The Calvo Parameter Revisited: An Unbiased Insight

Ali Elminejad*

Institute of Economic Studies, Charles University

Abstract

This study provides a meta-analysis of the Calvo parameter estimated within the new Keynesian Phillips curve using a data set of 509 estimates from 40 studies published in a quarter century. Novel linear and non-linear techniques suggest publication bias distorting the reported estimates towards typical values of the Calvo parameter used for calibration. Moreover, Bayesian model averaging results indicate that the reported estimates are systematically affected by various aspects of research design, particularly the choice of forcing variable in the NKPC, instrument selection, and authors' affiliation.

Keywords: Calvo parameter, New Keynesian Phillips Curve, Meta-analysis, Publication bias, Bayesian model averaging

JEL: C11, C83, E31

*Institute of Economic Studies, Faculty of Social Sciences, Charles University Opletalova 26, 110 00, Prague, Czech Republic (email: ali.elminejad@fsv.cuni.cz). Elminejad acknowledges support from the Czech Science Foundation grant "Spurious Precision in Meta-Analysis of Social Science Research" (#23-05227M).

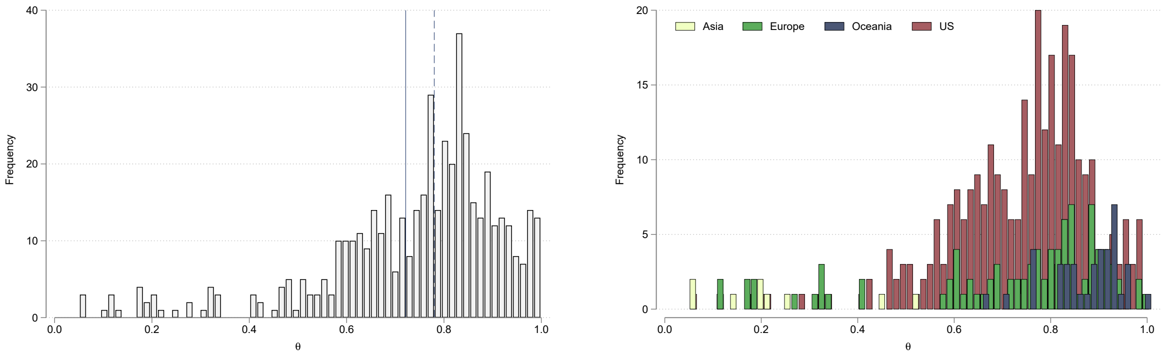
1 Introduction

Standard Calvo-based New Keynesian Phillips Curve (NKPC) is one of the central components in dynamic macroeconomic modeling. According to Calvo (1983), there is a constant probability $(1 - \theta)$ that in each period, a typical firm adjusts its price, and its price remains unchanged with probability θ, which is usually referred to as the Calvo parameter or price rigidity. Empirical examinations of the NKPC based on the Calvo pricing model result in a wide range of values. For example, Galí and Gertler (1999) and Galí et al. (2001) find price rigidity between 0.42 to 0.92 within the estimated NKPC. However, for calibration, researchers usually rely on the vast body of literature suggesting typical values such as 0.75 (average price duration of 4 quarters). For instance, Smets and Wouters (2003), as one of the influential studies in DSGE modeling, uses the typical value of 0.75 a priori for the Calvo parameter.

Two natural questions arise facing the Calvo parameter: First, is the parameter value consistent with the microeconomic data? Second, how does the estimated/calibrated parameter differ from the rest of the reported values in the literature? Extensive literature addresses the first question by comparing estimates based on the Calvo pricing model and microeconomic evidence. Alvarez and Burriel (2010) show that the standard Calvo model fails to capture the distribution of price durations found in microeconomic data. In contrast, Dufour et al. (2010) show that conditional on instrument selections, the price durations estimated by the Calvo-based NKPC are consistent with the US micro data. However, Nakamura and Steinsson (2013) argue that relying solely on the frequency of price changes might be misleading without considering other factors such as sales and cross-sectional heterogeneity. This paper mainly focuses on the second question by conducting a meta-study of a quarter-century literature. Meta-studies have become a widely accepted practice in economics since they are crucial in explaining the variation of results between individual studies (see, e.g., Chetty et al., 2013; Gechert et al., 2022 and Havranek et al., 2022). Similarly, this paper studies how different sources of heterogeneity affect the Calvo parameter estimated within the structural NKPC. To do so, I use a dataset of 509 reported estimates from 40 studies over the last 24 years. The results obtained from various techniques imply the presence of publication bias in the literature. Furthermore, using the Bayesian averaging model (BMA), I show that choice of forcing variables, authors' affiliation, and a set of research characteristics systematically affect the estimates of the Calvo parameter. To my knowledge, this is the first meta-analysis investigating the sources of variation among estimated Calvo parameters.

2 Data

Figure 1: Patterns in the data



Notes: The solid line is the mean estimate, and the dashed line denotes the median estimate reported in primary studies. Outlier estimates (i.e., negative or larger than 1) are excluded from the sub-figures.

