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

Simona Malovaná[a], Martin Hodula[b], Zuzana Gric[d], Josef Bajzík[c]

[a] *Czech National Bank and Prague University of Economics and Business*
[b] *Czech National Bank and Technical University of Ostrava*
[c] *Czech National Bank and Charles University in Prague*
[d] *Czech National Bank and Masaryk University in Brno*

Abstract

The ever-so-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 of the policy and its effectiveness. 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 a 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 battery of empirical tests, we verify the presence of a strong publication bias, especially against positive and not statistically 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 the underlying data characteristics.

Keywords: Bayesian model averaging, borrower-based measures, macroprudential policy, meta-analysis, publication bias
JEL Classification: C83, E58, G21, G28, G51

Email addresses: simona.malovana@cnb.cz (Simona Malovaná), martin.hodula@cnb.cz (Martin Hodula), zuzana.gric@cnb.cz (Zuzana Gric), josef.bajzik@cnb.cz (Josef Bajzík)

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 of 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 the borrower-based measures taking their place amongst the most popular policy tools both in advanced and emerging market economies (Cerutti et al., 2017a; Alam et al., 2019). The IMF iMaPP database covers 134 countries and it records 87 borrower-based policy actions over the 1990-2007 period and 316 policy actions over the 2008–2020 period.[2] Borrower-based measures, such as loan-to-value, debt-to-income, and debt service-to-income limits, are applied on the level of individual loans, and therefore can directly restrict the amount of credit available to the private sector. The limits are meant to increase borrowers' and lenders' resilience, lowering potential bank losses during economic and financial downturns.

Despite the growing use of borrower-based measures, there is 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 the 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 boom typically precede crises (Mendoza and Terrones, 2008; Jordà et al., 2011; Schularick and Taylor, 2012).[3] Placing the bank credit in the centre of policy and academic analyses is further supported by the rich evolving literature on the interaction between monetary and macroprudential policy. While the policies have different aims, both of them have direct or indirect impact on bank credit (Galati and Moessner, 2018) which can, in certain situations, be at odds (Malovana and Frait, 2017).

The fragmentation is due in part to a reliance on dummy-type policy action indices, which prevents from estimating quantitative effects of the policy. They assign a value of one when a macroprudential policy tool is implemented or changed and zero otherwise. In addition, one can suspect non-negligible endogeneity problems to hamper a proper assessment of macroprudential policy effects. The endogeneity bias can stem from the fact

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

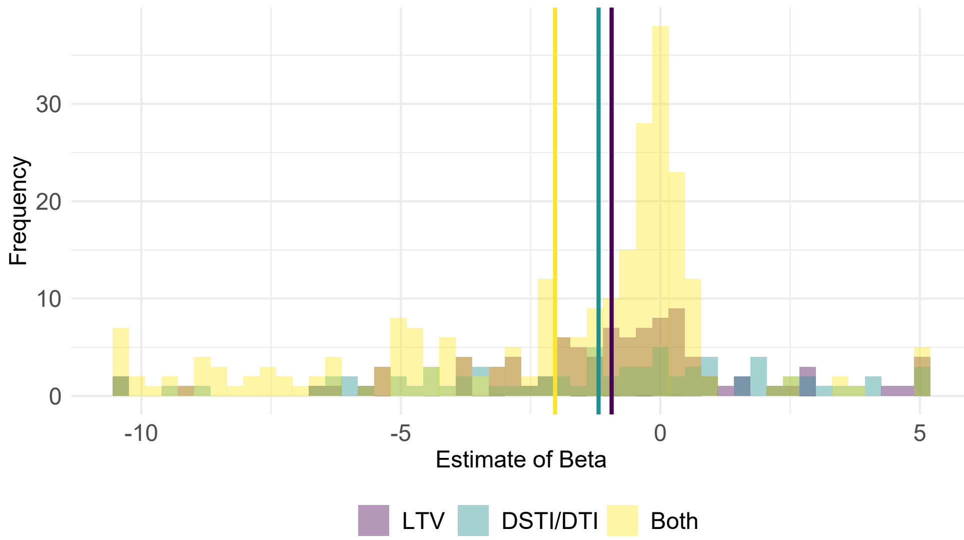
[3]See also Fuster et al. (2010), Geanakoplos (2010) ,Boz and Mendoza (2014), Gennaioli et al. (2015), Boissay et al. (2016) and Bordalo et al. (2018) for stylized models that try to capture the boom-bust cycle in credit associated with financial crises.

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 effects of using borrower-based measures individually and in combination with one another. For instance, there is emerging literature that provides evidence of differential impact 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 the impact 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, 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 the borrower-based measures are becoming a standard policy tool to address imbalances in the residential mortgage loan market, the estimated average (general) effect of individual tools could be of value to policy practitioners when deciding over the regulatory policy placement. However, beyond average effects, the policy makers, 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 to describe what the most commonly employed empirical strategy is. We then use state-of-the-art meta-analytic techniques to estimate the true effect of imposing borrower-based measures on bank lending, as well as the model averaging methods used to identify the significant drivers of the heterogeneity of the observed estimates.

The meta-analysis employed in the paper is well suited to tackle the heterogeneity of surveyed estimates. If the subject of the meta-analysis is deep/structural parameter in a correct model (i.e. the model that correctly describes the data generating process), then the heterogeneity of estimates should be fairly low and given by econometric issues only (estimation method, variables transformation, etc.). If the subject of the meta-analysis is a reduced-form parameter of a correct model, then higher heterogeneity of estimates should be expected because different policy regimes imply different value of the parameter. So, the heterogeneity of surveyed estimates should be explained also by the 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 of the impact of borrower-based measures on bank credit, the heterogeneity of estimates would be even more profound and tightly related to model specification. All the above-mentioned reasons for heterogeneity apply for the exercises presented in the paper.

Figure 1: The Mean Effect of Borrower-based Measures on Annual Credit Growth Is Negative



Note: The figure showcases a histogram of the estimated beta for all collected estimates on the relationship between bank annual credit growth and borrower-based measures expressed as index or dummy variable (see equation 1). The beta gives percentage point change in annual credit growth in response to one-unit increase in the borrower-based index or dummy. LTV stands for loan-to-value ratio; DSTI/DTI stands for debt service-to-income ratio or debt-to-income ratio; both stands for measures encompassing multiple borrower based measures. The solid vertical lines indicate the mean of winsorized estimates; 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 the regulatory change, the identified effect does not refer to the intensity or binding nature of the measure taken, but only to its introduction or change. Second, the mean effect of imposing value-based individually (LTV) and in combination with income-based (DTI, DSTI) limits on bank lending is observably different in terms of size of the reported coefficient. 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 the LTV limit. We find the application of DTI and DSTI limits to reduce the credit growth more than the LTV limit. 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 the 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 with 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, the study design and the choice of a 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" to report 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 reported results. The most important are the ones related to model specification, estimation methods, and data characteristics. Among other, our results points to the presence of a strong endogenenity bias related to the estimation of the effects of borrower-based measures on bank lending. Endogenenity of credit and macroprudential policy biases estimates towards zero. We find that studies that took measures to assure exogenenity of the policy shock are found to report more negative estimates. This may be linked to our another finding that studies using confidential data also report more negative effect. Confidential data are usually capturing development at the level of a single entity or a product (e.g. bank, firm or loan) and studies using such data can 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 those specifying lagged model, 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. Most widely cited papers in the relevant literature build on cross-country data and capture the occurrence of macroprudential policy measures (including the borrower-based measures) 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 lower the real credit growth rate and slow-down house price growth. While the dummy-approach allows to compare the effects of regulation across countries, it falls short on capturing the intensity of the given measure. For instance, decreasing the LTV limit from 100% to 80% is treated the same as changing the limit 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 failing 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 to examine the systematic dependencies of reported results on study design and to filter out the effects of misspecifications. It thus informs the policymakers and a wider professional public on 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 few policy tools in a single country. While the micro-founded evidence offers more precise estimates of the effects taking place, it represents 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 then serve as a benchmark against which country-specific studies may assess their estimates.

There are several other studies that sums 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 7 Latin America countries. Araujo et al. (2020) summarizes estimates of macroprudential policy effects 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, they highlight that the reported estimates of elasticities 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 data from 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 incorporates loan-to-value limit, a cap that directly restricts the amount of the loan to a certain fraction of the total value of the residential real estate. The second category includes income-based limits, such the debt-to-income and the debt-service-to-income limit. The aim is to prevent the pledge 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 that are usually country-specific. These include but are not limited to taxes of property gains, direct quantitative limits on mortgage lending or change 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 share of studies (31 out of 34) focus on modelling the impact of borrower-based measures on the stock of credit. This somewhat contrasts the fact that borrower-based measures are applied on new business, i.e., on flow, while authors study the impact on the stock of loans. One can therefore expect slower and more gradual impact of the measures on the changes of the stock of loans. About 60% of collected estimates is based on the same variable transformation – credit growth and the borrower-based index or dummy (see Table 2). One prevailing variable transformation allows us to provide an empirical summary of the effect of changes to the borrower-based limits on bank lending which provides a significant benefit compared to a unitless standardized effect approach (Doucouliagos and Laroche, 2003; Ahmadov, 2014; Valickova et al., 2015; Bajzik, 2021).[4] 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 heterogenity in the estimates.

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

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

where $\%\Delta L_{it}$ is annual credit growth, BBM_{it} is 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

[4] Estimates 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.

