening measures have smaller effects than tightening ones, but the difference is negligible in downturns. In relation to housing credit, Kuttner and Shim (2016) find the effects of tightening to be significant, while those of loosening are not.

Fourth, the average estimated effect changes over time. Figure 2 shows that the median beta estimate depends not only on the type or combination of borrower-based tools, but also on the start or median year of the estimation sample. In general, more recent studies report more negative estimates of the impact of borrower-based measures on bank lending. There are three possible explanations for this. One, studies incorporating more recent data capture more macroprudential borrower-based actions, since their use has increased rapidly since the GFC (Cerutti et al., 2017a; Alam et al., 2019). Two, it may reflect the fact that the borrower-based measures that have been used since the GFC are more binding and/or that the regulatory shocks are getting bigger. Three, the larger negative effect of borrower-based measures on bank lending may be related to the large structural changes that have occurred in the post-GFC period. For instance, low interest rates have depressed bank profitability and may have made banks more exposed to regulatory shocks. A similar piece of evidence was obtained in a meta-analytical study of capital-based macroprudential measures (Malovaná et al., 2021). Similarly, continuously growing house prices may have increased the demand for, and the volume of, real estate mortgage loans, making them more responsive to the introduction of borrower-based limits.

Fifth, the effect of borrower-based measures on bank lending is somewhat larger in the long term (-1.53 pp) than in the short term (-1.15 pp), but the difference is not as significant as might be expected. This finding touches upon the ongoing debate on whether borrower-based measures should be used as a cyclical or structural macroprudential policy tool. If used as a structural tool, borrower-based measures are introduced and kept unchanged through the cycle, while as a cyclical tool, they are changed more frequently and usually released in bad times. The fact that the literature does not identify a major difference between the short- and long-term effects on credit growth is supportive of the view that borrower-based measures can, in principle, be used as a structural tool with no major impact on long-term credit provisioning.

Figure 2: The Reported Estimates Decline Over Time



Note: The median year is calculated for each collected elasticity based on the time period of the data sample used in the estimation.

3. Publication Bias

The expected effect of a tightening of LTV, DTI or DSTI limits on the provision of bank loans is clearly negative. This strong prior knowledge about the direction of the estimated effect might lead researchers (or publishing outlets) to question or even disregard estimates that are not in line with economic logic. Estimates that are above zero might therefore act as a psychological barrier, suggesting that the data or model employed is incorrect and the estimates are thus unpublishable. Researchers may also get a similar feeling about estimates that are weak in statistical power or small in magnitude, despite them having the expected negative sign.

Publication bias is a phenomenon that arises when researchers do not publish all their estimates, but only those that are significant or have the "correct" sign. Such selection, accompanied by consequent exaggeration of the reported estimates, affects the field of economics to a large degree (e.g. Ioannidis et al., 2017; Astakhov et al., 2019). The following extracts from the sample studies show that selection bias might be a problem even in seminal papers estimating the effects of borrower-based measures on bank credit. For instance, the highly influential study by Akinci and Olmstead-Rumsey (2018) states:

> *"Turning to housing credit, we expect to find that LTV and DSTI caps as well as other housing measures reduce housing credit growth, and we do find the borrower-targeted policies (LTV and DSTI) are associated with lower housing credit growth..." (p. 12)*

In another study by Kuttner and Shim (2016), the authors reveal their a priori expectation about the regression outcome when commenting on the results:

> *"All have the correct (negative) sign, indicating that a policy tightening (coded as +1) reduces credit growth and a loosening (coded as -1) increases credit growth." (p. 12)*

Even the workhorse study in the area of macroprudential policy measures by Cerutti et al. (2017b) suggests that the expected effect of policy tightening on bank credit is negative:

> *"And, importantly, macroprudential policies are meant to be mostly ex-ante tools, that is, they should help reduce the boom part of the financial cycle." (p. 13)*

Given the numerous options as regards both study design and the choice of proper estimation approach, one can always "try harder" to find significant and/or "correctly signed" estimates. In the context of the literature on the effects of the minimum wage on employment, Card and Krueger (1995) argue that some studies have been affected by specification-searching and publication biases. Publication bias has been found not only in economics (Campos et al., 2019; Gechert et al., 2022) and market-based finance (Kim et al., 2019; Gric et al., 2021), but also in some bank-related studies (Zigraiova and Havranek, 2016; Campos et al., 2019). To find out whether the literature describing the effects of borrower-based measures on bank credit might also suffer from these issues, we examine publication bias using a series of graphical and econometric tests. The graphical tests typically use a funnel plot or histogram to show the distribution of the estimated elasticities and related measures of estimate precision, such as standard errors, p-values or t-statistics.

Besides graphical visualization, a wide range of econometric tests can be run to determine whether publication bias is present among the estimates. Specifically, we use several linear techniques (ordinary least squares, weighted least squares, fixed effects, and the hierarchical Bayesian approach). Linear estimation techniques provide solid ground for testing for the presence of publication bias as well as allowing us to estimate the size of the publication bias. They are used to test for publication bias by exploiting the association between the estimated elasticity $\hat{\beta}_{i,j}$ and its standard error $SE_{i,j}$ for each study j (Stanley et al., 2013; Stanley, 2005):

$$\hat{\beta}_{i,j} = \alpha + \gamma SE_{i,j} + \varepsilon_{i,j} \tag{2}$$

where α is the effect beyond bias (the "true" or corrected mean effect) and γ is the intensity of the publication bias. If the γ coefficient is statistically significant, publication bias is present.

Table 4 presents the results of the tests for publication bias. We obtain negative and statistically significant estimates of publication bias across different specifications and estimation methods. The estimates suggest that researchers strongly prefer negative estimates and may have a tendency to discard positive ones. While looking across estimation methods, we record significant estimates of the effect beyond bias especially when accounting for unobserved study-specific characteristics (columns 2 and 4). The bias-corrected coefficients are about half the size of the uncorrected mean of the collected estimates (-1.6). This is consistent with the "rule of thumb" in the economic literature. Ioannidis et al. (2017) suggest that, in economics, publication selection inflates the mean reported coefficients twofold. Overall, the baseline regression shows that estimates of the effect of borrower-based measures may be systemically exaggerated due to the presence of publication selection.

The baseline regression, informative as it already is, tells us little about the sources of publication bias. In theory, there can be two channels at work:

1. Researchers (or publishing outlets) prefer to report estimates with the "correct" negative sign.

2. Researchers (or publishing outlets) mainly want to publish estimates that are statistically significant.

We explore these two channels in the next subsection with a series of graphical and empirical tests.

Table 4: Linear Methods Indicate Strong Negative Publication Bias

	(1) OLS	(2) Study	(3) Precision	(4) FE	(5) Bayes
Constant (effect beyond bias)	-0.44	-0.92**	-0.49	-0.89**	-0.63*
	(0.47)	(0.4)	(0.61)	(0.37)	(0.38)
SE (publication bias)	-0.77***	-0.53**	-0.74***	-0.49***	-0.86***
	(0.16)	(0.27)	(0.26)	(0.13)	(0.29)
Observations	422	422	422	422	422
Studies	23	23	23	23	23

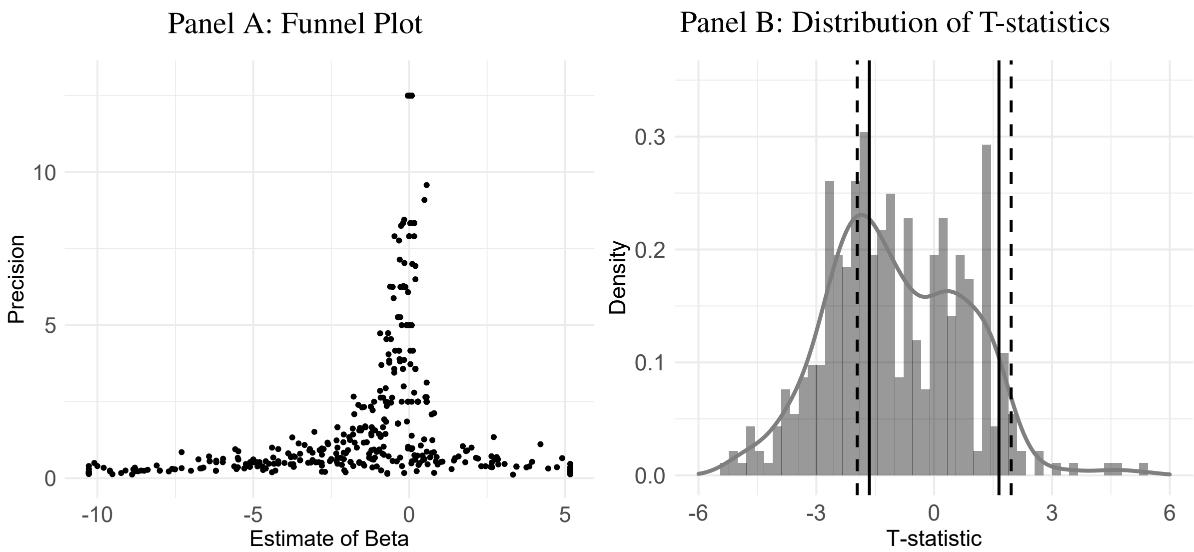
Note: The table presents the results of the regression of equation (2). The standard errors, reported in parentheses, are clustered at both the level of the study and the type of borrower-based measure used in the primary study. OLS – ordinary least squares. Study – the inverse of the number of estimates reported per study is used as the weight. Precision – the inverse of the reported estimate's standard error is used as the weight. FE – study-level fixed effects. Bayes – hierarchical Bayesian approach. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.1 Sources of Publication Bias

In Figure 3, we search for the incidence of the first and second potential sources of publication bias. Panel A in Figure 3 is a funnel plot which depicts the estimates' magnitude on the horizontal axis against their precision (the inverted standard error) on the vertical axis. The most precise estimates should be scattered around the "true" mean effect in the top part of the plot, while less precise estimates should form the tails of the distribution at the bottom. Thus, in the absence of publication selection, the funnel plot is approximately symmetric. Asymmetries would indicate the presence of publication bias. In our case, the funnel plot is asymmetric and visibly skewed toward the negative spectrum of the distribution, showing a clear preference for negative values. Positive estimates appear but are mostly associated with low precision. Hence, the first visual test indicates that search for the "correct" (negative) sign contributes to the documented publication bias.

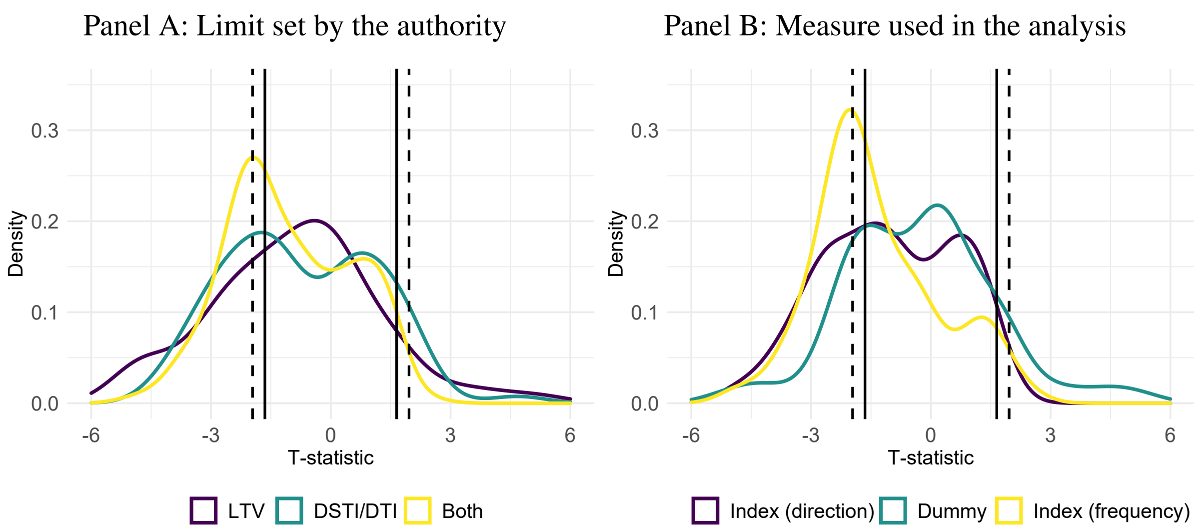
Panel B in Figure 3 is a histogram showing the distribution of the t-statistics. A stylized fact of empirical economics is the under-reporting of estimates that are not statistically significant: researchers prefer to report significant estimates (Cazachevici et al., 2020; Ehrenbergerova et al., 2021; Gechert et al., 2022). In the absence of publication bias, the distribution of the t-statistics should be approximately normal (Egger et al., 1997). We record a two-humped (bimodal) distribution of the reported t-statistics. The distribution is consistent with the view that researchers strongly prefer negative estimates that are significant at least at the 10% level and discard significant positive estimates. Further, there is evident over-reporting of negative estimates that lie between the 5% and 10% significance levels, suggesting that researchers search for "at least some significance" of their negative estimates. It is also apparent from the observed distribution of the t-statistics that when researchers do report positive estimates, they lie outside the region of statistical significance. Overall, we find that both sources of publication bias might be at work.