I use the Google Scholar search engine to find relevant estimates of the Calvo parameter in the literature. This database provides a powerful tool for full-text search. An online appendix provides details on the search process for collected estimates, which is consistent with the current protocol for meta-analysis (Havránek et al., 2020). The final dataset used in this paper covers 24 years of research from 1999 to 2022. It includes 509 estimates from 40 primary studies. All collected parameters are estimated within the NKPC equation. The Calvo-based NKPC is typically given by:

$$\pi_t = \beta E[\pi_{t+1}] + \lambda mc_t, \tag{1}$$

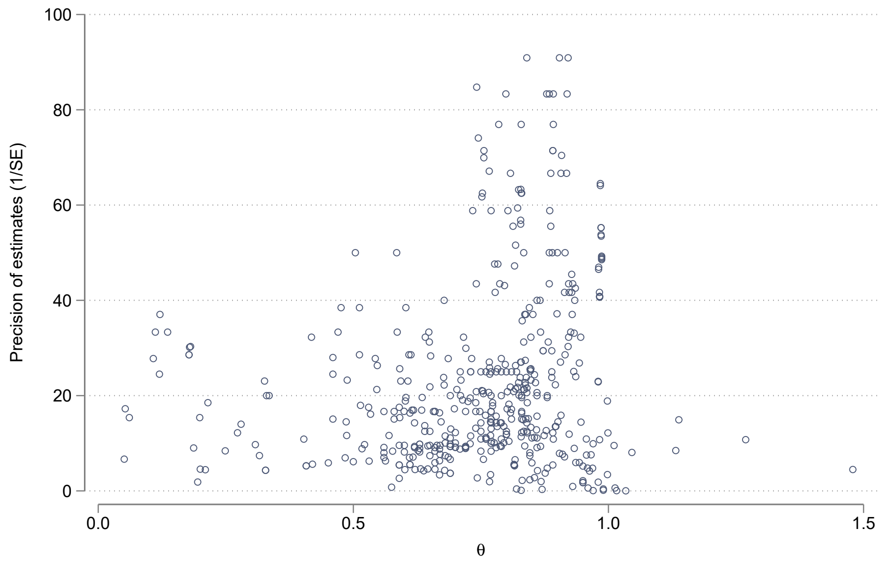
where β is a subjective discount factor, mc_t is real marginal costs, and $\lambda = (1-\theta)(1-\beta\theta)/\theta$. Hence, collected values of θ are obtained by structural estimates of the NKPC. Moreover, estimates are collected with their corresponding standard errors. Therefore, estimates reported without standard errors are excluded from the dataset. In addition to reported estimates and standard errors, the dataset includes 26 extra explanatory variables reflecting the framework in which estimates are reported: data characteristics, model specifications, estimation techniques, and publication characteristics. The final dataset consists of more than 14,000 manually collected data points. The search termination date is December 31, 2022. As of the search termination date, all the studies received 9679 citations combined, indicating the importance of primary studies. Table C1 in the online appendix provides more details on explanatory variables.

The left-hand side of Figure 1 illustrates the distribution of reported estimates in the literature. Overall, estimates are concentrated mainly around 0.70 and 0.85. However, there are several outliers on both sides of the distribution. Therefore, the data are winsorized at the 5% level. The mean point estimate is 0.72 (solid line), marginally smaller than the median, 0.78 (dashed line). The mean estimate is slightly lower than the typical value (0.75) used in the calibration of the dynamic model. The right-hand side of Figure 1 pictures the differences in the distribution of estimates from different regions. There is substantial variation among estimates if we consider different regions, which are not necessarily consistent with the microeconomic data. In the case of the US, the mean estimate and implied average price duration are 0.74 and 3.8 quarters, respectively. These numbers are in agreement with the part of the empirical literature on the US data (Nakamura and Steinsson, 2013; Cravino et al., 2020. Additionally, the mean point estimate based on European data is 0.70, which implies an average price duration of around ten months. This value is inconsistent with some of the microeconomic evidence from the euro area (Alvarez et al., 2006). However, this mean estimate is in line with more recent microeconomic evidence from the euro area. For example, in a recent study, Gautier et al. (2022) find an average price duration in 11 countries in the euro area between 3.39 and 5.15 quarters, depending on the inclusion of sales in the data, which is partially consistent with the mean estimate of European data.

3 Publication Bias

Publication bias significantly affects reported estimates in different fields of science, including economics. Researchers systematically tend to report estimates that are statistically significant and avoid estimates that are either insignificant or with a wrong sign. Hence, one can interpret the relationship between reported estimates and their standard errors as publication bias. Figure 2 shows the relationship between the reported Calvo parameters and their corresponding precision (the inverse of standard error). Without publication bias, the scatter plot (funnel plot) should form a symmetric inverted funnel since the most precise estimates would be around the average effect, and the estimates with lower precision would be more dispersed. Therefore, since there is a noticeable asymmetry, this visual tool suggests the presence of publication bias in the literature.