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 active borrower-based measures across the analysed period. The literature uses two basic indexes – frequency index and direction index. The frequency index, as the name suggests, adds up all 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, assumes negative values (-1) for loosening actions and positive values (+1) for tightening actions. Given the similar construction of these three measurements, we can interpret the estimated elasticities $\hat{\beta}$ as a percentage point change in annual credit growth in response to 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 type and definition of credit variable or borrower-based measures, different set of control variables or estimation approach. For instance, our sample studies consider two different types of borrower-based measures – value-based (limits on loan-to-value ratio; LTV) and income-based (limits on debt service-to-income ratio and debt-to-income ratio; DSTI and DTI). We use the state of the art meta-analytical techniques to construct summaries of the estimated elasticities, aiming to verify the presence of a 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 that are relevant for our topic. We employed Google Scholar as the main database for our search and screened the first 350 articles that were 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*

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

After reviewing the first set of articles, we continued with so-called "snowballing" method. As such, we went through all the citations in each of the relevant studies and identified 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 abstract and additional 66 due to lack of correspondence or data (see Figure A1). We ended up with 34 primary studies from which we collected 722 point estimates.[5] In the process of data collection, we closely followed the guidelines established in Havránek et al. (2020).[6]

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 "just" working papers might be frown upon, we need to keep in mind that in the central banking industry, working papers are one of the main communications tools. As such, they go through a rigorous peer-reviewing process, much like journal articles (Malovaná et al., 2020). During the screening process, we searched for working paper (journal) version for 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 version that were unique. Second, we disregarded all the elasticities that were reported without standard error, p-value, t statistic or confidence intervals. Third, we harvest elasticities for both borrower-based measure and its interaction with dummy variable[7] in case the specification employs it. We then treat these elasticities as separate observations in our data set and distinguish the two by an indicator variable. We use similar approach for specifications with additional lags of borrower-based measure. That is, we collect all the elasticities for multiple lags and employ each of them in the analysis together with an indicator variables to examine 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 harvest the elasticities for immediate, after-one-period and the maximal responses from each IRF. Again we then use all three observations as "stand-alone" elasticities and differentiate between them in the model averaging exercise (see Section 4) by employing corresponding dummies.[8]

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

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

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

[8]Fidrmuc and Korhonen (2006) used 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	Gerls 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 – the versions differ in time period examined; C – different number of countries is studied; O – the journal and working paper version differ in a sense that only one version contains elasticities of interest. Estimates that differ between journal article and 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 collected estimates to prepare them for the analysis and to achieve their overall comparability. First, we manually computed standards errors for all the observations for which it was not reported, using the information on t statistics, p-value or confidence intervals. Second, we unified all the elasticties that capture effect on credit growth to yearly effects, e.g. we multiplied the non-annualized quarterly estimates by four, etc. Third, we converted the cumulative elasticities to average one-period estimates by dividing them by 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 also to 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 variables' transformation. This is because we do not have any prior knowledge about the most common model specifications for the studied relationship. The first part of Table 2 informs about the most frequent combinations of left-hand side and right-hand side variables in our initial data set. 59% of all the estimates denote the effect of a one-unit change in the borrower-based index or dummy on annual credit growth. All other pairs of variables' transformations are rather minor and the runner-up combination represents slightly more than 12% share. All in all, the majority of primary studies worked with the same typology of borrower-based measure – with either borrower-based index or a dummy (BBM index/dummy). We consider these two classes of variables in one category because when employed in the estimation, the effect of both is measured as the effect of 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 of all collected estimates converted to partial correlation coefficients (PCCs), which is a method of standardisation commonly used in meta-analyses (Doucouliagos, 2005; Havranek and Irsova, 2010; Havranek et al., 2016). PCCs allow to compare estimates with different units of measurement which would not be directly comparable. However, by taking this transformation, we cannot make any conclusions about the true size of the estimate, only about its direction. PCC from i^{th} estimate of the j^{th} study can be derived from the t-statistics of the reported estimates and residual degrees of freedom: $PCC_{ij} = t_{ij}/\sqrt{(t_{ij}^2 - df_{ij})}$. BBM index/dummy refers to borrower-based measures reported in a form of 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 variables' transformation make the comparison of resulting elasticities impossible. The partial correlation coefficient (PCC) method is able to address a part of this issue, because it uses 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 and as such, PCC reflects only the direction, not the true size of the effect. The second part of Table 2 reports the results of PCC method. The two most common transformations as well as the total sample consistently lean toward negative correlation between borrower-based measure and credit. Although the other transformations exhibit 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 but also to quantify the *size* of the effect, we focus only on the most frequent variables' transformation – Credit growth $\sim \beta \times$ BBM index/dummy, which encompasses 424 observations drown from 23 studies. As such, we can directly employ collected estimates instead of PCCs in the analysis. Elasticities corresponding to selected transformation represent the percentage-point effect of one-unit increase in borrower-based index or dummy on annual credit growth.

2.2. Early View of the Collected Elasticities

Early view of the collected elasticities suggest five stylized facts. First, the average estimated effect of introducing or tightening of borrower-based measures on bank lending is highly negative: A unit increase in borrower-based index or dummy is associated with a decrease in annual bank lending growth by 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 the 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 deems exploration, despite the clear negative sign on the relationship. Reported estimates vary both within and across studies (see Figure A2).

Second, the mean estimated effect of value-based (LTV) and income-based (DTI, DSTI) limits on bank lending is observably different 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 of 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 DSTI limit on credit growth to be negative while finding much smaller effect for the LTV limit. Additionally, this result echoes to the literature showing that while LTV limits can be successful in lowering demand for and 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 collateral value 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 at -2 pp. The "belt and braces" approach through the imposition of value-based LTV

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 analyse						
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 of the narrow sample of estimates related to primary studies using in the estimation either dummy variable or index for capturing the effect of borrower-based measures. If the "loosening" dummy or index takes on positive value in the primary study, we multiply the collected estimate by -1 to make the direction of effects comparable relative 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 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 the estimates obtained based on dynamic model specification. As such, we are able to calculate 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 one: $CreditGrowth_t = \alpha \cdot CreditGrowth_{t-1} + \beta \cdot BBM_t + \gamma \cdot X_t + \epsilon_t$.

limits combined with income-based DTI and DSTI limits thus appears to have the highest 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 real estate housing market. Evidence that LTV, DTI and DSTI limits complement each other were also found in case of Slovakia (Jurča et al., 2020; Cesnak et al., 2021).

Third, the choice of researcher on how to capture the 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 smaller effect than considering a dummy-based index that sums up all borrower-based actions during the analysed period. 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. To this purpose, many studies use a cumulative policy index which sums up 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 indexes – one that marks one the application of any policy measure (regardless of the tightening/easing nature of said measure) and thus captures the general usage of macroprudential policies (frequency index) and other that assigns value of 1 to tightening and -1 to easing actions which is then more respective of capturing 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 of the direction index is -1.1 pp. Moreover, we find that the mean of studies considering only the tightening measures (-0.71 pp) is fairly close to the mean effect in studies considering loosening measures only (0.84 pp), suggesting a symmetric effect of borrower-based measures. This insight from our empirical summary of the literature somewhat at odds with single studies which generally arrive on somewhat 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, but the difference is negligible in downturns. Kuttner and Shim (2016) find the effects of tightening to be significant in their relations to housing credit, 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.

The explanation can be threefold. One, studies incorporating more recent data capture more macroprudential borrower-based actions since their use has increased rapidly following the emergence of the GFC (Cerutti et al., 2017a; Alam et al., 2019). Two, it could be the reflection of the fact that borrower-based measure that have been used after the GFC are more binding and/or that the regulatory shocks are getting bigger. Three, higher negative effect of borrower-based measures on bank lending can be related to the large structural changes that has occurred in the post-GFC period. For instance, low interest rates that has depressed bank profitability could have made banks more exposed to regulatory shocks. 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 demand and the volume of real estate mortgage loans, making them more responsive to the introduction of borrower-based limits.

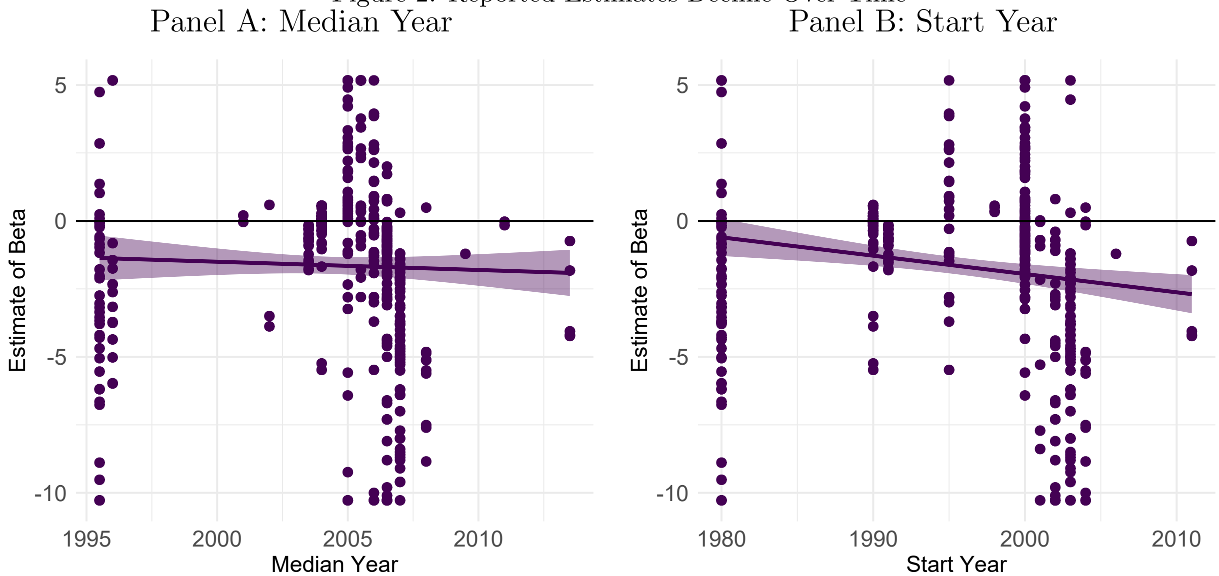
Fifth, the effect of borrower-based measures on bank lending is somewhat larger in the long term (-1.53 pp) than in the short term (-1.15 pp) but the difference is not as significant as might be expected. This finding touches upon the on-going debate on whether the borrower-based measures should be used 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.