Figure 3: Positive and Non-Significant Estimates Are Under-Reported



Note: Panel A: Precision is calculated as the inverse of the standard error. In the absence of publication bias, the funnel should be symmetric around the most precise estimates. We exclude estimates with extreme magnitude or precision from the figure but include all in the regressions. Panel B: The vertical lines denote the critical values associated with 5% (dashed line) and 10% (solid line) statistical significance. We exclude estimates with large t-statistics from the figure but include all in the regressions. In the absence of publication bias, the distribution of the t-statistics should be approximately normal.

Figure 4: Two Sources of Publication Bias by Borrower-Based Measures



Note: The vertical lines denote the critical values associated with 5% (dashed line) and 10% (solid line) statistical significance. We exclude estimates with large t-statistics from the figure but include all in the regressions. In the absence of publication bias, the distribution of the t-statistics should be approximately normal.

Panel A in Figure 4 shows that the bimodal distribution is driven by estimates of the effect of DSTI/DTI limits or the effect of their combination with LTV limits. This shows that positive non-significant estimates appear mainly for the introduction or tightening of DSTI/DTI limits. Beyond publication selectivity, this may reflect the fact that the use of DSTI/DTI limits is much less frequent than the use of LTV limits and the low variance may lead to an attenuation bias. It holds that researchers mostly report estimates of both LTV and DSTI/DTI limits that are negative and statistically significant. Inspecting panel B of Figure 4, we find that studies using a frequency-based index to capture borrower-based measures are the main source of the first publication bias (search for correctly signed estimates with at least some significance).

The graphical evidence is supported by empirical tests, the results of which are shown in Table 5. We aim to verify the hypothesis that researchers prefer significant negative estimates over non-significant positive ones. First, we extend equation (2) by a dummy variable which equals one if the estimate is statistically significant at the 10% level. Second, we alter the dummy to equal one if the estimate is significant at the 10% level and has a negative sign. Then we gradually regress the estimate on each of the two dummy variables and the interaction of the dummy with the estimate's standard error. In both specifications, the parameter of the interaction term captures the strength of the publication bias. The estimates in panels A and B of Table 5 confirm our hypothesis. The documented publication bias is found to be driven by elasticities that have the "correct" negative sign and are "just" significant at the 10% level.

Table 5: Publication Bias Is Driven by Selection of Both Sign and Statistical Significance

	(1) OLS	(2) Study	(3) Precision	(4) FE	(5) Bayes
Panel A: significant at 10% level					
Constant	0.09	-0.03	0.04	-0.52	-0.28
	(0.1)	(0.27)	(NaN)	(0.57)	(0.32)
SE	-0.26***	-0.08	-0.22**	-0.18	-0.38
	(0.07)	(0.25)	(0.11)	(0.14)	(0.26)
I(t-stat<1.65)	-1.12*	-1.3**	-1.39	-0.11	-0.73*
	(0.58)	(0.6)	(1.05)	(0.55)	(0.41)
SE×I(t-stat<1.65)	-0.98***	-0.92**	-0.83**	-0.84***	-0.51
	(0.21)	(0.36)	(0.39)	(0.17)	(0.35)
Observations	422	422	422	422	422
Studies	23	23	23	23	23
Panel B: significant at 10% level and negative					
Constant	0.18	0.11	0.06	0.05	-0.07
	(0.2)	(0.28)	(0.04)	(0.56)	(0.3)
SE	-0.07	0.15	0.01	-0.1	-0.25
	(0.12)	(0.14)	(0.14)	(0.14)	(0.26)
I(t-stat<1.65, β<0)	-1.3**	-1.3**	-1.84	-0.92	-1.19***
	(0.53)	(0.64)	(1.18)	(0.58)	(0.45)
SE×I(t-stat<1.65, β<0)	-1.71***	-1.93***	-1.4**	-1.43***	-1.3***
	(0.23)	(0.23)	(0.55)	(0.13)	(0.35)
Observations	422	422	422	422	422
Studies	23	23	23	23	23

Note: The table presents the results of the regression of equation (2) extended by additional dummy variables for collected elasticities significant at the 10% level (I(t-stat<1.65)) and elasticities that are negative at the same time (I(t-stat<1.65, β<0)). The standard errors, reported in parentheses, are clustered at both the level of the study and the type of borrower-based measure used in the primary study. OLS – ordinary least squares. Study – the inverse of the number of estimates reported per study is used as the weight. Precision – the inverse of the reported estimate's standard error is used as the weight. FE – study-level fixed effects. Bayes – hierarchical Bayesian approach. *** p < 0.01, ** p < 0.05, * p < 0.1.

3.2 Dummy- versus Intensity-Adjusted Approach to Capturing Macroprudential Policy Measures

The dominant stream of literature relies on the dummy-coded approach to capturing the introduction of, or changes to, borrower-based measures. As we have already mentioned in this paper, these dummy-type indices indicate the direction of a policy change (at best) and lack information on the intensity of the change. While this makes the economic interpretation of the estimated coefficients challenging (but not impossible), it actually represents the second-best solution.

Ideally, one would measure the intensity of borrower-based measures. For example, for the LTV limit it is possible to use the actual percentage change to the requirement, for instance, a 10 pp change from a cap of 90% to one of 80%. However, one needs to keep in mind that such regulation can take different forms and tightening or easing might not always transmit to a change in the numerical setting of the limit. In countries such as Korea and Hong Kong, which have used LTV limits actively, different borrowers face different LTV limits based on certain characteristics (for example, the location of the property, whether it is the first or second home/mortgage, or the volume of the mortgage). The settings of these characteristics also change quite frequently. Using data on macroprudential policy actions from Alam et al. (2019), we find that only about half of LTV tightening actions contain a change to the numerical LTV limit. Thus, it is not a straightforward task to capture the overall intensity of an LTV cap in a single country, much like it is difficult to obtain a cross-country comparable measure. Attaching a value to the degree of intensity of a given regulation inevitably involves a degree of subjectivity, notwithstanding the fact that without access to detailed mortgage market data, one can hardly know how much a policy is actually binding, which may create attenuation bias.[8]

Bearing in mind the limitations of both the dummy-based and the intensity-adjusted approach, we undertake efforts to facilitate economic interpretation of our "true effect" (beyond bias) estimates. Specifically, we consult the integrated Macroprudential Policy (iMaPP) database maintained by IMF staff, which offers a unique numerical indicator of regulatory limits on the LTV ratio for over 60 countries over the 1990–2019 period (Alam et al., 2019). The time period mimics the maximum time span of the studies included in our sample. The database combines information from five databases that were dominantly used in our surveyed studies, so we expect the country coverage to be high as well. We focus on the application of LTV limits because the database only contains information about the intensity of this measure (out of the three main borrower-based measures, LTV, LTI, and DSTI). Moreover, it is changed the most frequently (Cerutti et al., 2016; Alam et al., 2019). From the database, we draw information on the intensity of LTV changes (tightening) in individual countries during the 1990–2019 period and sort the policy actions into six buckets – 5, 10, 15, 20, 25, 30 – according to the size of the LTV change. If the distribution of changes to LTV turns out to be non-normal and visibly skewed (toward smaller or larger changes), it can help place the estimated true effect coefficient into the economic context and verify its meaningfulness for policymakers.

Table 6 shows the frequency of changes to LTV grouped into the six buckets. Apparently, the most frequent change to the LTV limit is one of less than 10 pp, while changes of approximately 5 pp are the most frequent. Under the (strong) assumption that numerical changes to the LTV limit drive the

[8] In fact, since the dummy approach imperfectly measures the intensity of policy changes, it is also prone to attenuation bias, which should bias the coefficients of the estimated borrower-based measures toward zero. The fact that the literature using the dummy-based approach finds a significant relationship between borrower-based measures and bank lending (despite the measurement error) is encouraging.

estimates in the literature, the "true effect" estimate (a decline in credit growth of about 0.9 pp) can be attributed to a change to the LTV limit of less than 10 pp, or even more likely of less than 5 pp.

Table 6: Intensity-Adjusted LTV Shock

LTV change (in pp)	Frequency	% of actions	No. of countries
≈ 5	29 (51)	34% (48%)	16 (21)
≈ 10	24	28% (22%)	16
≈ 15	8	9% (7%)	7
≈ 20	11	13% (10%)	10
≈ 25	7	8% (7%)	7
≈ 30	6	7% (6%)	6
Total	85 (107)		62 (67)

Note: The table shows the individual numerical changes to the LTV limit grouped into six buckets. The buckets are specified using the following rules. The first bucket contains a percentage point change that falls within the interval (2.5, 7.5), while the second bucket is defined as (7.5, 12.5), the third as (12.5, 17.5), the fourth as (17.5, 22.5), the fifth as (22.5, 27.5), and the sixth as (27.5, 32.5). Numbers in brackets represent a different version of the rules where the first bucket lies within the interval (1.0, 7.5) and hence also counts very small, gradual changes to the LTV limit. For example, in the Netherlands, the LTV limit was applied gradually, with a change of 1 pp per year.

3.3 Robustness Tests

So far, we have identified two sources of publication bias: positive significant and negative non-significant estimates are under-reported in the existing literature. This finding may reflect either a strong personal a priori belief of the researcher or pressure to publish significant estimates that are in line with the theory or the dominant stream of literature. We now discuss and test the robustness of our results by addressing shortcomings stemming from the linear regression technique employed.

First, linear estimation assumes – by design – a linear relationship between the reported elasticity and the standard error. A useful analogy appears in Nansen McCloskey and Ziliak (2019), who liken publication bias to the Lombard effect in psychoacoustics: speakers increase their effort in the presence of noise. However, the linearity condition might not always be valid. For instance, Stanley et al. (2010) show that the most precise estimates are less likely to be contaminated by publication bias, which violates the linear relation between publication bias and the standard error. This finding has led to the development of new estimation methods that relax the assumption of linearity (see Stanley et al., 2010; Ioannidis et al., 2017; Furukawa, 2019; Bom and Rachinger, 2019; Andrews and Petroulakis, 2019). These methods build on the assumption that more precise estimates are less likely to suffer from publication bias; therefore, they try to isolate them and use them to compute the average effect.[9] Although this approach has proved to be useful in many studies (see, for instance, Bajzik et al., 2020; Havranek et al., 2021), it is not suitable for us due to the presence of two

[9] Ioannidis et al. (2017) propose a procedure that focuses only on estimates with statistical power above 80%. A similar approach is proposed by Stanley et al. (2010), who suggest focusing on the top 10% of estimates in terms of precision (the Top 10 method). The stem-based method of Furukawa (2019) suggests using only the stem of the funnel plot, that is, a portion of the most precise estimates. This portion is determined by minimizing the trade-off between bias (increasing the number of imprecise estimates included) and variance (reducing the number of estimates included). The kinked method proposed by Bom and Rachinger (2019) builds on the idea that estimates are automatically reported if they cross a certain precision threshold; therefore, they introduce an "endogenous kink" technique that estimates this threshold. The selection model by Andrews and Petroulakis (2019) first identifies the conditional publication probability (the probability of publication as a function of a study's results) and then uses it to correct for publication bias. The underlying intuition involves jumps in publication probability at conventional p-value cut-offs.

sources of publication bias. By applying these methods, we would focus only on the first source – over-reporting of "just" significant negative results – while ignoring the second source – under-reporting of significant positive results. Moreover, we hesitate to employ these methods given the relatively modest number of observations at our disposal. Most of these methods work by searching (exogenously or endogenously) for a precision threshold below which estimates are discarded. As a result, they remove a substantial number of imprecise estimates, in our case imprecise positive elasticities, shifting the entire distribution to the negative region and yielding a potentially inflated negative effect.

Second, the standard error is assumed to be exogeneous. Even the exogeneity condition might not hold, since publication bias can work through both point estimates and standard errors, which are computed using different approaches in different studies. Stanley (2005) argues that since the standard error itself is estimated, estimates of equation 2 might suffer from attenuation bias. We test the exogeneity condition by employing various methods that are robust to the endogeneity of the standard error. The estimates are given in Appendix B. These methods are based on verifying the presence of "p-hacking", i.e., a situation where authors prefer to include significant estimates rather than non-significant ones. These methods include the Caliper test first proposed by Gerber and Malhotra (2008a,b), the p-curve method developed by Simonsohn et al. (2014a,b), and the Elliott et al. (2022) tests.