3. Publication Bias

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

Publication bias is a phenomena that discusses the case when researchers do not publish all their estimates, but only those that are significant or with the “correct” sign. Such a selection accompanied with consequent exaggeration of the reported estimates affects the economic field in large degree (e. g. Ioannidis et al., 2017; Astakhov et al., 2019). The

Figure 2: Reported Estimates Decline Over Time



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

following fragments from sample studies shows that the selection bias might be a problem even in seminal papers estimating the effects of borrower-based measures on bank credit. For instance, a 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 area of macroprudential policy measures by Cerutti et al. (2017b) suggests that the anticipated 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 numerous possibilities in both, the study design and the choice of a 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. The publication bias was confirmed not only in economics (Campos et al., 2019; Gechert et al., 2020) and market-based finance (Kim et al., 2019; Gric et al., 2021), but in some bank-related studies as well (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 the publication bias by a series of graphical and econometric tests. The graphical tests typically use a funnel plot or histogram to show the distribution of the estimated elasticites and the related measures of estimate precision, such as the standard error, p-values or t-statistics.

Besides graphical visualization, a wide range of econometric tests can be run for clarifying the presence of publication bias among the estimates. Specifically, we use several linear techniques (ordinary least squares, weighted least squares, fixed effect model and hierarchical Bayesian approach). The linear estimation techniques provide a solid ground for the assessment of the presence of a publication bias as well as allowing to estimate the size of the publication bias. They are used to test for the 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):

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

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

Table 4 presents the results of the tests for publication bias. We obtain negative and statistically significant estimate of publication bias across different specifications and estimation methods. 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), which 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 can be systemically exaggerated due to the presence of a publication selection.

The baseline regression, informative as it already is, tells us little about the sources of the 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) want to mainly 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

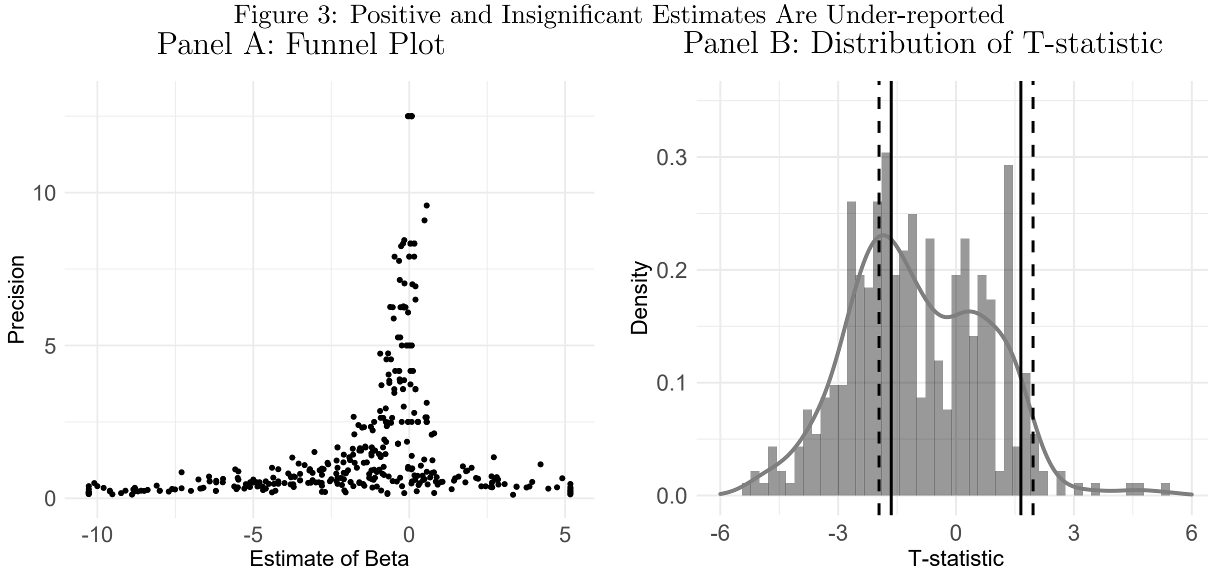
Note: The table presents the results of regression of equation (2). The standard errors, reported in parentheses, are clustered at both the level of the study and type of the borrower-based measure used in the primary study. OLS – ordinary least squares. Study – the inverse of the number of estimates 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 the Publication Bias

In Figure 3, we search for the incidence of the first and the second potential source of the publication bias. Figure 3, panel A is a funnel plot which depicts the estimates' magnitude on the horizontal axis against the estimates' 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 the less precise estimates should form 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 asymmetrical and visibly skewed towards negative spectrum of the distribution, showing a clear preference for negative values. Positive estimates appear but are mostly associated with small precision. Hence, the first visual test indicates that the search for the "correct" (negative) sign contributes to the documented publication bias.

Figure 3, panel B is a histogram showing the distribution of 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., 2020). In the absence of the publication bias, the distribution of the t-statistics should be approximately normal (Egger et al., 1997). We record 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 just between the 5% and 10% significance level, suggesting that researchers search for "at least some significance" of their negative estimates. From the observed distribution of t-statistics is also apparent that when researchers report positive estimates, they lie outside the statistical significance region. Overall, we find that both sources of the publication bias might be at work.

Figure 3: Positive and Insignificant Estimates Are Under-reported

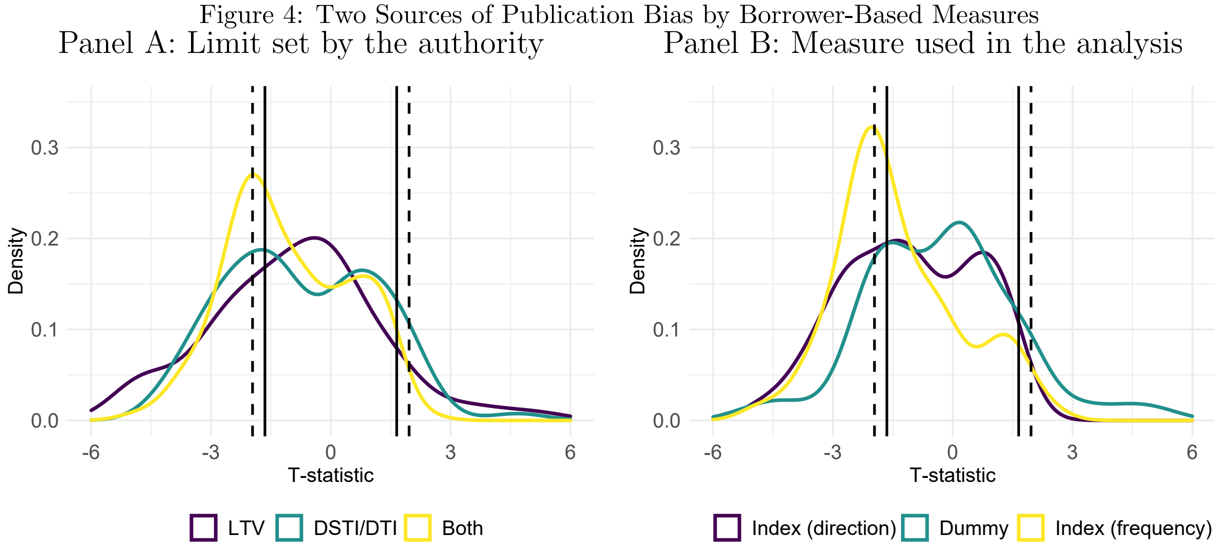


Note: Panel A: Precision is calculated as an inverse of standard error. In the absence of publication bias the funnel should be symmetrical around the most precise estimates. We exclude estimates with extreme magnitude or precision from the figure but include all in the regressions. Panel B: The vertical lines denote the critical value 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, panel A 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 not significant estimates appear mainly for the introduction or tightening of DSTI/DTI limits. Beyond publication selectivity, this can reflect the fact that the use of DSTI/DTI limits is much less frequent than the use of LTV limits and the little variance may lead to an attenuation bias. It holds that researchers report mostly estimates of both, the

LTV and DSTI/DTI limits, that are negative and statistically significant. While inspecting panel B of Figure 4, we find that studies using 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).

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



Note: The vertical lines denote the critical value 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.

The graphical evidence is further supported by empirical tests, results of which are shown in Table 5. We aim to verify the hypothesis that researchers prefer significant negative estimates over the non-significant and positive ones. First, we extend the equation (2) by a dummy variables 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 have a negative sign. Then we gradually regress the estimate on each of the two dummy variables and an interaction of the dummy with the estimate's standard error. In both specifications, the parameter of the interaction term captures the strength of publication bias. Estimates from Table 5, panels A and B, 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 Both Selection of 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 regression of equation (2) extended by additional dummy variables for collected elasticities significant at 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 type of the 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 Capture Macroprudential Policy Measures

The dominant stream of the literature relies on dummy-coded approach to capture the introduction 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, it is possible for the LTV limit to use the actual percentage change to the requirement, for instance, tracking a 10 pp change from LTV cap of 90% to 80%. However, one needs to keep in mind that the regulation can take different forms and tightening or easing might not always be transmitted to a change in the numerical setting of the limit. In countries like Korea and Hong Kong, which have used LTV limit 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 setting of these characteristics also change quite frequently. Using data on macroprudential policy action from Alam et al. (2019), we find that only about half of LTV tightening actions contain change to the numerical LTV limit. Thus, it is not a straightforward task to capture the overall intensity of LTV cap in a single country, much like to obtain a cross-country comparable measure. Attaching a value to the degree of intensity of a given regulation requires unavoidably a certain degree of subjectivity. Notwithstanding the fact that, without access to detailed mortgage market data, one can hardly know how much the policy is actually binding which may create attenuation bias.[9]

Bearing in mind the limitation of both, the dummy- and the intensity-adjusted approach, we undertake efforts to ease economic interpretation of our "true effect" (beyond bias) estimates. Specifically, we consult the integrated Macroprudential Policy (iMaPP) database maintained by IMF staff that offers a unique numerical indicator of regulatory limits on the loan-to-value 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 existing databases that were dominantly used in our surveyed studies thus we expect the country coverage to be high as

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

well. We focus on the application of LTV limits because only this limit (out of the three main borrower-based measures, LTV, LTI, and DSTI) has information about its intensity available in the database. 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 value change. If the distribution of changes to LTV turns out to be non-normal and visibly skewed (towards smaller or higher changes), it can help to place the estimated true effect coefficient into economic context and verify its meaningfulness for policymakers.