The Caliper test specifies narrow bounds surrounding the thresholds commonly employed for t-statistics (1.645, 1.96, 2.58). We use caliper sizes of 0.1, 0.2, and 0.3. In the absence of p-hacking, the distribution above and below the threshold should be close to identical. The results, presented in Table B1, show that publication selection is present mainly for negative estimates. If we test the parameter value against the value of 0.4, which conforms to the 60:40 distribution around the threshold (instead of 50:50) as reasoned in Bruns et al. (2019), evidence for publication selection is present for negative estimates and also when all the estimates (positive and negative) are tested together.

The p-curve method and the Elliott et al. (2022) tests are centered around testing the distribution of p-values. They both study the distribution of t-statistics and p-values around classical significance thresholds (1%, 5%, 10%). When authors try to find significant estimates, the surroundings around these values are skewed. We first employ the p-curve method, which tests the null hypothesis that the literature has no evidential value. The evidential value thus conforms to the existence of strong prior knowledge on the likely effect of borrower-based measures on bank lending. In other words, there should be no effect of borrower-based measures on bank lending beyond publication bias. The only objective of testing for evidential value is to rule out selective reporting as a likely explanation for a set of statistically significant findings (Simonsohn et al., 2014b). Based on Figure B1, we obtain evidence of strong evidential value, as the p-value distribution is visibly left-skewed. Figure B2 further shows that p-hacking mostly concerns negative estimates at a 5% and 10% level of significance and also suggests the existence of over-reporting of non-significant positive estimates. Table B2 shows the results of the Elliott tests (a series of binomial, Fisher, and density discontinuity tests) with a null hypothesis of no p-hacking. Details on the method are available in Appendix B. We reject the null of no p-hacking after considering more powerful statistical tests.

As a final robustness check, we re-run the analyses using the whole sample of collected elasticities, transformed using the PPCs. The estimates are presented in Appendix C. In general, they largely conform to the conclusion stemming from the baseline estimates. Even when using the whole sample of elasticities, we continue to find strong empirical support for the hypothesis that the estimated effect of the introduction or tightening of borrower-based measures on annual bank credit growth is prone to publication bias driven by both selection of the "correct" sign and search for statistical significance (see Figure C1 and Table C2).

4. Drivers of Heterogeneity

Given the differences observed across the primary studies, one can object that the heterogeneity of the estimates might have a different origin than publication bias. For instance, it might be caused by different data, methods, and information about the country of origin employed in the primary studies. Therefore, we aim to identify the factors that drive the proposed relationship the most. In doing so, we use model-averaging methods that address both model uncertainty and omitted variable bias issues. Moreover, these methods allow us to order the factors that drive the heterogeneity in the collected estimates by their importance.

Based on the above-mentioned reasons, we controlled for more than 20 primary study characteristics in three areas – data characteristics, model specification and estimation, and publication characteristics. Before arriving at the final set of control variables, we collected a broader set of potential characteristics. We employed two simple decision rules to remove unsuitable variables. First, the primary study characteristic must be present in at least 20% of all articles or collected estimates. Second, the correlation of all control variables must be below 80%. Although we set these criteria somewhat arbitrarily, we also tried different thresholds without observing any major impact on the results. The final set of characteristics, including definitions and summary statistics, is shown in Table 7. Next, we provide a brief reasoning for the inclusion of specific control variables.

Data characteristics. Information about the effects of different data characteristics on the relationship between borrower-based measures and bank lending is scarce. The closest meta-analytical studies are by Malovaná et al. (2021) on the relationship between capital-based measures and lending, and by Havranek and Rusnak (2012) and Ehrenbergerova et al. (2021) on monetary policy transmission. Malovaná et al. (2021) find, among other things, that the expression of the dependent variable (in their application – different expressions of the bank capital ratio), the type of credit considered, and different midpoint of the data matter. Havranek and Rusnak (2012) and Ehrenbergerova et al. (2021) reveal several discrepancies caused by different length of the data sample and by different data frequency. Thus, in our study, we control for the type of credit used as the dependent variable, the type of borrower-based measure used as the independent variable, data frequency, the midpoint of the data, the number of countries in the estimation sample, the region of the analysis, and data confidentiality.

Model specification and estimation. Other meta-analytical papers reveal a key impact of study design on the size and direction of the estimated effect (Malovaná et al., 2021; Zigraiova et al., 2021). Therefore, we control for a number of characteristics related to model specification and estimation. First, we distinguish between static and dynamic models. Dynamic model specifications contain a lagged dependent variable, which is supposed to capture persistence in credit dynamics.

Table 7: Variable Definitions and Summary Statistics

Variable	Definition	Mean	SD	W. Mean	W. SD
Estimate	The reported estimate of the beta coefficient.	-1.46	3.15	-1.7	3.53
Std. error	The reported standard error of the beta coefficient.	1.47	1.54	1.64	1.7
Data characteristics					
LTV	= 1 if the borrower-based measure (independent variable) captures the effect of the limit on the loan-to-value ratio.	0.24	0.42	0.38	0.49
DSTI/DTI	= 1 if the borrower-based measure (independent variable) captures the effect of the limit on the debt service-to-income ratio or debt-to-income ratio.	0.19	0.39	0.19	0.4
Index (direction)	= 1 if the borrower-based measure (independent variable) is a direction index.	0.47	0.5	0.41	0.49
Index (frequency)	= 1 if the borrower-based measure (independent variable) is a frequency index.	0.27	0.45	0.24	0.43
Household loans	= 1 if household credit is used as the dependent variable.	0.5	0.5	0.69	0.46
Midpoint	The logarithm of the midpoint of the data sample.	2.64	0.34	2.72	0.31
No. of countries	The logarithm of the number of countries in the data sample.	2.84	1.38	2.71	1.53
Confidential data	= 1 if confidential (supervisory) data are used (as opposed to publicly available data).	0.23	0.42	0.31	0.47
Annual frequency	= 1 if the data frequency is annual.	0.33	0.47	0.34	0.48
Europe	= 1 if the study covers a country or group of countries from Europe.	0.35	0.48	0.23	0.42
Model specification and estimation					
Lagged by 1Y or more	= 1 if the estimate is lagged by a year (4 quarters) or more.	0.25	0.43	0.13	0.33
Contemporaneous	= 1 if the estimate is contemporaneous (not lagged at all).	0.21	0.4	0.26	0.44
House prices in eq.	= 1 if the model includes house prices as a control variable.	0.17	0.38	0.14	0.35
Some interaction in eq.	= 1 if the model contains some interaction term (discrete or continuous) with the borrower-based measure.	0.34	0.47	0.16	0.37
Add. regulatory var. in eq.	= 1 if the model contains an additional regulatory variable (borrower-based measure or capital-based measure) on top of the studied borrower-based measure.	0.23	0.42	0.24	0.43
Dynamic model	= 1 if the model is dynamic, i.e., contains a lagged dependent variable.	0.72	0.45	0.8	0.4
GMM method	= 1 if the general method of moments (GMM) is used.	0.18	0.39	0.39	0.49
Fixed-effects method	= 1 if a fixed-effects (FE) regression method is used.	0.34	0.47	0.3	0.46
Time fixed effects incl.	= 1 if time fixed effects are included.	0.48	0.5	0.27	0.45
Publication characteristics					
Publication year	The logarithm of the publication year of the primary study minus the earliest publication year in our dataset plus one.	1.73	0.71	1.56	0.71
Impact factor	The recursive impact factor.	0.76	0.5	0.83	0.56
Citations	The logarithm of the number of citations divided by the number of years from publication until 2021.	2.75	1.05	2.56	1.47
Published	= 1 if the primary study was published in a journal with an impact factor.	0.39	0.49	0.36	0.48

Note: The table presents definitions and summary statistics of primary study characteristics in the analysis of heterogeneity. W. Mean – weighted mean; W. SD – weighted standard deviation; weights are calculated as the inverse of the number of estimates reported per study.

Second, we account for different lag structures and missing key control variables. We also search for more specific factors, such as the presence of additional regulatory variables[10] and interaction

[10] Additional regulatory variables included in the same estimation equation on top of the borrower-based measures may distort the relationship studied. Specifically, the other regulatory variables may have the opposite sign to the borrower-based measure, making it hard to deduce the true size and sign of the relationship studied.

terms[11] in the model specification. Third, we scrutinize the impact of different estimation techniques. We distinguish between OLS, GMM, and fixed-effects regression.

Publication characteristics. These characteristics reveal the correlation between primary study estimates and unobserved features of primary study quality. Among these variables we include the publication year, the discounted recursive impact factor, the annualized number of citations, and dummy variables indicating whether the study was published in a journal. Other meta-analytical papers show that all these characteristics can have some impact on the estimated coefficients (Araujo et al., 2020; Bajzik et al., 2020; Valickova et al., 2015).

4.1 Estimation Method for Analyzing Drivers of Heterogeneity

Since we collect a large number of control variables, it is not convenient to use OLS for the estimation. First, if we included all of the variables, the model would not be parsimonious. Second, if we did not include all of the variables, the estimates might suffer from omitted variable bias or we would have to contend with model uncertainty issues or with best model selection. In order to tackle all of these possible problems, we instead use model-averaging approaches – both Bayesian and frequentist. Model-averaging approaches do not reject any of the possible explanatory variables in advance, which is crucial when we aim to clarify the heterogeneity among the primary estimates. Furthermore, Bayesian model averaging – besides supplying an estimate for each variable and its standard deviation – provides the probability of inclusion in the underlying model for each variable.

Bayesian model averaging (BMA) aims to provide the best possible approximation of the distribution of each regression parameter. Since our data provides 31 potential explanatory variables, we have 2^{31} model combinations (without interaction terms). To run such a process would be very time-consuming, so we employ the Markov chain Monte Carlo (MCMC) process with the Metropolis-Hastings algorithm. This algorithm then only goes through the most probable models (Zeugner and Feldkircher, 2015). In addition, the algorithm assigns a weight to each of the most probable models with respect to the goodness of fit of the other possible models. This weight is called the posterior model probability (PMP). From the PMPs of the all relevant models, the posterior inclusion probability (PIP) of each explanatory variable is estimated. This probability ranges from one, meaning the variable is included in every relevant model, to zero, meaning it does not influence the relationship of interest at all. The coefficient and standard deviation of each explanatory variable are derived and weighted using the variable coefficients from the relevant models and their PMPs as well.

The question naturally arises of how to find the most probable models or where to start in searching for them. Thus, BMA requires some prior information about the regression coefficients (the g-prior) and regarding the models (the model prior). Following Eicher et al. (2011), we use a unit information g-prior (UIP) and a uniform model prior in the baseline setting. This expresses our lack of knowledge about the particular probabilities of the parameter values, since in this setting the zero regression coefficient has the same weight as one observation in the data.

[11] Interaction terms explore the heterogeneity in the effect analyzed. The most common interactions are the size of the bank and GDP growth. The empirical literature suggests that the effect of macroprudential policies is significant for small and medium-sized enterprises, while at least LTV is much less effective for larger banks (Ayyagari et al., 2018; Morgan et al., 2019). Regarding GDP growth, the literature suggests that the policy instruments are countercyclical (Budnik, 2020; Lim et al., 2011).

We examine the sensitivity of our results by choosing different priors in the robustness checks. For example, we employ the dilution prior proposed by George (2010) as an alternative choice for the model prior. The dilution prior adjusts the model probabilities by the determinant of the correlation matrix of the variables included in the proposed model. Hence, in the case of high correlation, the determinant is close to one and assigns only a small weight, and in the case of low correlation it is close to zero and assigns a large weight. This approach has been used in recent meta-analyses in economics – see, for instance, Gechert et al. (2022) and Bajzik et al. (2020). We also employ a combination of the Hannan-Quinn (HQ) g-prior, which adjusts data quality, and the random model prior (Fernandez et al., 2001; Ley and Steel, 2009; Feldkircher and Zeugner, 2012; Zigraiova et al., 2021). Next, we employ a combination of the BRIC g-prior, which minimizes the prior effect on the result, and the random model prior, which attaches an equal probability to every model size (Zeugner and Feldkircher, 2015; Gechert et al., 2022). Last but not least, we use frequentist model averaging (FMA) and a frequentist check (OLS) as an additional two robustness checks.

Following the approach proposed in Eicher et al. (2011), in the following section we only interpret variables with a PIP above 0.5. The categorization is as follows. The variable is classified as decisive if its PIP is higher than 0.99, strong if it is between 0.95 and 0.99, substantial if it is between 0.75 and 0.95, and weak if it is between 0.5 and 0.75. The 0.5 threshold serves as a baseline for including variables in the OLS frequentist check.

Reference model. We define our reference model based on the predominant characteristics of the primary studies. This means that for each group of dummy variables (e.g., the contemporaneous effect, the effect lagged by less than one year, and the effect lagged by more than one year) we drop the most frequent one to avoid the dummy variable trap. The retained variables are then compared to the dropped one – the reference variable. Of course, we give ourselves the freedom to choose to include the more appropriate variable in the reference model. We use this option where the difference between two characteristics is small or in order to aid interpretation.

4.2 Results

The results of the Bayesian model averaging are visualized in Figure 5. In the figure, each column represents an individual regression model, while the column width captures the posterior model probability. The rows denote the individual variables included in each model. We order the variables by the value of their posterior inclusion probability from top to bottom in descending order. Red color (lighter in grayscale) indicates a negative sign on the variable's coefficient, i.e., the estimated effect is more negative. Blue color (darker in grayscale) indicates a positive sign, i.e., the effect is less negative. Where a variable is excluded from the particular model, the cell is left blank. Results using different priors are presented in Figure A3 in the Appendix.

The corresponding numerical results are reported in Table 8. We also show two alternative estimations to the baseline BMA. First, we estimate frequentist model averaging, including the same set of variables as used in the BMA. Second, we estimate simple OLS, excluding variables that turned out to be less important in the BMA exercise (i.e., whose posterior inclusion probability is below 0.5).

Publication bias. The presence of publication bias in the collected estimates is supported by the evidence across all the models we run. The reported negative elasticities are found to be systemically exaggerated due to publication bias even if we control for data and publication characteristics and estimation method. The size of the publication bias across the different methods is around -0.6 (Table 8), which corresponds to the estimated magnitude of the publication bias in

Section 3. We conclude that our previous finding of significant publication bias was not driven by omitting factors associated with heterogeneity. We further find additional factors that explain the observed heterogeneity in the estimated elasticities.

Model specification and estimation. Apart from the standard error, the highest PIP is recorded for several characteristics linked to the estimation method and model specification chosen by the researchers. For one, we find that those studies which employ the generalized method of moments (GMM) estimator record significantly more negative estimates of the effects of borrower-based measures on bank lending than those using the ordinary least squares (OLS) estimator. In theory, the GMM estimator should be more successful than the OLS estimator in reducing endogeneity bias (Baum et al., 2003). In fact, the GMM was used in most of the leading studies in our sample (based on a citation count) with the purpose to "mitigate endogeneity concerns" (Cerutti et al., 2017a; Akinci and Olmstead-Rumsey, 2018). Modeling of the impact of borrower-based measures on bank lending is prone to endogeneity bias, especially since decisions to take policy actions are made with respect to the current state and future development of the financial sector, so that in equation (1), $cov(BBM, \varepsilon) \neq 0$.[12] OLS estimates of the $\hat{\beta}$ elasticity in equation (1) may therefore be biased toward zero, given that $cov(BBM, \varepsilon) > 0$. Under the assumption that the GMM estimator reduces the (positive) endogeneity bias by considering a correct set of instruments (Stock et al., 2002), it is expected that GMM-based studies find more negative elasticity estimates.

We further find that contemporaneously specified models, meaning that researchers estimate the impact of the introduction or tightening of borrower-based measures at time t on bank lending within the same month, quarter or year, deliver lower negative elasticity estimates than models specified with a lag. In other words, the estimated effect of borrower-based measures is found to be stronger with a lag, suggesting that the impact of the measures takes time to materialize. Consider a hypothetical case where policymakers tend to tighten the borrower-based requirements when housing credit is already expanding rapidly. This would give rise to a positive correlation between the borrower-based dummy and credit, partially (or fully) offsetting (or miscalculating) the desired policy effect. Poghosyan (2020) also estimates the impact of borrower-based measures on bank credit to be delayed. In his setup, it reaches its peak only after three years. McDonald (2015) estimates the impact of LTV loosening to be significantly positive only when credit growth was previously strong.

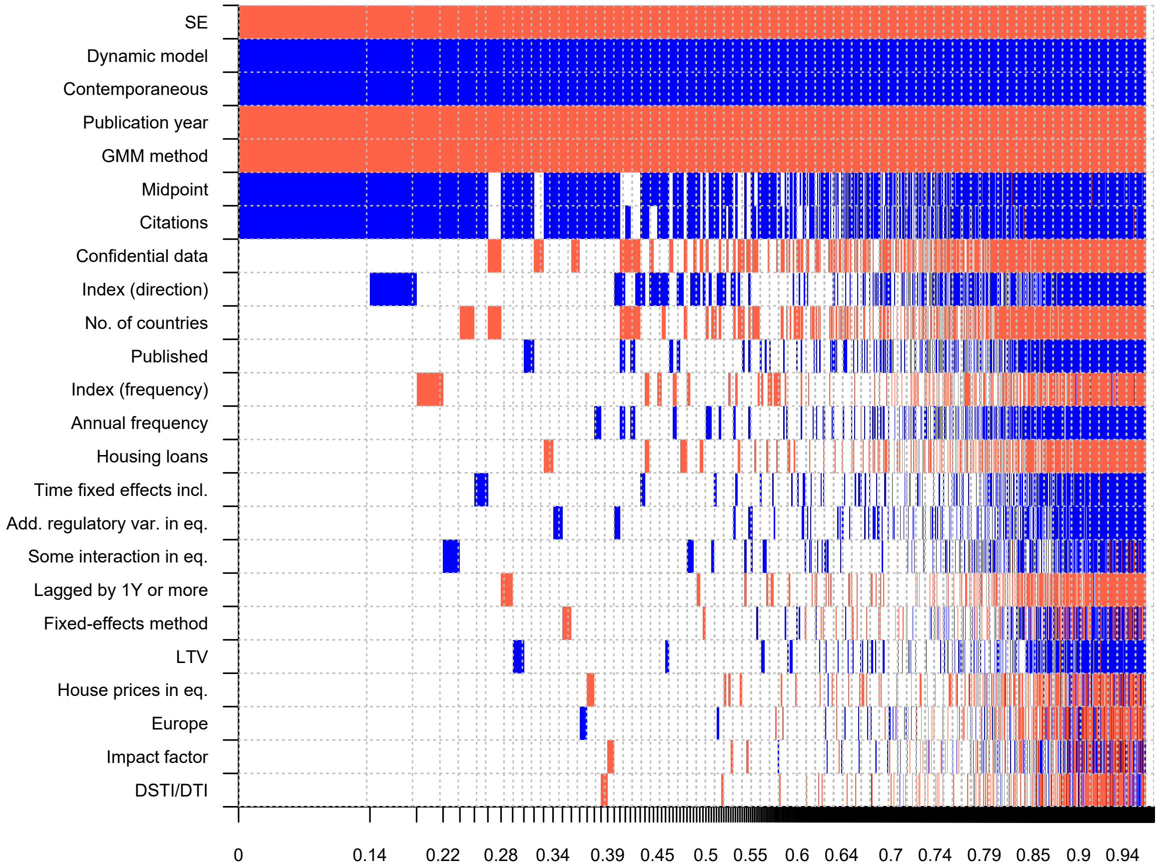
Last, we discover that the inclusion of a lagged dependent variable in the model specification (i.e., the use of a dynamic model specification) produces a lower negative estimate of the relationship between borrower-based measures and bank lending. The inclusion of a lagged dependent variable generally explains a lot of the variance in the credit dynamics. Given the high persistence of bank credit, a model which does not contain a lagged dependent variable could produce an overestimated negative elasticity linked to borrower-based measures.

Data characteristics. We find that papers using confidential data report a more negative effect of borrower-based measures on bank lending. Confidential data usually capture developments at the level of a single entity or product (e.g., bank, firm or loan) and are usually available to supervisory or regulatory institutions only. Studies using such data may in theory be more successful than others in correctly identifying exogenous regulatory shocks that are plausibly unrelated to lending

[12] The literature studying the effects of the application of borrower-based measures on the financial sector is, in general, highly sensitive to the presence of endogeneity bias. Borrower-based measures are set on a country level and are a direct response to developments in the real estate mortgage market. This is in contrast to the application of capital-based measures (and the literature studying the effects thereof), of which only the countercyclical buffer is used with respect to the financial cycle development.

opportunities due to more detailed data. As a result, they are less likely to suffer from the endogeneity bias described above.[13] What is more, macro-level studies capture averages across constrained and unconstrained agents, which can cause the effect of such regulation to be underestimated. We further find that studies using data from a larger sample of countries produce more negative elasticity estimates. Tracking the effect of the application of borrower-based measures in more countries can make up for the above-mentioned disadvantage of macro-level studies due to the large focus on the cross-sectional aspect.

Figure 5: Model Inclusion in Bayesian Model Averaging



Note: The response variable is the estimated effect of borrower-based measures on credit growth. Columns denote 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 color (darker in 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 color (lighter in 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.

[13] For example, the effectiveness of introducing borrower-based limits on lending growth 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 the potential bias.

Publication characteristics. The BMA estimates indicate a strong association of two publication characteristics with the collected estimates – the year of publication and the number of citations of the paper. We find that studies published more recently report more negative elasticity estimates. We interpret this relationship as a potential effect of quality: more recent studies have access to better data as well as better estimation methods that allow them to filter out pollution related to measurement errors or endogeneity that biases the estimates toward zero. Another explanation lies in the fact that the majority of borrower-based measures were brought in after the GFC of 2007–2009. Recent studies can thus exploit much more variation in their regulatory proxies. In addition, we find quite the opposite relationship when considering the number of citations, owing to the fact that newer studies will by definition have fewer citations on average. Highly cited documents are those with less negative elasticity estimates.

Table 8: What Drives the Heterogeneity of the Collected Estimates

	Bayesian model averaging			Frequentist model averaging			Frequentist check (OLS)		
	P.mean	P.SD	PIP	Coef.	SE	p-value	Coef.	SE	p-value
Constant	-6.572	-	1.000	-1.908	4.565	0.676	-9.703	3.944	0.014
SE	-0.599	0.105	1.000	-0.453	0.114	0.000	-0.643	0.087	0.000
Data characteristics									
Midpoint	2.142	1.456	0.743	0.800	1.321	0.545	2.989	1.059	0.005
Confidential data	-0.526	0.850	0.333	-1.825	0.990	0.065			
No. of countries	-0.090	0.176	0.269	-0.586	0.223	0.009			
Index (direction)	0.232	0.446	0.269	0.786	0.549	0.152			
Annual frequency	0.094	0.303	0.139	0.981	0.599	0.102			
Index (frequency)	-0.095	0.299	0.138	-0.215	0.657	0.744			
Housing loans	-0.068	0.228	0.127	-0.927	0.482	0.054			
LTV	0.026	0.133	0.076	0.994	0.541	0.066			
Europe	-0.020	0.171	0.071	-0.736	0.632	0.245			
DSTI/DTI	-0.005	0.095	0.052	0.683	0.567	0.228			
Model specification and estimation									
Dynamic model	2.571	0.467	1.000	2.601	0.567	0.000	2.730	0.400	0.000
Contemporaneous	1.865	0.400	1.000	1.793	0.452	0.000	1.855	0.542	0.001
GMM method	-2.442	0.684	0.995	-1.528	0.759	0.044	-2.841	0.826	0.001
Add. regulatory var. in eq.	0.050	0.205	0.098	0.751	0.461	0.104			
Time fixed effects incl.	0.048	0.196	0.098	0.421	0.656	0.522			
Some interaction in eq.	0.034	0.156	0.089	0.142	0.535	0.791			
Lagged by 1Y or more	-0.038	0.184	0.082	-0.507	0.491	0.301			
Fixed-effects method	0.010	0.202	0.079	0.196	0.681	0.773			
House prices in eq.	-0.029	0.187	0.073	-0.413	0.614	0.501			
Publication characteristics									
Publication year	-1.414	0.312	0.997	-1.153	0.400	0.004	-1.385	0.386	0.000
Citations	0.441	0.327	0.715	0.054	0.286	0.851	0.673	0.313	0.032
Published	0.103	0.300	0.154	0.993	0.492	0.044			
Impact factor	-0.004	0.093	0.057	0.017	0.412	0.967			

Note: P. mean – posterior mean, P. SD – posterior standard deviation, PIP – posterior inclusion probability, SE – standard error. Bayesian model averaging employs the combination of the uniform model prior and the unit information g-prior recommended by Eicher et al. (2011). Frequentist model averaging applies Mallow's weights (Hansen, 2007) using orthogonalization of the covariate space as suggested by Amini and Parmeter (2012) to reduce the number of models estimated. The frequentist check (OLS) includes the variables with a PIP estimated by BMA of 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 7.

5. Concluding Remarks

We present the first quantitative synthesis of the empirical literature on the effects of macroprudential borrower-based measures on the extension of bank credit. Borrower-based measures rank among the most commonly used macroprudential policy measures worldwide and have seen increased use in the post-GFC period. Despite the broad application of LTV, DTI, and DSTI limits, there is little consensus on how the toolkit works in practice. The existing literature is especially interested in the relationships between the use of these policies and developments in credit and housing markets, with a view to analyzing the effectiveness of these policies in managing credit and financial cycles.