Table 6 shows the frequency of changes to LTV grouped into six buckets. Apparently, the most frequent change to the LTV limit is by less than 10 pp while changes of approximately 5 pp points are the most frequent. Under the (strong) assumption that numerical changes to LTV limit drive the estimates in the literature, the "true effect" estimate (about -0.9 pp decline in credit growth) would be in response to a change to LTV limit by less than 10 pp, or even more likely by 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), counting also very little gradual changes to the LTV limit. For example, in 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 not significant estimates are under-reported in the existing literature. This finding can reflect both, the strong personal a priori belief of a researcher or a pressure to publish significant estimates that are in line with the theory or a dominant stream of literature. We now discuss and test the robustness of our results by addressing shortcomings stemming from the employed linear regression.

First, the 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 be always 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 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.[10] Although this approach 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 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 a relatively modest number of observations at our disposal. Most of these methods work by searching (exogenously or endogenously) for a precision threshold below which the estimates are discarded. As such, they remove a substantial number of imprecise estimates, in our case imprecise positive elasticities, shifting the entire distribution to the negative region and yielding potentially inflated negative effect.

Second, the standard error is assumed to be exogeneous. Even the exogeneity condition might not hold since the 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. Estimates are stored in Appendix

[10] Ioannidis et al. (2017) propose a procedure that focuses only on estimates with statistical power above 80%. Similar approach is proposed by Stanley et al. (2010) who suggest to focus on the top 10% of the most precise estimates (the Top10 method). Also 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 (rising number of imprecise estimates that are included) and variance (decreasing 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 so-called conditional publication probability (the probability of publication as a function of a study's results) and then use it to correct for publication bias. The underlying intuition involves jumps in publication probability at conventional p-value cut-offs.

Appendix B. These methods are based on verifying the presence of the so-called p-hacking, i.e. a situation where authors prefer to include significant estimates rather than the not 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. (2014b,a) and the Elliott et al. (2021) tests.

The Caliper test specify narrow bounds surrounding the commonly employed thresholds for the 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, stored in Table A1, 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 the 50:50) as reasoned in Bruns et al. (2019), then the evidence of publication selection is present for negative estimates and also when all estimates (positive and negative) are tested together.

The p-curve method and the Elliott et al. (2021) tests are centered around testing the distribution of p-values. They jointly 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 would be skewed. We first employ the p-curve method which test the null hypothesis that the literature has no evidential value. The evidential value thus conforms to the existence of a strong prior knowledge on the likely effect of the borrower-based measures on bank lending. In another words, there should be no effect of borrower-based measures on bank lending beyond the 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., 2014a). Based on Figure A4, we obtain evidence of the existence of a strong evidential value as the p-value distribution is visibly left-skewed. Figure A5 further shows that p-hacking mostly concerns negative estimates at a 5% and 10% level of significance and also suggest the existence of an over-reporting of not significant positive estimates. Table A2 shows results of the so-called Elliott tests (a series of binomial, Fisher's, and density discontinuity tests) with a null hypothesis of no p-hacking. Details on the method are available in the Appendix Appendix B. We reject the null of no p-hacking while 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. Estimates are stored in Appendix 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 to the hypothesis, that the estimated effect of the introduction or tightening of borrower-based measures on annual bank credit growth is prone to the publication bias which is driven by both selection of the "correct" sign and the search for statistical significance (see Figure A6 and Table A4).

4. Drivers of Heterogeneity

Given the observed differences across the primary studies, one can object that the heterogeneity of the estimates might have different origin than the 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 drives 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 to order the factors that drives the heterogeneity in collected estimates by their importance.

Based on the above mentioned reasons, we control for more than 20 primary study characteristics in three areas – data characteristics, model specification and estimation, and publication characteristics. Before arriving to 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%. Albeit we set up these criteria somewhat arbitrary, we tried different thresholds as well without major impact on the results. The final set of characteristics, including the description and summary statistics, is shown in Table 7. Next, we provide a brief reasoning for the inclusion of specific control variables.

Data characteristics. The information about the effect of different data characteristics on the relationship between borrower based measures and bank lending are 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), among other, find that the expression of the dependent variable (in their application – different expressions of the bank capital ratio), type of credit considered as well as different midpoint of the data matters. 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 type of credit used as a dependent variable, type of borrower-based measure as an independent variable, data frequency, midpoint of the data, number of countries in the estimation sample, the region of the analysis, and data confidentiality.

Table 7: Variable Definitions and Summary Statistics

Variable	Definition	Mean	SD	W. Mean	W. SD
Estimate	The reported estimate of the coefficient beta.	-1.46	3.15	-1.7	3.53
St. error	The reported standard error of the coefficient beta.	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 direction index.	0.47	0.5	0.41	0.49
Index (frequency)	= 1 if the borrower-based measure (independent variable) is frequency index.	0.27	0.45	0.24	0.43
Household loans	= 1 if household credit is used as 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 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 data frequency is annual.	0.33	0.47	0.34	0.48
Europe	= 1 if the study covers a country or a 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 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 lagged dependent variable.	0.72	0.45	0.8	0.4
GMM method	= 1 if general method of moments (GMM) is used.	0.18	0.39	0.39	0.49
Fixed-effects method	= 1 if 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 number of citation divided by number of years from its 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.

Model specification and estimation. Other meta-analytical papers reveal a key impact of the study design on the size and the 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 a static and a dynamic model. Dynamic model specification contains lagged dependent variable, which is supposed to capture persistence in credit dynamics. 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[11] and interaction terms[12] 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 the primary study quality. Among these variables we include the publication year, discounted recursive impact factor, annualized number of citations and dummy variable indicators whether the study was published in the 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 to Analyze Drivers of Heterogenenity

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

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

[12] Interaction terms explore heterogeneity in the analyzed effect. The most common interactions are the size of the bank and GDP growth. Empirical literature suggests that the effect of macroprudential policies is significant for small and medium 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 suggest that the policy instruments are countercyclical (Budnik, 2020; Lim et al., 2011).

underlying model.

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

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

We examine the sensitivity of our results through different prior choice in the robustness checks. For example, we employ dilution prior proposed by George (2010) as alternative choice for 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 case of high correlation, the determinant is close to one and has assigned just little weight and vice versa. Such approach was used in recent meta-analyses in economics, see for instance Gechert et al. (2020) and Bajzik et al. (2020). We also employ combination of the Hannan-Quinn (HQ) g-prior that 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 combination of the BRIC g-prior that minimizes the prior effect on the result and the random model prior that provides equal probability to every model size (Zeugner and Feldkircher, 2015; Gechert et al., 2020). Last but not least we use frequentist model averaging (FMA) and frequentist check (OLS) as additional two robustness checks.

Following the approach proposed in Eicher et al. (2011), we interpret in the following section only the variables with PIP above 0.5. The categorization is as follows. The variable is classified as decisive if the PIP is higher than 0.99. In range between 0.95 and 0.99 the variable is marked as strong. If the PIP falls between 0.75 and 0.95 the variable is deemed substantial and between 0.5 and 0.75 is called weak. Besides the PIP 0.5 threshold serves us baseline for variables inclusion in OLS frequentist check.

Reference model. We define our reference model based on the predominant characteristics of the primary studies. It means that for each dummy variable group (e.g. contemporaneous effect, effect lagged by less than one year, and 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 let ourselves the freedom of discretion to choose to include the more proper variable in the reference model. We use this option in case that the difference between two characteristics is small or in order to smooth the interpretation.

4.2. Results

Results of the Bayesian model averaging are visualized in Figure 5. In the figure, each column represents the individual regression models while the column width captures the posterior model probability. Each row denotes 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. The red color (lighter in grayscale) indicates a negative sign of the variable's coefficient, i.e. the estimated effect is more negative. The blue color (darker in grayscale) indicates a positive sign, i.e. the effect is less negative. In case a variable is excluded from the particular model, the cell is left blank. Results using different priors are stored in the Appendix, Figure A3.

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 a simple OLS, excluding variables that turned out less important in the BMA exercise (i.e. the pposterior inclusion probability is below 0.5).