Given the complexity of the transmission of macroprudential policy tools, the meta-analysis performed in this paper provides an important service to both researchers and policymakers. We synthesize the empirical evidence from a unique dataset of more than 700 estimates of the elasticity between the application of borrower-based measures and bank lending collected from 34 primary studies. We provide the following key findings.

The use of different borrower-based limits may affect bank lending in a dissimilar fashion. On inspecting the basic statistical properties of the dataset, we find that income-based DTI/DSTI limits may be more effective than the value-based LTV limit in terms of reducing credit growth. The mean negative effect of introducing or tightening DTI and DSTI limits on bank lending is about 20% stronger than that of the LTV limit. Furthermore, the statistics indicate that joint application of both LTV and income-based limits is most effective, with a negative effect on annual credit growth of about -2.1 pp (almost two times the effect of using the limits individually). Nevertheless, the model-averaging techniques did not confirm that these differences are statistically significant. While it may be too early to arrive at a definitive conclusion on the strength of the effects of individual borrower-based tools, the summary statistics from the meta-analytical dataset offer much-needed early evidence on the matter and highlight a fruitful ground for future research.

Furthermore, the meta-analytical summary of the collected elasticities shows that the effect of borrower-based limits on bank lending is growing stronger over time, as more recent studies report more negative estimates. This can be explained either by the effect of quality, where more recent studies employ better methods, or by the effect of quantity, where newer studies can exploit greater variation in borrower-based policy. It may also reflect the changing macro-financial environment after the Global Financial Crisis of 2007-2009, where the onset of a low interest rate environment depressed banks' profitability and capital, making them more likely to respond to regulatory tightening.

Using a series of empirical tests, we next find that estimates of the impact of borrower-based measures are prone to publication selection. This stems from the fact that borrower-based measures are expected to have a negative effect on bank lending. Positive or statistically non-significant estimates might thus act as a psychological barrier, suggesting that the data or model employed are incorrect and the estimates are thus unpublishable. Using a plethora of methods, we find that when publishing results, researchers show a preference for elasticities that have the "correct" negative sign and are "just" significant at the 10% level. The identified publication bias exaggerates the mean estimate by almost 40%.

Using Bayesian and frequentist model averaging and controlling for a large number of confounders – study characteristics – we confirm that the single most important variable for explaining the variation in the reported elasticities is the standard error. A large standard error is found to be associated with large negative estimates. In the absence of publication selection, the observed effects of borrower-based measures should vary randomly around the "true" value, independently of the standard error (Stanley, 2005). The violation of independence suggests a preference for large estimates that compensate for large standard errors. On top of that, we find that a large part of the variation in the reported elasticities can be explained by model specification, estimation method, and underlying data characteristics.

In particular, we find that studies which take steps to reduce endogeneity bias report more negative estimates. Endogeneity of the borrower-based policy variable is likely to bias estimates of the impact of borrower-based policy on bank lending toward zero, given the fact that decisions to take policy actions are made with respect to the state of the financial sector. The existence of endogeneity bias in the literature is supported by the fact that we find the use of confidential data to be robustly associated with the study results: studies built on confidential data report more negative estimates.

Our evidence also suggests that contemporaneous models might be misspecified and under-estimate the effect of borrower-based measures on bank lending. The estimated effect of borrower-based measures is found to be stronger with a lag, suggesting that the impact of the measures takes time to materialize. Tracking the impact of borrower-based measures on bank lending within the same month or quarter in which they were applied can lead to a positive correlation between the variables, as the measures are usually applied in the growth phase of the financial cycle.

Our meta-analytical evidence has important policy implications. Overall, the empirical summary of the literature provided in the paper shows that borrower-based measures are effective policy tools in terms of directly restricting (mortgage) credit growth. However, we find that the existing negative estimates are systemically exaggerated due to the presence of publication selection and insufficient identification power of the modeling framework employed. A central banker wishing to calibrate the effect of introducing or tightening borrower-based measures in her stress-testing framework would have a difficult job finding the correct elasticity value. The evidence provided in our paper can serve as a useful benchmark against which countries with no historical experience with the use of borrower-based measures can check their forecasting or stress-testing models.

References

ACHARYA, V. V., K. BERGANT, M. CROSIGNANI, T. EISERT, AND F. MCCANN (2020): "The Anatomy of the Transmission of Macroprudential Policies." IMF Working Papers 20/58, International Monetary Fund.

ACHARYA, V. V., K. BERGANT, M. CROSIGNANI, T. EISERT, AND F. J. MCCANN (2020): "The Anatomy of the Transmission of Macroprudential Policies." Working Paper 27292, National Bureau of Economic Research.

AFANASIEFF, T. S., F. CARVALHO, E. C. DE CASTRO, R. L. COELHO, J. GREGÓRIO, ET AL. (2015): "Implementing Loan-to-Value Ratios: The Case of Auto Loans in Brazil (2010-11)." Working Paper 380, Banco Central do Brasil.

AHMADOV, A. K. (2014): "Oil, Democracy, and Context: A Meta-Analysis." *Comparative Political Studies,* 47(9):1238–1267.

AHUJA, A. AND M. NABAR (2011): "Safeguarding Banks and Containing Property Booms: Cross-Country Evidence on Macroprudential Policies and Lessons From Hong Kong SAR." IMF Working Paper WP/11/284, International Monetary Fund.

AKINCI, O. AND J. OLMSTEAD-RUMSEY (2015): "How Effective Are Macroprudential Policies? An Empirical Investigation." International Finance Discussion Papers 1135, Board of Governors of the Federal Reserve System.

AKINCI, O. AND J. OLMSTEAD-RUMSEY (2018): "How Effective Are Macroprudential Policies? An Empirical Investigation." *Journal of Financial Intermediation,* 33:33–57.

ALAM, Z., M. A. ALTER, J. EISEMAN, M. R. GELOS, M. H. KANG, M. M. NARITA, E. NIER, AND N. WANG (2019): "Digging Deeper – Evidence on the Effects of Macroprudential Policies From a New Database." IMF Working Paper WP/19/66, International Monetary Fund.

AMINI, S. M. AND C. F. PARMETER (2012): "Comparison of Model Averaging Techniques: Assessing Growth Determinants." *Journal of Applied Econometrics,* 27(5):870–876.

ANDREWS, D. AND F. PETROULAKIS (2019): "Breaking the Shackles: Zombie Firms, Weak Banks and Depressed Restructuring in Europe." Working Paper Series 2240, European Central Bank.

ARAUJO, J., M. PATNAM, A. POPESCU, F. VALENCIA, AND W. YAO (2020): "Effects of Macroprudential Policy: Evidence From Over 6,000 Estimates." IMF Working Paper WP/20/67, International Monetary Fund.

ARMSTRONG, J., H. SKILLING, AND F. YAO (2019): "Loan-to-Value Ratio Restrictions and House Prices: Micro Evidence From New Zealand." *Journal of Housing Economics,* 44: 88–98.

ASTAKHOV, A., T. HAVRANEK, AND J. NOVAK (2019): "Firm Size And Stock Returns: A Quantitative Survey." *Journal of Economic Surveys,* 33(5):1463–1492.

AYYAGARI, M., T. BECK, AND M. S. MARTINEZ PERIA (2017): "Credit Growth and Macroprudential Policies: Preliminary Evidence on the Firm Level." BIS Papers 91a, Bank for International Settlements.

AYYAGARI, M., T. BECK, AND M. M. S. M. PERIA (2018): "The Micro Impact of Macroprudential Policies: Firm-Level Evidence." IMF Working Paper WP/18/267, International Monetary Fund.

BACHMANN, R. AND S. RÜTH (2017): "Systematic Monetary Policy and the Macroeconomic Effects of Shifts in Loan-to-Value Ratios." CESifo Working Paper 6458, CESifo.

BACHMANN, R. AND S. K. RÜTH (2020): "Systematic Monetary Policy and the Macroeconomic Effects of Shifts in Residential Loan-to-Value Ratios." *International Economic Review,* 61 (2):503–530.

BAJZIK, J. (2021): "Trading Volume and Stock Returns: A Meta-Analysis." *International Review of Financial Analysis,* 78:101923.

BAJZIK, J., T. HAVRANEK, Z. IRSOVA, AND J. SCHWARZ (2020): "Estimating the Armington Elasticity: The Importance of Data Choice and Publication Bias." *Journal of International Economics,* 127(C):103383.

BAUM, C. F., M. E. SCHAFFER, AND S. STILLMAN (2003): "Instrumental Variables and GMM: Estimation and Testing." *The Stata Journal,* 3(1):1–31.

BOISSAY, F., F. COLLARD, AND F. SMETS (2016): "Booms and Banking Crises." *Journal of Political Economy,* 124(2):489–538.

BOM, P. R. AND H. RACHINGER (2019): "A Kinked Meta-Regression Model for Publication Bias Correction." *Research Synthesis Methods,* 10(4):497–514.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2018): "Diagnostic Expectations and Credit Cycles." *The Journal of Finance,* 73(1):199–227.

BOZ, E. AND E. G. MENDOZA (2014): "Financial Innovation, the Discovery of Risk, and the US Credit Crisis." *Journal of Monetary Economics,* 62:1–22.

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):103796.

BUDNIK, K. B. (2020): "The Effect of Macroprudential Policies on Credit Developments in Europe 1995-2017." ECB Working Paper No. 2462, European Central Bank.

CAMPOS, N. F., J. FIDRMUC, AND I. KORHONEN (2019): "Business Cycle Synchronisation and Currency Unions: A Review of the Econometric Evidence Using Meta-Analysis." *International Review of Financial Analysis,* 61:274–283.

CARD, D. AND A. B. KRUEGER (1995): "Time-Series Minimum-Wage Studies: A Meta-Analysis." *The American Economic Review,* 85(2):238–243.

CARRERAS, O., E. P. DAVIS, AND R. PIGGOTT (2018): "Assessing Macroprudential Tools in OECD Countries Within a Cointegration Framework." *Journal of Financial Stability,* 37: 112–130.

CATTANEO, M. D., M. JANSSON, AND X. MA (2020): "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association,* 115(531):1449–1455.

CAZACHEVICI, A., T. HAVRANEK, AND R. HORVATH (2020): "Remittances and Economic Growth: A Meta-Analysis." *World Development,* 134:105021.

Cerutti, E., S. Claessens, and L. Laeven (2016): "The Use and Effectiveness of Macroprudential Policies." BIS Paper No. 86, Bank of International Settlements.

Cerutti, E., S. Claessens, and L. Laeven (2017): "The Use and Effectiveness of Macroprudential Policies: New Evidence." *Journal of Financial Stability,* 28:203–224.

Cerutti, E., R. Correa, E. Fiorentino, and E. Segalla (2017): "Changes in Prudential Policy Instruments—A New Cross-Country Database." *International Journal of Central Banking,* 13(1):477–503.

Cesnak, M., J. Klacso, and R. Vasil (2021): "Analysis of the Impact of Borrower-Based Measures." NBS Occasional Paper 3/2021, National Bank of Slovakia.

Claessens, S., S. R. Ghosh, and R. Mihet (2013): "Macro-Prudential Policies to Mitigate Financial System Vulnerabilities." *Journal of International Money and Finance,* 39:153–185.

de Araujo, D. K. G., J. B. R. Barroso, and R. B. Gonzalez (2017): "Loan-to-Value Policy and Housing Finance: Effects on Constrained Borrowers." BIS Working Papers 673, Bank for International Settlements.

de Araujo, D. K. G., J. B. R. B. Barroso, and R. B. Gonzalez (2020): "Loan-to-Value Policy and Housing Finance: Effects on Constrained Borrowers." *Journal of Financial Intermediation,* 42:100830.

Doucouliagos, C. (2005): "Publication Bias in the Economic Freedom and Economic Growth Literature." *Journal of Economic Surveys,* 19(3):367–387.

Doucouliagos, C. and P. Laroche (2003): "What Do Unions Do to Productivity? A Meta-Analysis." *Industrial Relations: A Journal of Economy and Society,* 42(4):650–691.

Egger, M., G. D. Smith, and C. Minder (1997): "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *Journal of Economic Surveys,* 316:629–634.

Ehrenbergerova, D., J. Bajzik, and T. Havranek (2021): "When Does Monetary Policy Sway House Prices? A Meta-Analysis." *IMF Economic Review,* forthcoming.

Eicher, T. S., C. Papageorgiou, and A. E. Raftery (2011): "Default Priors and Predictive Performance in Bayesian Model Averaging, With Application to Growth Determinants." *Journal of Applied Econometrics,* 26(1):30–55.

Elliott, G., N. Kudrin, and K. Wüthrich (2022): "Detecting p-Hacking." *Econometrica,* 90(2):887–906.

Epure, M., I. Mihai, C. Minoiu, and J.-L. Peydró (2018): "Household Credit, Global Financial Cycle, and Macroprudential Policies: Credit Register Evidence From an Emerging Country." Working Papers Series 1006, Universitat Pompeu Fabra.

Feldkircher, M. and S. Zeugner (2012): "The Impact of Data Revisions on the Robustness of Growth Determinants — A Note on 'Determinants of Economic Growth: Will Data Tell?'." *Journal of Applied Econometrics,* 27(4):686–694.

Fernandez, C., E. Ley, and M. F. J. Steel (2001): "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics,* 16(5):563–576.

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

FURUKAWA, C. (2019): "Publication Bias Under Aggregation Frictions: Theory, Evidence, and a New Correction Method." unpublished manuscript.

FUSTER, A., D. LAIBSON, AND B. MENDEL (2010): "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives,* 24(4):67–84.

GADATSCH, N., L. MANN, AND I. SCHNABEL (2017): "A New IV Approach for Estimating the Efficacy of Macroprudential Measures." Arbeitspapier 05/2017, German Council of Economic Experts.

GADATSCH, N., L. MANN, AND I. SCHNABEL (2018): "A New IV Approach for Estimating the Efficacy of Macroprudential Measures." *Economics Letters,* 168:107–109.

GALATI, G. AND R. MOESSNER (2013): "Macroprudential Policy – A Literature Review." *Journal of Economic Surveys,* 27(5):846–878.

GALATI, G. AND R. MOESSNER (2018): "What Do We Know About the Effects of Macroprudential Policy?" *Economica,* 85(340):735–770.

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

GEANAKOPLOS, J. (2010): "The Leverage Cycle." *NBER Macroeconomics Annual,* 24(1):1–66.

GECHERT, S., T. HAVRANEK, Z. IRSOVA, AND D. KOLCUNOVA (2022): "Measuring Capital-Labor Substitution: The Importance of Method Choices and Publication Bias." *Review of Economic Dynamics,* 45:55–82.

GENNAIOLI, N., A. SHLEIFER, AND R. VISHNY (2015): "Neglected Risks: The Psychology of Financial Crises." *American Economic Review,* 105(5):310–14.

GEORGE, E. I. (2010): "Dilution Priors: Compensating for Model Space Redundancy." Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown. IMS Collections, Vol. 6., Institute of Mathematical Statistics.

GERBER, A. AND N. MALHOTRA (2008): "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science,* 3(3):313–326.

GERBER, A. S. AND N. MALHOTRA (2008): "Publication Bias in Empirical Sociological Research: Do Arbitrary Significance Levels Distort Published Results?" *Sociological Methods & Research,* 37(1):3–30.

GERSL, A. AND M. JASOVA (2012): "Measures to Tame Credit Growth: Are They Effective?" Working Papers IES 2012/28, Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies.

GERŠL, A. AND M. JAŠOVÁ (2014): "Measures to Tame Credit Growth: Are They Effective?" *Economic Systems,* 38(1):7–25.

GRIC, Z., J. BAJZIK, O. BADURA, ET AL. (2021): "Does Sentiment Affect Stock Returns? A Meta-Analysis Across Survey-Based Measures." CNB Working Paper No. 10/2021, Czech National Bank.

GRODECKA, A. (2020): "On the Effectiveness of Loan-to-Value Regulation in a Multiconstraint Framework." *Journal of Money, Credit and Banking,* 52(5):1231–1270.

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

HAVRANEK, T. AND Z. IRSOVA (2010): "Meta-Analysis of Intra-Industry FDI Spillovers: Updated Evidence." *Czech Journal of Economics and Finance,* 60(2):151–174.

HAVRANEK, T. AND M. RUSNAK (2012): "Transmission Lags of Monetary Policy: A Meta-Analysis." IES Working Paper, No. 27/2012, Charles University in Prague, Institute of Economic Studies (IES), Prague.

HAVRANEK, T., R. HORVATH, AND A. ZEYNALOV (2016): "Natural Resources and Economic Growth: A Meta-Analysis." *World Development,* 88:134–151.

HAVRÁNEK, T., T. STANLEY, H. DOUCOULIAGOS, P. BOM, J. GEYER-KLINGEBERG, I. IWASAKI, W. R. REED, K. ROST, AND R. VAN AERT (2020): "Reporting Guidelines for Meta-Analysis in Economics." *Journal of Economic Surveys,* 34(3):469–475.

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

HODULA, M., M. MELECKÝ, L. PFEIFER, AND M. SZABO (2021): "Cooling Down the Mortgage Loan Market: The Effect of Recommended Borrower-Based Limits on New Mortgage Lending." CNB Working Paper No. 3/2022, Czech National Bank.

IGAN, M. D. AND M. H. KANG (2011): "Do Loan-to-Value and Debt-to-Income Limits Work? Evidence from Korea." IMF Working Paper WP/11/297, International Monetary Fund.

IOANNIDIS, J. P., T. STANLEY, AND H. DOUCOULIAGOS (2017): "The Power of Bias in Economics Research." *Economic Journal,* 127(605):236–265.

JÁCOME, L. I. AND S. MITRA (2015): "LTV and DTI Limits – Going Granular." IMF Working Paper WP/15/154, International Monetary Fund.

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

JURČA, P., J. KLACSO, E. TEREANU, M. FORLETTA, AND M. GROSS (2020): "The Effectiveness of Borrower-Based Macroprudential Measures: A Quantitative Analysis for Slovakia." IMF Working Paper No. 20/134, International Monetary Fund.

KIM, J., H. DOUCOULIAGOS, AND T. STANLEY (2019): *Market Efficiency in Asian and Australasian Stock Markets: A Fresh Look at the Evidence.* In *International Financial Markets,* pages 382–419. Routledge.

KRONICK, J. (2015): "Do Loan-to-Value Ratio Regulation Changes Affect Canadian Mortgage Credit?" MPRA Working Papers 73761, Munich Personal RePEc Archive.

KRZNAR, M. I. AND M. J. MORSINK (2014): "With Great Power Comes Great Responsibility: Macroprudential Tools at Work in Canada." IMF Working Paper WP/14/083, International Monetary Fund.

KUTTNER, K. N. AND I. SHIM (2013): "Can Non-Interest Rate Policies Stabilize Housing Markets? Evidence From a Panel of 57 Economies." BIS Working Papers 433, Bank for International Settlements.

KUTTNER, K. N. AND I. SHIM (2016): "Can Non-Interest Rate Policies Stabilize Housing Markets? Evidence From a Panel of 57 Economies." *Journal of Financial Stability,* 26: 31–44.

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

LIM, C. H., A. COSTA, F. COLUMBA, P. KONGSAMUT, A. OTANI, M. SAIYID, T. WEZEL, AND X. WU (2011): “Macroprudential Policy: What Instruments and How to Use Them? Lessons from Country Experiences.” IMF Working Paper WP/11/238, International Monetary Fund.

MALOVANÁ, S. AND J. FRAIT (2017): “Monetary Policy and Macroprudential Policy: Rivals or Teammates?” *Journal of Financial Stability,* 32:1–16.

MALOVANÁ, S., M. HODULA, AND Z. RAKOVSKÁ (2020): “Researching the Research: A Central Banking Edition.” CNB Research and Policy Note 3/2020, Czech National Bank.

MALOVANÁ, S., J. BAJZÍK, Z. GRIC, AND M. HODULA (2021): “A Tale of Different Capital Ratios: How To Correctly Assess The Impact of Capital Regulation on Lending.” CNB Working Paper No. 8/2021, Czech National Bank.

MCDONALD, C. (2015): “When is Macroprudential Policy Effective?” BIS Working Papers 496, Bank for International Settlements.

MENDOZA, E. G. AND M. E. TERRONES (2008): “An Anatomy of Credit Booms: Evidence from Macro Aggregates and Micro Data.” Working Paper 14049, National Bureau of Economic Research.

MORGAN, P., P. REGIS, AND N. SALIKE (2015): “Loan-to-Value Policy as a Macroprudential Tool: The Case of Residential Mortgage Loans in Asia.” ADBI Working Paper 528, ADBInstitute.

MORGAN, P. J., P. J. REGIS, AND N. SALIKE (2019): “LTV Policy as a Macroprudential Tool and Its Effects on Residential Mortgage Loans.” *Journal of Financial Intermediation,* 37: 89–103.

NANSEN MCCLOSKEY, D. AND 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 Education,* 50(4):356–361.

NEAGU, F., L. TATARICI, AND I. MIHAI (2015): “Implementing Loan-to-Value and Debt Service-To-Income Measures: A Decade of Romanian Experience.” Occasional Papers 15, Bank of Romania.

POGHOSYAN, T. (2019): “How Effective is Macroprudential Policy? Evidence From Lending Restriction Measures in EU Countries.” IMF Working Papers WP/19/045, International Monetary Fund.

POGHOSYAN, T. (2020): “How Effective is Macroprudential Policy? Evidence From Lending Restriction Measures in EU Countries.” *Journal of Housing Economics,* 49:101694.

RICHTER, B., M. SCHULARICK, AND I. SHIM (2018): “The Macroeconomic Effects of Macroprudential Policy.” BIS Working Papers 740, Bank for International Settlements.

RICHTER, B., M. SCHULARICK, AND I. SHIM (2019): “The Costs of Macroprudential Policy.” *Journal of International Economics,* 118:263–282.

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

SIMONSOHN, U., L. D. NELSON, AND J. P. SIMMONS (2014): "P-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results." *Perspectives on Psychological Science,* 9(6):666–681.

SIMONSOHN, U., L. D. NELSON, AND J. P. SIMMONS (2014): "P-Curve: A Key to the File-Drawer." *Journal of Experimental Psychology: General,* 143(2):534–547.

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

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

STANLEY, T. D., H. DOUCOULIAGOS, M. GILES, J. H. HECKEMEYER, R. J. JOHNSTON, P. LAROCHE, J. P. NELSON, M. PALDAM, J. POOT, AND G. PUGH (2013): "Meta-Analysis of Economics Research Reporting Guidelines." *Journal of Economic Surveys,* 27 (2):390–394.

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business & Economic Statistics,* 20(4):518–529.

TANTASITH, C., N. ANANCHOTIKUL, C. CHOTANAKARN, V. LIMJAROENRAT, AND R. PONGSAPARN (2020): "The Impact of LTV Policy on Bank Lending: Evidence From Thailand." BIS Working Papers 110, Bank for International Settlements.

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

VAN BEKKUM, S., M. GABARRO, R. M. IRANI, AND J.-L. PEYDRÓ (2019): "Take It to the Limit? The Effects of Household Leverage Caps." Economic Working Paper Series No. 1682, Universitat Pompeu Fabra, Department of Economics and Business.

WANG, M. B. AND T. SUN (2013): "How Effective are Macroprudential Policies in China?" IMF Working Paper WP/13/075, International Monetary Fund.

ZEUGNER, S. AND M. FELDKIRCHER (2015): "Bayesian Model Averaging Employing Fixed and Flexible Priors: The BMS Package for R." *Journal of Statistical Software,* 68(4):1–37.

ZHANG, L. AND E. ZOLI (2016a): "Leaning Against the Wind: Macroprudential Policy in Asia." IMF Working Paper Series WP/14/022, International Monetary Fund.

ZHANG, L. AND E. ZOLI (2016b): "Leaning Against the Wind: Macroprudential Policy in Asia." *Journal of Asian Economics,* 42:33–52.

ZHANG, Y. AND T. TRESSEL (2017): "Effectiveness and Channels of Macroprudential Policies: Lessons From the Euro Area." *Journal of Financial Regulation and Compliance.*

ZIGRAIOVA, D. AND T. HAVRANEK (2016): "Bank Competition and Financial Stability: Much Ado About Nothing?" *Journal of Economic Surveys,* 30(5):944–981.