Publication bias. The presence of the publication bias in the collected estimates is supported by an 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 the estimation method. The size of the publication bias across different methods is around -0.6 (Table 8) which corresponds to the estimated magnitude of the publication bias from 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 an ordinary least squares (OLS) estimator. In theory, the GMM estimator should be more successful in reducing the endogeneity bias when compared to the OLS estimator (Baum et al., 2003). In fact, the GMM has been 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). The modelling of the impact of borrower-based measures on bank lending is prone to the endogeneity bias, especially since decisions to take policy actions are taken with respect to the concurrent state of the financial sector as well as its future development, so that in equation (1), $cov(BBM, \epsilon) \neq 0$.[13] OLS estimates of the $\hat{\beta}$ elasticity in equation (1) can therefore be biased towards zero, given that $cov(BBM, \epsilon) > 0$. 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 the 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 those specifying lagged model. In another words, the estimated effect of borrower-based measures is found to have stronger effect with a lag, suggesting that the impact of the measures takes time to materialize. Consider a case when policymakers tended to tighten the borrower-based requirements when housing credit was already expanding rapidly. This 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 and in his setup, it reaches its peak only after three years. McDonald (2015) estimates the impact of LTV loosening to be significantly positive

[13] 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 the endogeneity bias. Borrower-based measures are set on a country-level and are a direct response to the development in the real estate mortgage market. This is in contrast to the application of capital-based measures – and the literature studying its effects thereof – from which only the counter-cyclical buffer is used with respect to the financial cycle development.

only when credit growth was previously strong.

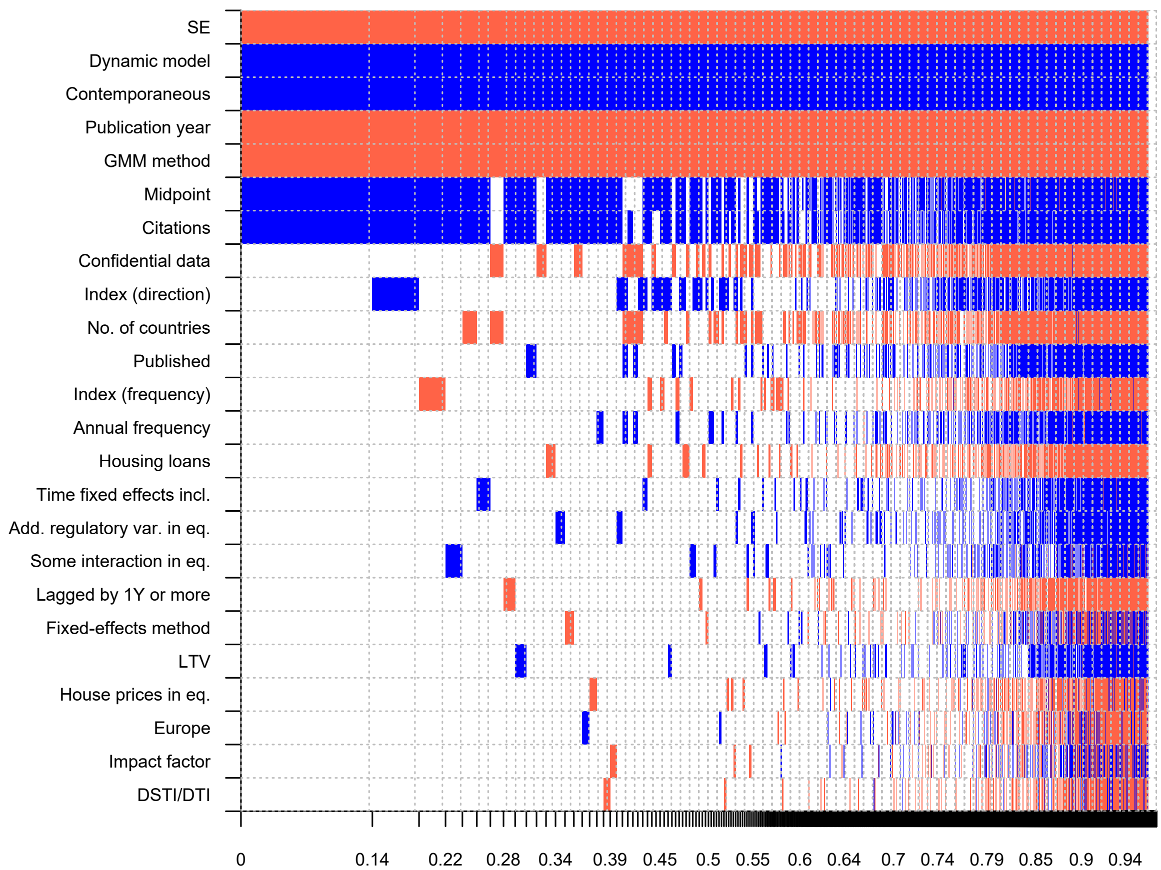
Last, we discover that the inclusion of a lagged dependent variable into the model specification (i.e. dynamic model specification) produces lower negative estimate of the relationship between borrower-based measures and bank lending. The inclusion of the lagged dependent variable generally explains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

Data characteristics. We find that papers using confidential data report more negative effect of borrower-based measures on bank lending. Confidential data are usually capturing development at the level of a single entity or a product (e.g. bank, firm or loan) and are usually available to supervisory or regulatory institutions only. Studies using such data can in theory be more successful than others in correctly identifying exogenous regulatory shocks that are plausibly unrelated to lending opportunities due to more detailed data. As such, they are less likely to suffer from the endogeneity bias described above.[14] Notwithstanding the fact that macro-level studies capture averages across constrained and unconstrained agents which can under-estimate the regulation's effect. 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.

Publication characteristics. BMA estimates indicate strong association of two publication characteristics with the collected estimates – the year of the 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 the potential effect of quality: more recent studies have access to better data as well as better estimation methods that allows them to filter out pollution related to measurement errors or endogeneity that biases estimates towards zero. Another explain lies in the fact that majority of borrower-based regulation occurred 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 number of citations, owing to the fact that newer studies will have by definition and on average less citations. Highly cited documents are those with less negative elasticity estimates.

[14]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.

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.

Table 8: What Drives the Heterogeneity of 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 covariate space suggested by Amini and Parmeter (2012) to reduce the number of estimated models. The frequentist check (OLS) includes the variables with PIP estimated by BMA above 0.5 and is estimated using standard errors clustered at the study level. Description and summary of all 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 amongst the most commonly used macroprudential policy measures worldwide and have seen an increased use in the post-GFC period. Despite the broad usage 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 important service to both researchers and policy-makers. 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. While inspecting basic statistical properties of the collected 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 of DTI and DSTI limits on bank lending is about 20% stronger than that of LTV limit. Furthermore, the statistics indicate that a joint application of both, LTV and income-based limits is the most effective with the negative effect on the 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 on a definitive conclusion on the strength of the effects of individual borrower-based tool, the summary statistics from the meta-analytical dataset offers a much-needed early evidence on the matter and highlights 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 grows 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 the effect of quantity where newest studies can exploit larger variation of borrower-based policy. It can also be a reflection of a changing macro-financial environment after the Global financial crisis of 2007–2009 where entering low interest rate environment depressed bank profitability and capital, making them more likely to respond to regulatory tightening.

Using series of empirical tests, we next find that estimates of the impact of borrower-based measures are prone to the publication selection. This stems from the fact that borrower-based measures are expected to have a negative effect on bank credit provision. Positive or not statistically significant estimates might thus act as a psychological barrier suggesting that the employed data or model are incorrect and thus, the estimates are un-publishable. Using plethora of methods, we find that researchers show preference when publishing results for elasticities that have the "correct" negative sign and are "just" significant at the 10% level. The identified publication bias results in an exaggeration of 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 the explanation of the variation in the reported elasticities is the standard error. Large standard error is found to be associated with large negative estimates. In the absence of publication selection, 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 the underlying data characteristics.

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

Our evidence also suggest that contemporaneous models might be misspecified and underestimate the effect of borrower-based measures on bank lending. The estimated effect of borrower-based measures is found to have stronger effect 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 growing 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 employed modelling framework. A central banker, wishing to calibrate the effect of introduction or tightening of borrower-based measures into her stress-testing framework, would have a difficult job finding the correct elasticity value. The evidence provided in the paper can serve as a useful benchmark against which countries – with no historical experience with the use of borrower-based measures – can back up their forecasting or stress-testing models models.

References

Acharya, V.V., Bergant, K., Crosignani, M., Eisert, T., McCann, F., 2020a. The Anatomy of the Transmission of Macroprudential Policies. IMF Working Papers 20/58. International Monetary Fund.

Acharya, V.V., Bergant, K., Crosignani, M., Eisert, T., McCann, F.J., 2020b. The anatomy of the transmission of macroprudential policies. Technical Report. National Bureau of Economic Research.

Afanasieff, T.S., Carvalho, F., de Castro, E.C., Coelho, R.L., Gregório, J., 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, 1238–1267.

Ahuja, A., Nabar, M., 2011. Safeguarding Banks and Containing Property Booms: Cross-Country Evidence on Macroprudential Policies and Lessons from Hong Kong SAR. IMF Working Paper. International Monetary Fund.

Akinci, O., Olmstead-Rumsey, J., 2015. How Effective Are Macroprudential Policies? An Empirical Investigation. International Finance Discussion Papers 1135. Board of Governors of the Federal Reserve System.

Akinci, O., Olmstead-Rumsey, J., 2018. How effective are macroprudential policies? an empirical investigation. Journal of Financial Intermediation 33, 33–57.

Alam, Z., Alter, M.A., Eiseman, J., Gelos, M.R., Kang, M.H., Narita, M.M., Nier, E., Wang, N., 2019. Digging Deeper – Evidence on the Effects of Macroprudential Policies from a New Database. IMF Working Papers. International Monetary Fund.

Amini, S.M., Parmeter, C.F., 2012. Comparison of model averaging techniques: Assessing growth determinants. Journal of Applied Econometrics 27, 870–876.

Andrews, D., Petroulakis, F., 2019. Breaking the Shackles: Zombie Firms, Weak Banks and Depressed Restructuring in Europe. Working Paper Series 2240. European Central Bank.

de Araujo, D.K.G., Barroso, J.B.R., Gonzalez, R.B., 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., Barroso, J.B.R.B., Gonzalez, R.B., 2020. Loan-to-value policy and housing finance: Effects on constrained borrowers. Journal of Financial Intermediation 42, 100830.

Araujo, J., Patnam, M., Popescu, A., Valencia, F., Yao, W., 2020. Effects of macroprudential policy: Evidence from over 6,000 estimates .

Armstrong, J., Skilling, H., Yao, F., 2019. Loan-to-value ratio restrictions and house prices: Micro evidence from new zealand. Journal of Housing Economics 44, 88–98.

Astakhov, A., Havranek, T., Novak, J., 2019. Firm size and stock returns: A quantitative survey. Journal of Economic Surveys 33, 1463–1492.

Ayyagari, M., Beck, T., Martinez Peria, M.S., 2017. Credit Growth and Macroprudential Policies: Preliminary Evidence on the Firm Level. BIS Papers 91a. Bank for International Settlements.

Ayyagari, M., Beck, T., Peria, M.M.S.M., 2018. The Micro Impact of Macroprudential Policies: Firm-Level Evidence. IMF Working Papers. International Monetary Fund.

Bachmann, R., Rüth, S., 2017. Systematic Monetary Policy and the Macroeconomic Effects of Shifts in Loan-to-Value Ratios. CESifo Working Paper. CESifo.

Bachmann, R., Rüth, S.K., 2020. Systematic monetary policy and the macroeconomic effects of shifts in residential loan-to-value ratios. International Economic Review 61, 503–530.

Bajzik, J., 2021. Trading volume and stock returns: A meta-analysis. International Review of Financial Analysis 78, 101923.

Bajzik, J., Havranek, T., Irsova, Z., Schwarz, J., 2020. Estimating the armington elasticity: The importance of data choice and publication bias. Journal of International Economics 127, 103383.

Baum, C.F., Schaffer, M.E., Stillman, S., 2003. Instrumental variables and gmm: Estimation and testing. The Stata Journal 3, 1–31.

Boissay, F., Collard, F., Smets, F., 2016. Booms and banking crises. Journal of Political Economy 124, 489–538.

Bom, P.R., Rachinger, H., 2019. A kinked meta-regression model for publication bias correction. Research Synthesis Methods 10, 497–514.

Bordalo, P., Gennaioli, N., Shleifer, A., 2018. Diagnostic expectations and credit cycles. The Journal of Finance 73, 199–227.

Boz, E., Mendoza, E.G., 2014. Financial innovation, the discovery of risk, and the us credit crisis. Journal of Monetary Economics 62, 1–22.

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

Budnik, K.B., 2020. The effect of macroprudential policies on credit developments in europe 1995-2017 .

Campos, N.F., Fidrmuc, J., Korhonen, I., 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., Krueger, A.B., 1995. Time-series minimum-wage studies: a meta-analysis. The American Economic Review 85, 238–243.

Carreras, O., Davis, E.P., Piggott, R., 2018. Assessing macroprudential tools in oecd countries within a cointegration framework. Journal of Financial Stability 37, 112–130.

Cattaneo, M.D., Jansson, M., Ma, X., 2020. Simple local polynomial density estimators. Journal of the American Statistical Association 115, 1449–1455.

Cazachevici, A., Havranek, T., Horvath, R., 2020. Remittances and economic growth: A meta-analysis. World Development 134, 105021.

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

Cerutti, E., Claessens, S., Laeven, L., 2017a. The use and effectiveness of macroprudential policies: New evidence. Journal of Financial Stability 28, 203–224.

Cerutti, E., Correa, R., Fiorentino, E., Segalla, E., 2017b. Changes in prudential policy instruments—a new cross-country database. International Journal of Central Banking 13, 477–503.

Cesnak, M., Klacso, J., Vasil, R., et al., 2021. Analysis of the Impact of Borrower-Based Measures. NBS Occasional Paper 3/2021. National Bank of Slovakia.

Claessens, S., Ghosh, S.R., Mihet, R., 2013. Macro-prudential policies to mitigate financial system vulnerabilities. Journal of International Money and Finance 39, 153–185.

Doucouliagos, C., 2005. Publication bias in the economic freedom and economic growth literature. Journal of Economic Surveys 19, 367–387.

Doucouliagos, C., Laroche, P., 2003. What do unions do to productivity? a meta-analysis. Industrial Relations: A Journal of Economy and Society 42, 650–691.

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

Ehrenbergerova, D., Bajzik, J., Havranek, T., 2021. When does monetary policy sway house prices? a meta-analysis .

Eicher, T.S., Papageorgiou, C., Raftery, A.E., 2011. Default priors and predictive performance in bayesian model averaging, with application to growth determinants. Journal of Applied Econometrics 26, 30–55.

Elliott, G., Kudrin, N., Wuthrich, K., 2021. Detecting p-hacking. Econometrica .

Epure, M., Mihai, I., Minoiu, C., Peydró, J.L., 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., Zeugner, S., 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, 686–694.

Fernandez, C., Ley, E., Steel, M.F.J., 2001. Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics 16, 563–576.

Fidrmuc, J., Korhonen, I., 2006. Meta-analysis of the business cycle correlation between the euro area and the ceecs. Journal of Comparative Economics 34, 518–537.

Furukawa, C., 2019. Publication bias under aggregation frictions: Theory, evidence, and a new correction method. Evidence, and a New Correction Method (March 29, 2019) .

Fuster, A., Laibson, D., Mendel, B., 2010. Natural expectations and macroeconomic fluctuations. Journal of Economic Perspectives 24, 67–84.

Gadatsch, N., Mann, L., Schnabel, I., 2017. A New IV Approach for Estimating the Efficacy of Macroprudential Measures. Arbeitspapier 05/2017. German Council of Economic Experts.

Gadatsch, N., Mann, L., Schnabel, I., 2018. A new iv approach for estimating the efficacy of macroprudential measures. Economics Letters 168, 107–109.

Galati, G., Moessner, R., 2013. Macroprudential policy–a literature review. Journal of Economic Surveys 27, 846–878.

Galati, G., Moessner, R., 2018. What do we know about the effects of macroprudential policy? Economica 85, 735–770.

Gambacorta, L., Murcia, A., 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, 1–66.

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

Gennaioli, N., Shleifer, A., Vishny, R., 2015. Neglected risks: The psychology of financial crises. American Economic Review 105, 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., Malhotra, N., 2008a. Do statistical reporting standards affect what is published? publication bias in two leading political science journals. Quarterly Journal of Political Science 3, 313–326.

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

Gerls, A., Jasova, M., 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., Jašová, M., 2014. Measures to tame credit growth: Are they effective? Economic Systems 38, 7–25.

Gric, Z., Bajzik, J., Badura, O., et al., 2021. Does sentiment affect stock returns? a meta-analysis across survey-based measures .

Grodecka, A., 2020. On the effectiveness of loan-to-value regulation in a multiconstraint framework. Journal of Money, Credit and Banking 52, 1231–1270.

Hansen, B., 2007. Least squares model averaging. Econometrica 75, 1175–1189.

Havranek, T., Horvath, R., Zeynalov, A., 2016. Natural resources and economic growth: A meta-analysis. World Development 88, 134–151.

Havranek, T., Irsova, Z., 2010. Meta-analysis of intra-industry fdi spillovers: Updated evidence. Czech journal of Economics and Finance 60, 151–174.

Havranek, T., Irsova, Z., Laslopova, L., Zeynalova, O., 2021. Skilled and Unskilled Labor Are Less Substitutable than Commonly Thought. CEPR Discussion Papers 15724. CEPR.

Havranek, T., Rusnak, M., 2012. Transmission lags of monetary policy: A meta-analysis .

Havránek, T., Stanley, T., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W.R., Rost, K., Van Aert, R., 2020. Reporting guidelines for meta-analysis in economics. Journal of Economic Surveys 34, 469–475.

Hodula, M., Melecký, M., Pfeifer, L., Szabo, M., 2021. Cooling Down the Mortgage Loan Market: The Effect of Recommended Borrower-Based Limits on New Mortgage Lending. CNB Working Paper. forthcoming.

Igan, M.D., Kang, M.H., 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., Stanley, T., Doucouliagos, H., 2017. The power of bias in economics research. Economic Journal 127, 236–265.

Jácome, L.I., Mitra, S., 2015. LTV and DTI Limits – Going Granular. IMF Working Paper WP/15/154. International Monetary Fund.

Jordà, Ò., Schularick, M., Taylor, A.M., 2011. Financial crises, credit booms, and external imbalances: 140 years of lessons. IMF Economic Review 59, 340–378.

Jurča, P., Klacso, J., Tereanu, E., Forletta, M., Gross, M., 2020. The effectiveness of borrower-based macroprudential measures: A quantitative analysis for slovakia .

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

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., Morsink, M.J., 2014. With great power comes great responsibility: Macroprudential tools at work in Canada. IMF Working Papers. International Monetary Fund.

Kuttner, K.N., Shim, I., 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., Shim, I., 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., Steel, M.F., 2009. On the effect of prior assumptions in bayesian model averaging with applications to growth regression. Applied Econometrics 24, 651–674.

Lim, C.H., Costa, A., Columba, F., Kongsamut, P., Otani, A., Saiyid, M., Wezel, T., Wu, X., 2011. Macroprudential Policy: What Instruments and How to Use Them? Lessons from Country Experiences. IMF working paper WP/11/238. International Monetary Fund.

Malovana, S., Frait, J., 2017. Monetary policy and macroprudential policy: Rivals or teammates? Journal of Financial Stability 32, 1–16.

Malovaná, S., Hodula, M., Rakovská, Z., et al., 2020. Researching the Research: A Central Banking Edition. CNB Research and Policy Note 3/2020. Czech National Bank.

Malovaná, S., Bajzík, J., Gric, Z., Hodula, M., 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.When is Macroprudential Policy Effective?