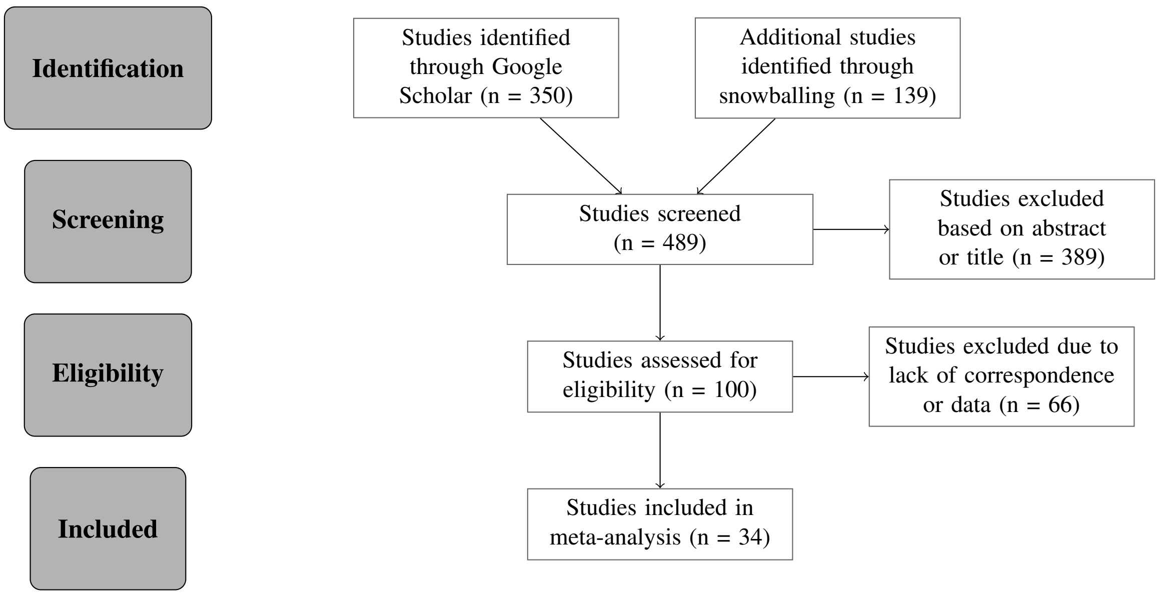
ZIGRAIOVA, D., T. HAVRANEK, Z. IRSOVA, AND J. NOVAK (2021): "How Puzzling is the Forward Premium Puzzle? A Meta-Analysis." *European Economic Review,* 134:103714.

Appendix A: Additional Charts

A.1 PRISMA Diagram

Figure A1 depicts the overall process employed in the selection of primary studies. As this article is a sister article to Malovaná et al. (2021), the selection process is quite similar in both cases. In the *identification* phase, we surveyed the first 350 research articles returned by Google Scholar given a tailor-made search query. The query was limited to papers published in or after 2010 (see Section 2.1). Next, we snowballed all the citations in each of the relevant studies. In this way we identified an additional 139 articles. Thus, in total we *screened* 489 articles. In the next step, we scrutinized all the titles and abstracts and rejected all the studies that were not acceptable even from a high-level perspective. We thereby eliminated 389 studies. The remaining 100 were assessed for *eligibility*. During this step, we investigated each of these articles in detail and dropped 66 due to a lack of correspondence or data. The main elimination criteria were: (1) the study must report numerical results; (2) the estimated elasticities must be presented together with corresponding test statistics – standard errors, t-statistics, p-values or exact confidence intervals; (3) the effect cannot be a cross-border effect; and finally, (4) the measure of lending cannot be expressed as a ratio to some other continuous variable such as total loans or total bank assets. We ended up with 34 primary studies *included* in the meta-analysis.

Figure A1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Diagram

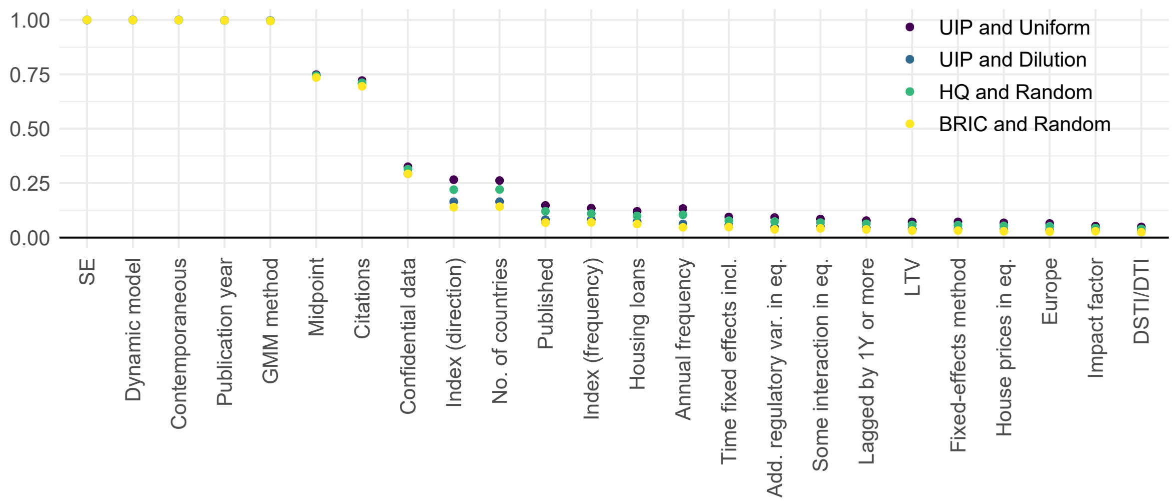


A.2 Variation of Collected Estimates

Figure A2: Reported Estimates Vary Both Within and Across Studies

Note: 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 vertical line denotes unitary elasticity. LTV stands for loan-to-value ratio; DSTI/DTI stands for debt service-to-income ratio; Both stands for measures encompassing multiple borrower-based measures.

Figure A3: Bayesian Model Averaging – Prior Sensitivity



Note: The figures show the 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 the 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 the random model prior and a combination of the BRIC g-prior and the 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).

Appendix B: Extensions to the Publication Bias Tests

We report three additional tests as an extension to the analysis of publication bias presented in Section 3: the caliper test proposed by Gerber and Malhotra (2008a,b) and recently adjusted by Bruns et al. (2019), the p-curve developed by Simonsohn et al. (2014a,b), and a set of statistical tests used by Elliott et al. (2022).

B.1 The Caliper Test

Caliper tests are based on the intuition that the reported t-statistics should be evenly distributed around the conventional significance thresholds (e.g., 1.65 for 10% significance and 1.96 for 5% significance). In other words, the number of reported t-statistics above a threshold ("over caliper") should not be statistically different from the number of reported t-statistics below the threshold ("under caliper") and the resulting ratio should be equal to or lower than 0.5 (50:50). Bruns et al. (2019) proposes to follow a more lenient rule when the over-to-under caliper ratio is 0.4 (60:40). They claim that the original 50:50 null hypothesis is too conservative, as large-scale evidence suggests that economic research is frequently under-powered (Ioannidis et al., 2017). Hence, the frequency of the t-statistic values is likely to decrease with the magnitude of the t-statistic values if reporting biases are absent.

The results are presented in Table B1. We inspect two significance thresholds, corresponding to the 90% and 95% confidence intervals. We use three caliper sizes (0.1, 0.2, and 0.3). The results show that publication selection is present for negative estimates where we reject the null of no p-hacking for both the conservative threshold of 0.5 and the more lenient threshold of 0.4. We also obtain weak evidence of p-hacking for positive estimates at the 10% level of significance, which conforms to our baseline finding that researchers do not publish positive estimates, since they go against economic logic.

Table B1: The Caliper Test Confirms P-Hacking

T-stat	C	All		Negative		Positive	
1.96	0.1	*0.574*	*(0.452)*	**0.641**	**(0.510)**	0.250	(-0.060)
	0.2	0.444	(0.361)	*0.506*	*(0.413)*	0.167	(0.009)
	0.3	0.400	(0.332)	*0.491*	*(0.412)*	0.091	(0.005)
1.65	0.1	**0.686**	**(0.576)**	**0.658**	**(0.526)**	**0.769**	**(0.552)**
	0.2	*0.505*	*(0.424)*	**0.606**	**(0.505)**	0.341	(0.215)
	0.3	*0.514*	*(0.444)*	**0.600**	**(0.516)**	0.340	(0.223)

Note: The table shows the results of the caliper test for three caliper sizes 0.1, 0.2, and 0.3. The reported numbers represent the share of observations in the narrow interval above the significance threshold, i.e., the share of observations above 1.96 or 1.65. Formally, the ratio C is calculated as the number of observations above the given significance threshold ("over caliper") over the total number of observations. We test two one-sided null hypotheses of no p-hacking: C is lower than or equal to 0.5 and C is lower than or equal to 0.4. Significant results are shown in bold (H0: $C \leq 0.5$) and italics (H0: $C \leq 0.4$). Lower 95% confidence intervals are reported in brackets.

B.2 Tests Based on the Distribution of P-Values

Next, we analyze the distribution of p-values. First, we employ the p-curve method, which tests the null hypothesis that there is no evidential value, i.e., the introduction or tightening of borrower-based measures has no effect on bank lending beyond publication bias. In other words, the null hypothesis suggests a flat distribution of p-values. If the distribution is right-skewed, it implies a non-zero effect, with its size and strength presented on the y-axis of the chart. On the other hand, a left-skewed distribution would suggest that p-hacking is present at the given significance level.

The results are presented in Table B1. The figure shows that the distribution of p-values is visibly left-skewed, suggesting the presence of p-hacking. Looking at the individual estimates in terms of sign, we find that the p-hacking mainly concerns negative estimates. This is supported by the distribution of p-values plotted in Figure B2. Overall, the p-value distribution confirms the existence of a strong evidential value, that is, researchers have strong prior expectations about the direction of effect of the studies. In our case, researchers strongly expect to find negative effects of borrower-based measures on bank lending.

Second, we employ a series of tests used by Elliott et al. (2022). Elliott et al. (2022) test for the presence of p-hacking by looking at the distributions of p-values across multiple studies and identify novel additional testable restrictions for p-values based on t-tests. Specifically, they show that the p-curves based on t-tests are completely monotone in the absence of p-hacking, and their magnitude and the magnitude of their derivatives are restricted by upper bounds. Following Simonsohn et al. (2014b) and Cattaneo et al. (2020), the tests apply the binomial, Fisher, and density discontinuity approaches. But in addition to them, Elliott et al. (2022) proposed several new tests with more statistical power. The first is a histogram-based test for non-increasingness. The second is a histogram-based test for 2-monotonicity and bounds. And the third and last one is the least concave majorant (LCM) test, relying on concavity of the CDF of p-values.

The results are presented in Table B2. All of the tests have the null hypothesis of no p-hacking. We do not reject the null with less powerful tests (binomial and Fisher), but we can reject it with the tests for discontinuity, non-increasingness (CS1), and 2-monotonicity (CS2B). The tests imply the presence of p-hacking mainly for positive estimates.

Table B2: Tests Used by Elliott et al. (2022)

Test	All estimates	Negative estimates	Positive estimates
Binomial	1.000	1.000	1.000
Fisher	1.000	1.000	0.570
Discontinuity	**0.003**	**0.010**	**0.008**
CS1	0.170	0.349	**0.000**
CS2B	**0.040**	0.558	**0.000**
LCM	0.821	0.995	0.160
N	422	284	138

Note: The table presents the p-values of six different statistical tests with the null hypothesis of no p-hacking. CS1 is the test for non-increasingness. CS2B is the test for K-monotonicity. LCM is the test based on the concavity of the CDF of p-values. We run the tests at a threshold of t = 1.62, corresponding to 10% statistical significance. Values in bold indicate rejections of the hypothesis of no p-hacking at the 10% significance level.

Figure B1: P-Curve

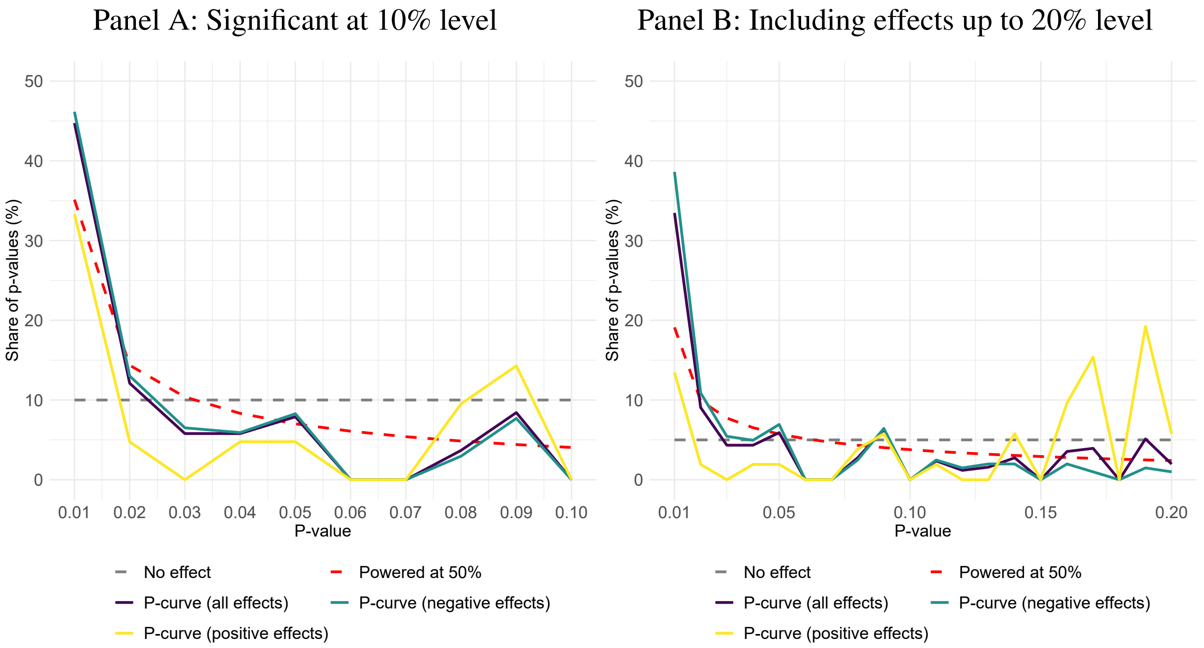
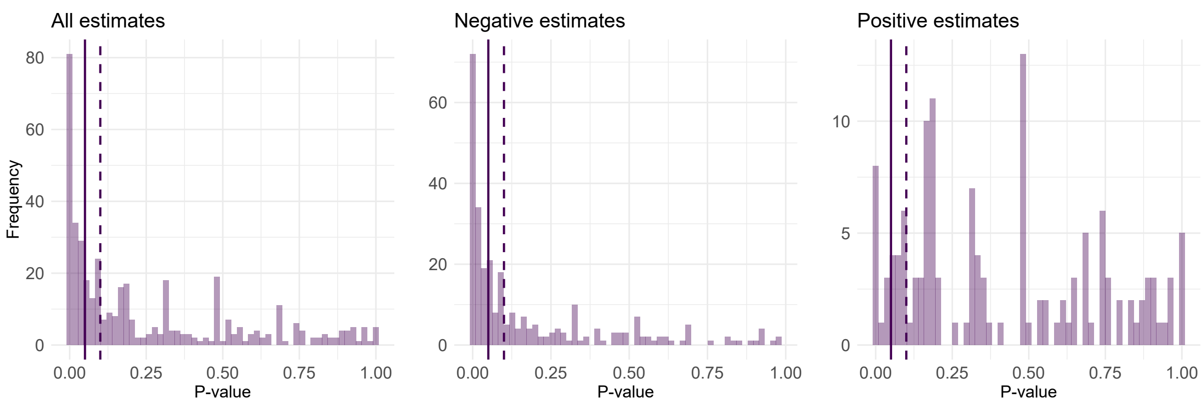