Mendoza, E.G., Terrones, M.E., 2008. An Anatomy of Credit Booms: Evidence from Macro Aggregates and Micro Data. Technical Report. National Bureau of Economic Research.

Morgan, P., Regis, P., Salike, N., 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., Regis, P.J., Salike, N., 2019. Ltv policy as a macroprudential tool and its effects on residential mortgage loans. Journal of Financial Intermediation 37, 89–103.

Nansen McCloskey, D., Ziliak, S.T., 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, 356–361.

Neagu, F., Tatarici, L., Mihai, I., 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., Schularick, M., Shim, I., 2018. The Macroeconomic Effects of Macroprudential Policy. BIS Working Papers 740. Bank for International Settlements.

Richter, B., Schularick, M., Shim, I., 2019. The costs of macroprudential policy. Journal of International Economics 118, 263–282.

Schularick, M., Taylor, A., 2012. Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. American Economic Review 102, 1029–1061.

Simonsohn, U., Nelson, L.D., Simmons, J.P., 2014a. P-curve: A key to the file-drawer. Journal of Experimental Psychology: General 143, 534–547.

Simonsohn, U., Nelson, L.D., Simmons, J.P., 2014b. P-curve and effect size: Correcting for publication bias using only significant results. Perspectives on Psychological Science 9, 666–681.

Stanley, T.D., 2005. Beyond publication bias. Journal of economic surveys 19, 309–345.

Stanley, T.D., Doucouliagos, H., Giles, M., Heckemeyer, J.H., Johnston, R.J., Laroche, P., Nelson, J.P., Paldam, M., Poot, J., Pugh, G., et al., 2013. Meta-analysis of economics research reporting guidelines. Journal of economic surveys 27, 390–394.

Stanley, T.D., Jarrell, S.B., Doucouliagos, H., 2010. Could it be better to discard 90% of the data? a statistical paradox. The American Statistician 64, 70–77.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics 20, 518–529.

Tantasith, C., Ananchotikul, N., Chotanakarn, C., Limjaroenrat, V., Pongsaparn, R., 2020. The impact of LTV Policy on Bank Lending: Evidence from Thailand. BIS Working Papers 110. Bank for International Settlements.

Valickova, P., Havranek, T., Horvath, R., 2015. Financial development and economic growth: A meta-analysis. Journal of Economic Surveys 29, 506–526.

Van Bekkum, S., Gabarro, M., Irani, R.M., Peydró, J.L., 2019. Take it to the limit? the effects of household leverage caps. Working Paper (December 11, 2019) .

Wang, M.B., Sun, T., 2013. How Effective are Macroprudential Policies in China? IMF Working Paper Series. International Monetary Fund.

Zeugner, S., Feldkircher, M., 2015. Bayesian model averaging employing fixed and flexible priors: The bms package for r. Journal of Statistical Software 68, 1–37.

Zhang, L., Zoli, E., 2016a. Leaning against the Wind: Macroprudential Policy in Asia. IMF Working Paper Series WP/14/022. International Monetary Fund.

Zhang, L., Zoli, E., 2016b. Leaning against the wind: Macroprudential policy in asia. Journal of Asian Economics 42, 33–52.

Zhang, Y., Tressel, T., 2017. Effectiveness and channels of macroprudential policies: Lessons from the euro area. Journal of Financial Regulation and Compliance .

Zigraiova, D., Havranek, T., 2016. Bank competition and financial stability: much ado about nothing? Journal of Economic Surveys 30, 944–981.

Zigraiova, D., Havranek, T., Irsova, Z., Novak, J., 2021. How puzzling is the forward premium puzzle? A meta-analysis. European Economic Review 134, 103714.

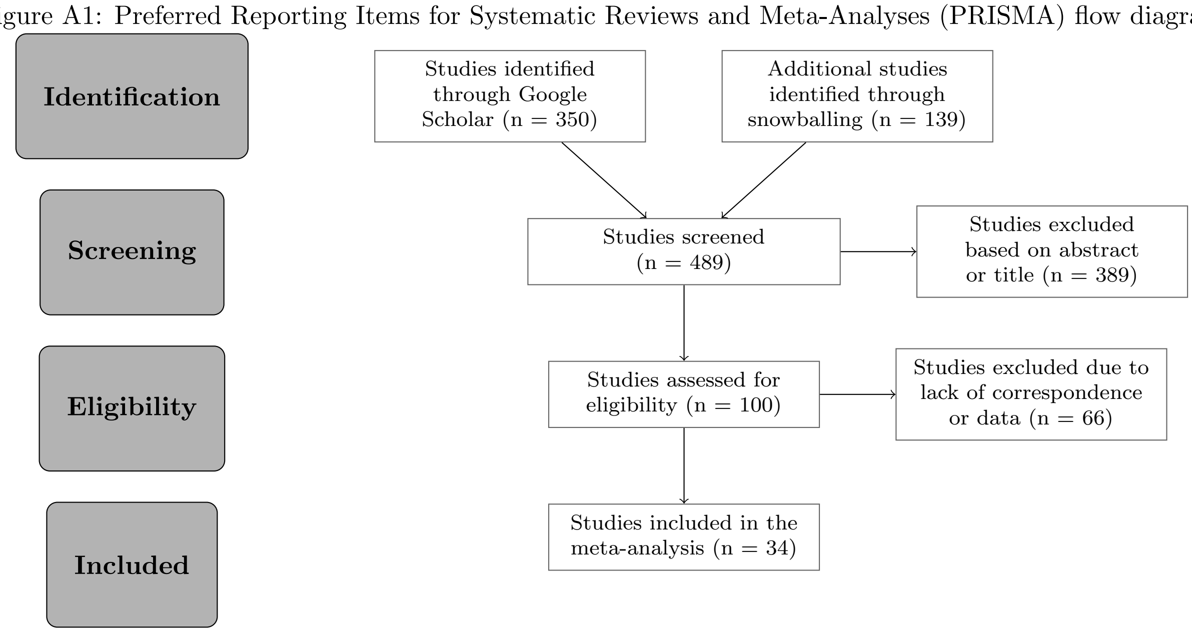
For Online Publication

Appendix A. Additional Charts

Appendix A.1. PRISMA Diagram

Figure A1 depicts the overall process employed in the selection of primary studies. As this article is sisterly 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 snowball all the citations in each of the relevant studies. Hence we identified additional 139 articles. Thus, in total we *screened* 489 articles. In the next step, we scrutinized all the titles and abstracts and reject all studies were not acceptable even from a high-level perspective. Hereby we eliminated 389 studies. The rest of 100 was assessed for *eligibility*. During this step, we investigate in detail each of these articles and dropped out 66 studies due to a lack of correspondence or data. The main elimination criteria were: (1) the study must report numerical results; (2) estimated elasticities must be presented together with the corresponding test statistic – standard error, t-statistic, p-value or exact confidence interval; (3) the effect is not a cross-boarder effect; 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. In sum, 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

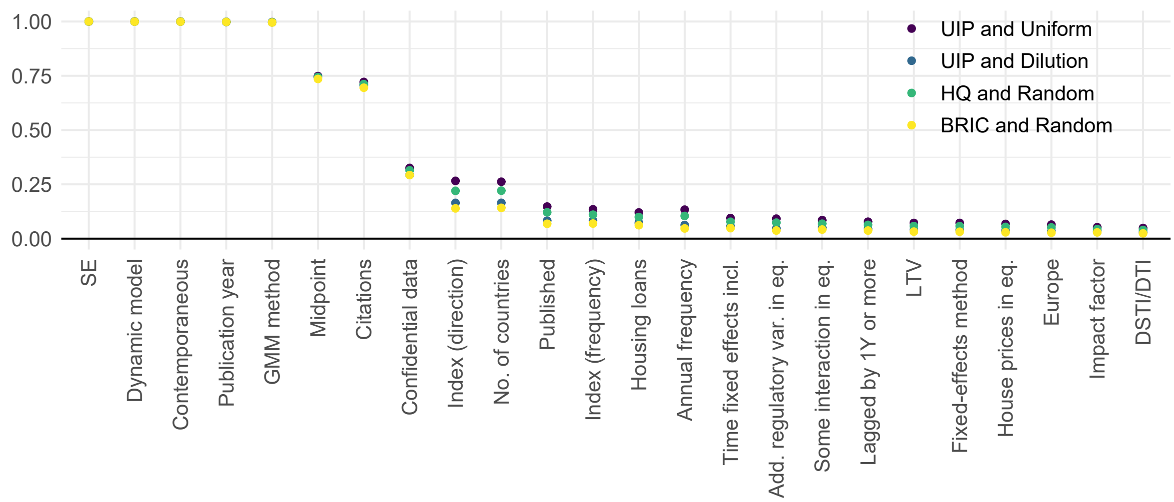


Appendix A.2. Variation of Collected Estimates

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 ration; both stands for measures encompassing multiple borrower based measures.

Figure A3: Bayesian Model Averaging – Prior Sensitivity



Note: The figures show posterior inclusion probability for different prior combinations. In our baseline, we use a unit information g-prior (UIP) and a uniform model prior which reflects our lack of prior knowledge. The uniform model prior gives each model the same prior probability, and the unit information g-prior provides the same information as one observation from the data. As a robustness check, we use a dilution model prior, as proposed by (George, 2010), to account for potential collinearity between explanatory variables. Next, we also employ a combination of the Hannan-Quinn (HQ) g-prior and random model prior and a combination of the BRIC g-prior and random model prior. The HQ g-prior adjusts data quality while the BRIC g-prior minimizes the prior effect on the results. The random model prior gives equal prior probability to every model size (Gechert et al., 2020).

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 by developed by Simonsohn et al. (2014b,a), and a set of statistical tests used by Elliott et al. (2021).

Appendix 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 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 underpowered (Ioannidis et al., 2017). As such, the frequency of t-statistic values is likely to decrease with the magnitude of t-statistic values if reporting biases are absent.