Figure B2: Distribution of P-Values



Appendix C: Additional Results Based on the Full Sample of Collected Estimates Transformed to Partial Correlation Coefficients

As different studies use different units of measurement, we use the partial correlation coefficient (PCC) method to make the estimates directly comparable. This transformation comes at the cost of losing some information, namely, the economic interpretability of the estimated parameters, but it increases the number of observations we can work with from 422 to 722. To this end, we use PCC transformation to verify the main conclusions presented in the main text.

First, we test for the presence of publication selection using linear estimation methods (see Tables C1 and C2). The estimates confirm the publication selection preference for estimates that are just significant at the 10% level and negative. Using a visual test in the form of a funnel plot, we next plot the distribution of PCCs and the associated t-statistics (Figure C1). The funnel plot of PCCs is visibly left-skewed, confirming the presence of negative publication bias. The funnel plot of t-statistics lends additional support to this view and adds that positive estimates, if published, are mostly not statistically significant. This conforms to our claim that publication selection exists and is driven by both selection of sign and statistical significance.

Table C1: Publication Bias – Linear Methods

	(1) OLS	(2) Study	(3) Precision	(4) FE	(5) Bayes
Constant (effect beyond bias)	-0.0066	-0.0291**	-0.0033***	-0.028	-0.0195
	(0.0142)	(0.0116)	(0.0012)	(0.0217)	(0.0686)
SE (publication bias)	0.0943	0.1672	-0.0041	0.2232	-0.2389
	(0.2097)	(0.257)	(0.242)	(0.3983)	(0.3971)
Observations	722	722	722	722	722
Studies	34	34	34	34	34

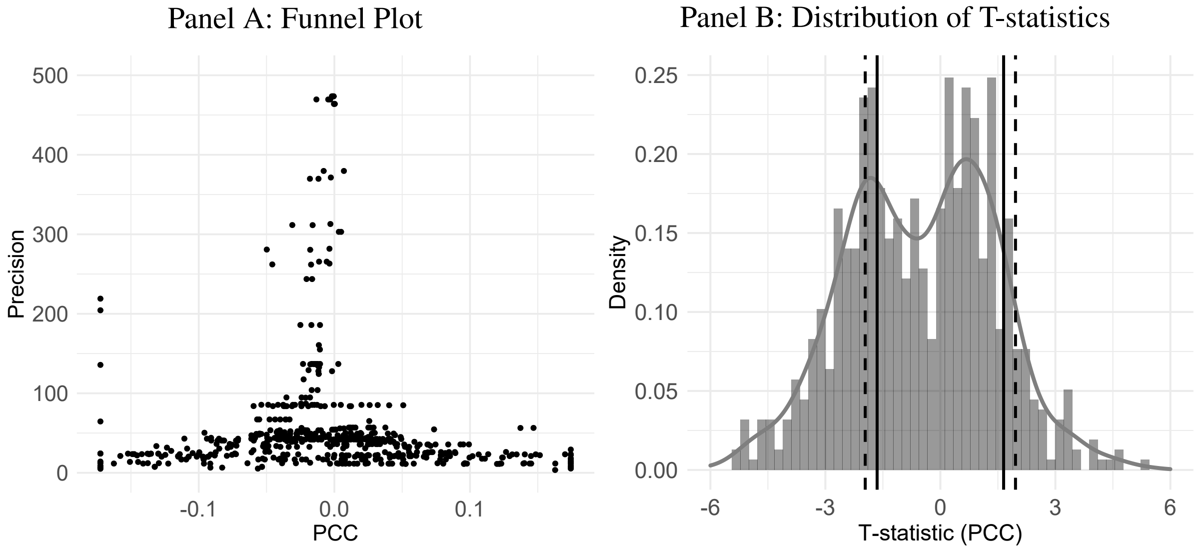
Note: The table presents the results of the regression of equation (2) using all the collected estimates transformed to PCCs. The standard errors, reported in parentheses, are clustered at both the level of the study and the type of borrower-based measure used in the primary study. OLS – ordinary least squares. Study – the inverse of the number of estimates reported per study is used as the weight. Precision – the inverse of the reported estimate's standard error is used as the weight. FE – study-level fixed effects. RE – study-level random effects. Bayes – hierarchical Bayesian approach. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C2: Publication Bias – Linear Methods and Interaction Terms

	(1) OLS	(2) Study	(3) Precision	(4) FE	(5) Bayes
Panel A: significant at 10% level					
Constant	0.0070	-0.0028	-0.0004	-0.0279	-0.0013
	(0.011)	(0.0124)	(0.0013)	(0.0246)	(0.0877)
SE	-0.0615	-0.0237	0.1213**	0.1366	-0.1729
	(0.2862)	(0.3137)	(0.0608)	(0.3393)	(0.391)
I(t-stat<1.65)	-0.0311**	-0.0476**	-0.0036	-0.0267**	-0.0209
	(0.0133)	(0.019)	(0.0028)	(0.0107)	(0.1037)
SE×I(t-stat<1.65)	0.4767	0.4236	-0.5025	1.1028*	-0.3291
	(0.8936)	(0.6414)	(0.8011)	(0.6049)	(0.7106)
Observations	722	722	722	722	722
Studies	34	34	34	34	34
Panel B: significant at 10% level and negative					
Constant	0.0187	0.0157	0.0000	0.0072	0.0054
	(0.0152)	(0.0133)	(0.0016)	(0.0229)	(0.0845)
SE	0.0425	-0.0130	0.5102***	0.0908	0.1184
	(0.2748)	(0.3099)	(0.1062)	(0.3758)	(0.399)
I(t-stat<1.65, β<0)	-0.0289*	-0.0422**	-0.0013	-0.0119	-0.0264
	(0.0163)	(0.0192)	(0.002)	(0.0083)	(0.109)
SE×I(t-stat<1.65, β<0)	-2.2258***	-1.8356***	-3.1341***	-2.265***	-2.0722***
	(0.4098)	(0.5739)	(0.2377)	(0.2802)	(0.5753)
Observations	722	722	722	722	722
Studies	34	34	34	34	34

Note: The table presents the results of the regression of equation (2) using all the collected estimates transformed to PCCs and extended by additional dummy variables for collected elasticities significant at the 10% level (I(t-stat<1.65)) and elasticities that are negative at the same time (I(t-stat<1.65, β<0)). The standard errors, reported in parentheses, are clustered at both the level of the study and the type of borrower-based measure used in the primary study. OLS – ordinary least squares. Study – the inverse of the number of estimates reported per study is used as the weight. Precision – the inverse of the reported estimate's standard error is used as the weight. FE – study-level fixed effects. RE – study-level random effects. Bayes – hierarchical Bayesian approach. *** p < 0.01, ** p < 0.05, * p < 0.1.

Figure C1: Publication Bias – Funnel Plot and Distribution of T-Statistics



Note: Panel A: Precision is calculated as the inverse of the standard error of the PCC. In the absence of publication bias, the funnel should be symmetric around the most precise PCCs. We exclude PCCs with extreme magnitude or precision from the figure but include all in the regressions. Panel B: The vertical lines denote the critical values associated with 5% (dashed line) and 10% (full line) statistical significance. We exclude PCCs with large t-statistics from the figure but include all in the regressions. In the absence of publication bias, the distribution of the t-statistics should be approximately normal.

CNB Working Paper Series (since 2021)

WP 8/2022	Simona Malovaná Martin Hodula Zuzana Gric Josef Bajzík	*Borrower-based macroprudential measures and credit growth: How biased is the existing literature?*
WP 7/2022	Martin Časta	*How credit improves the exchange rate forecast*
WP 6/2022	Milan Szabo	*Meeting investor outflows in Czech bond and equity funds: Horizontal or vertical?*
WP 5/2022	Róbert Ambriško	*Nowcasting macroeconomic variables using high-frequency fiscal data*
WP 4/2022	Jaromír Baxa Jan Žáček	*Monetary policy and the financial cycle: International evidence*
WP 3/2022	Martin Hodula Milan Szabo Lukáš Pfeifer Martin Melecký	*Cooling the mortgage loan market: The effect of recommended borrower-based limits on new mortgage lending*
WP 2/2022	Martin Veselý	*Application of quantum computers in foreign exchange reserves management*
WP 1/2022	Vojtěch Molnár	*Price level targeting with imperfect rationality: A heuristic approach*
WP 10/2021	Zuzana Gric Josef Bajzík Ondřej Badura	*Does sentiment affect stock returns? A meta-analysis across survey-based measures*
WP 9/2021	Jan Janků Ondřej Badura	*Non-linear effects of market concentration on the underwriting profitability of the non-life insurance sector in Europe*
WP 8/2021	Simona Malovaná Martin Hodula Josef Bajzík Zuzana Gric	*A tale of different capital ratios: How to correctly assess the impact of capital regulation on lending*
WP 7/2021	Volha Audzei	*Learning and cross-country correlations in a multi-country DSGE model*
WP 6/2021	Jin Cao Valeriya Dinger Tomás Gómez Zuzana Gric Martin Hodula Alejandro Jara Ragnar Juelsrud Karolis Liaudinskas Simona Malovaná Yaz Terajima	*Monetary policy spillover to small open economies: Is the transmission different under low interest rates?*
WP 5/2021	Martin Hodula Ngoc Anh Ngo	*Does macroprudential policy leak? Evidence from non-bank credit intermediation in EU countries*

WP 4/2021	Tomáš Adam Ondřej Michálek Aleš Michl Eva Slezáková	*The Rushin index: A weekly indicator of Czech economic activity*
WP 3/2021	Michal Franta Jan Libich	*Holding the economy by the tail: Analysis of short- and long-run macroeconomic risks*
WP 2/2021	Jakub Grossmann	*The effects of minimum wage increases in the Czech Republic*
WP 1/2021	Martin Časta	*Deriving equity risk premium using dividend futures*

CNB Research and Policy Notes (since 2021)

RPN 4/2021	Simona Malovaná Martin Hodula Zuzana Gric Josef Bajzík	*Macroprudential policy in central banks: Integrated or separate? Survey among academics and central bankers*
RPN 3/2021	Martin Hodula Jan Janků Lukáš Pfeifer	*Interaction of cyclical and structural systemic risks: Insights from around and after the global financial crisis*
RPN 2/2021	Karel Musil Stanislav Tvrz Jan Vlček	*News versus surprise in structural forecasting models: Central bankers' practical perspective*
RPN 1/2021	Miroslav Plašil	*Designing macro-financial scenarios: The New CNB framework and satellite models for property prices and credit*

CZECH NATIONAL BANK
Na Příkopě 28
115 03 Praha 1
Czech Republic

ECONOMIC RESEARCH DIVISION
Tel.: +420 224 412 321
Fax: +420 224 412 329
http://www.cnb.cz
e-mail: research@cnb.cz

ISSN 1803-7070