The results are presented in Table A1. We inspect two significance thresholds corresponding to the 90% and 95% confidence interval. 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 it goes against the economic logic.

Table A1: 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 that are 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 a total number of observations. We test two one-sided null hypotheses of no p-hacking: C is lower or equal to 0.5 and C is lower or equal to 0.4. Significant results are shown in bold (H0: $C \leq 0.5$) and italic (H0: $C \leq 0.4$). Lower 95% confidence intervals are reported in brackets.

Appendix B.2. Tests Based on the Distribution of P-values

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

The results are presented in Table A4. The figure shows that the distribution of p-values is visibly left-skewed, suggesting the presence of p-hacking. While looking at the individual estimates in terms of sign, we find that the p-hacking mainly concerns negative estimates. This is further supported by the distribution of p-values that is plotted in Figure A5. Overall, the p-values distribution confirms the existence of a strong evidential value, that is, researchers have strong prior expectations of the direction of the studies effect. In our case, researchers strongly expect to find negative effects of borrower-based measures on bank lending.

Second, we employ series of tests used by Elliott et al. (2021). Elliott et al. (2021) 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. (2014a) and Cattaneo et al. (2020) the tests apply binomial, Fisher's, and density discontinuity approaches. But in addition to them, Elliott et al. (2021) proposed several new tests with more statistical power. First of them is a histogram-based test for non-increasingness. Second one is a histogram-based test for 2-monotonicity and bounds. And third and the last one is the least concave majorant (LCM) test relying on concavity of the CDF of p-values.

The results are presented in Table A2. All of the tests have 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 test for discontinuity, non-increasingness (CS1), 2-monotonicity (CS2B). The tests imply presence of p-hacking mainly for positive estimates.

Figure A4: P-curve

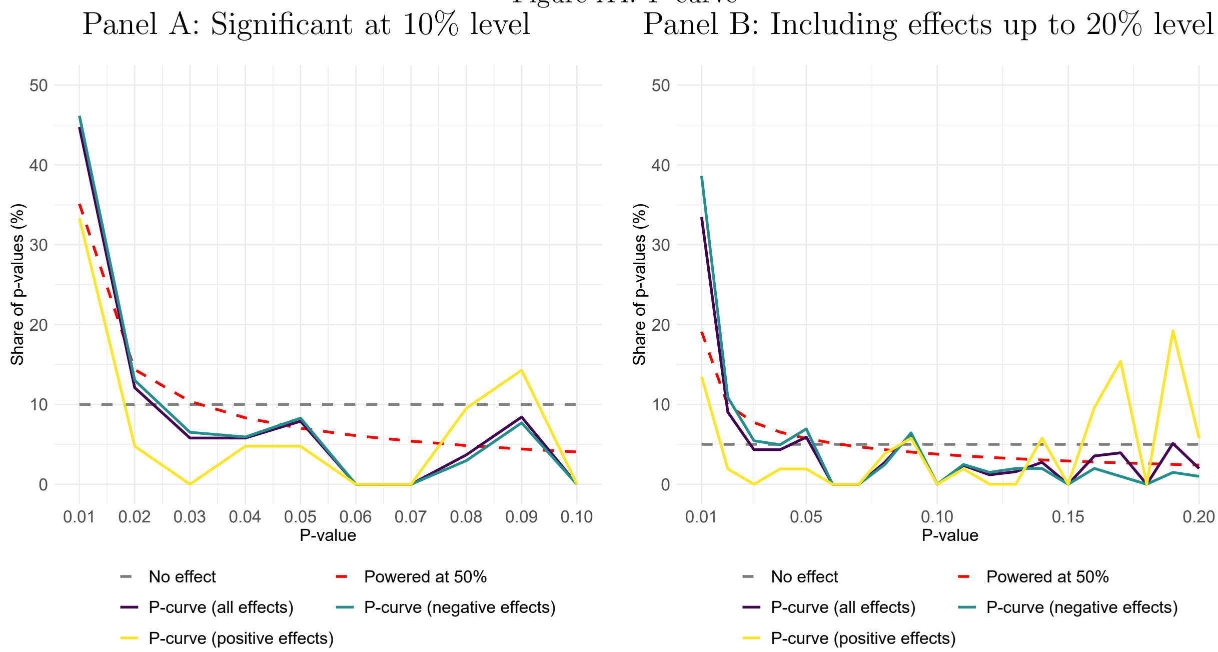


Figure A5: Distribution of P-values

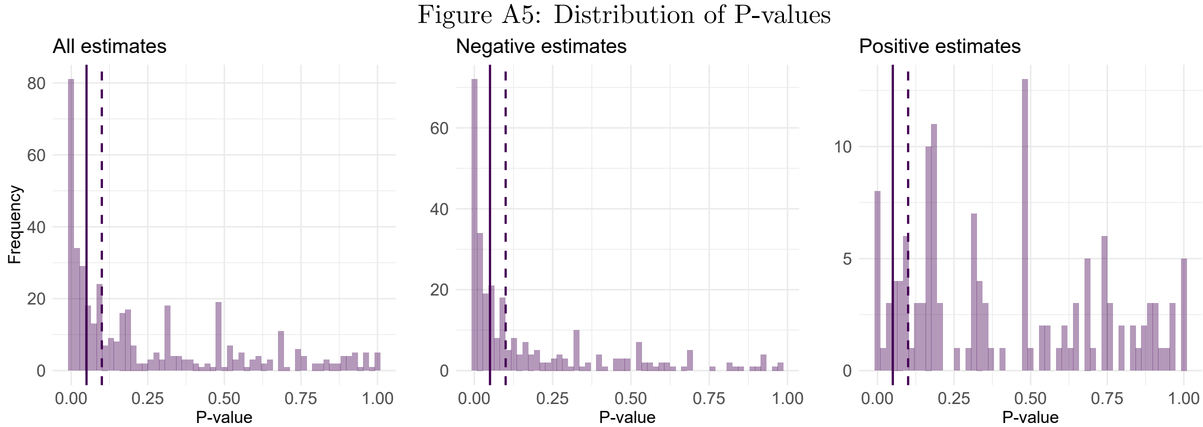


Table A2: Tests Used by Elliott et al. (2021)

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 p-values of six different statistical tests with 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 10% significance level.

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 at the same time, it increases the number of observations we can work with from 422 to 722. To this end, we use the PCCs 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 A3 and A4). The estimates confirm the preference of the publication selection for estimates that are just significant at 10% level and negative. Using visual test in a form of a funnel plot, we next plot the distribution of PCCs and associated t-statistics (Figure A6. The funnel plot of PCCs is visibly left-skewed, confirming the presence for negative publication bias. The funnel plot of t-statistics lays additional support to this view and adds that positive estimates, if published, are mostly not statistically significant. This conforms to our claim that the publication selection exists and is driven by both, selection of sign and statistical significance.

Table A3: 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

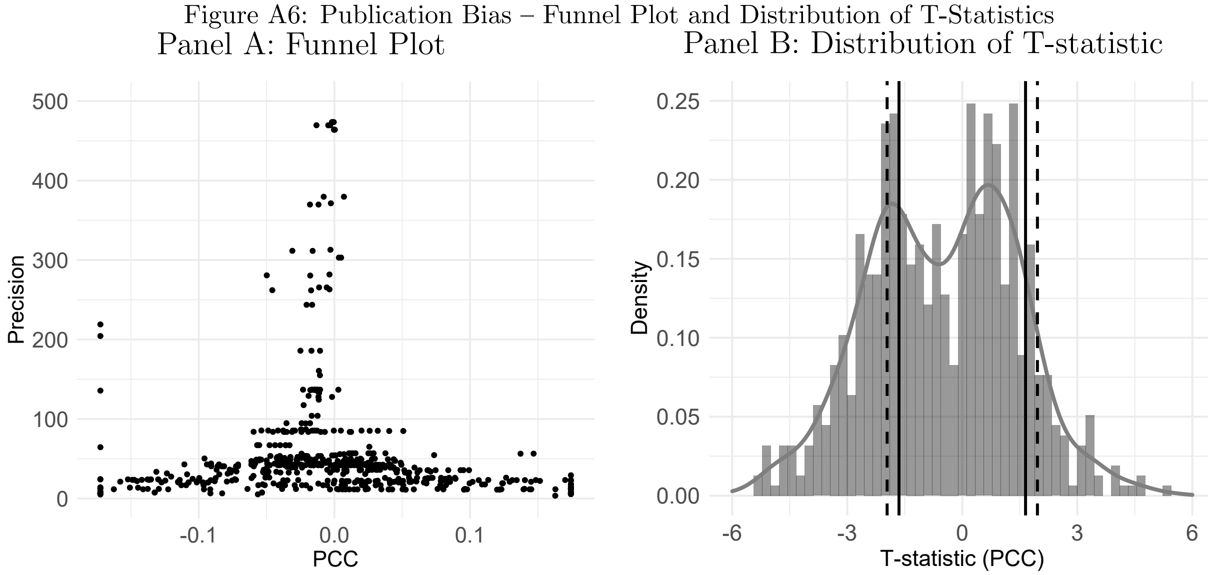
Note: The table presents the results of regression of equation (2) using all collected estimates transformed to PCCs. The standard errors, reported in parentheses, are clustered at both the level of the study and type of the borrower-based measure used in the primary study. OLS – ordinary least squares. 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 A4: 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 regression of equation (2) using all collected estimates transformed to PCCs and extended by additional dummy variables for collected elasticities significant at 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 type of the 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 A6: Publication Bias – Funnel Plot and Distribution of T-Statistics



Note: Panel A: Precision is calculated as an inverse of standard error of the PCC. In the absence of publication bias the funnel should be symmetrical 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 value 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.