Figure 2: Funnel plot suggests publication bias



Notes: In the absence of publication bias, the plot should resemble a symmetric inverted funnel. Outliers are excluded from the figure but included in the analysis.

Relying solely on a visual tool in which we assume a linear relationship between estimates and their precision is insufficient to conclude publication bias. Regressing estimates against their standard errors, one can extend assessing the asymmetry of the funnel plot to a regression-based test:

$$\hat{\theta}_{ij} = \theta_0 + \beta \cdot SE(\hat{\theta}_{ij}) + \epsilon_{ij}, \tag{2}$$

where $\hat{\theta}_{ij}$ is the i^{th} reported estimate in the j^{th} study and $SE(\hat{\theta}_{ij})$ is the corresponding standard error. In this regression setting, β denotes the size of publication bias, and the intercept can be interpreted as the mean value of the estimate corrected for publication bias. Panel A in Table 1 reports the regression results based on different specifications. Since the original regression is subject to heteroskedasticity, both sides of Equation 2 are divided by standard errors to give more weights to more precise estimates, which yields a weighted least squares estimator. Besides, standard errors are clustered at the study level since estimates within a study are not independent. The weighted estimator and additional specifications (except the study fixed effect specification) imply publication bias in estimating the Calvo parameter. Furthermore, the results indicate that if we exclude systematic publication bias, the mean corrected for bias will vary between 0.75 and 0.90, depending on the specification. This variation means that the average price rigidity can be 5% to 25% larger, which consequently implies an average price duration of up to twice a longer period (8 quarters) than what the mean estimate in the literature suggests.

Table 1: Linear and non-linear tests

Panel A: Linear tests	WLS	FE	BE	Study
Standard error	-1.749***	0.380	-3.289***	-2.205***
(*publication bias*)	(0.503)	(0.916)	(0.913)	(0.529)
	[-2.986, -0.735]			[-3.310, -1.130]
Constant	0.840***	0.754***	0.879***	0.842***
(*mean beyond bias*)	(0.022)	(0.037)	(0.025)	(0.0187)
	[0.790, 0.887]			[0.801, 0.881]
Implied duration (quarters)	6.250	4.065	8.264	6.329
Observations	509	509	509	509
Studies	40	40	40	40
Panel B: Non-linear tests	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.800***	0.785***	0.800***	0.819***
	(0.022)	(0.009)	(0.007)	(0.030)
Implied duration (quarters)	5.000	4.651	5.000	5.525
Observations	509	509	509	509
Studies	40	40	40	40

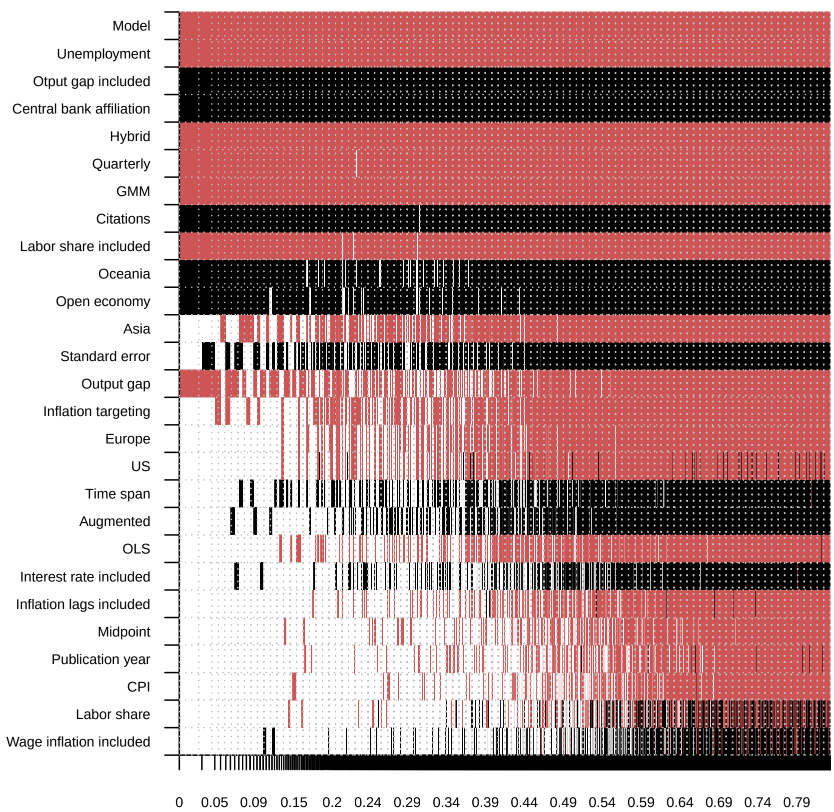
Notes: Panel A presents the results of Equation 2. WLS = weighted least squares. FE = study fixed effects. Study = estimates are weighted by the inverse of the number of estimates reported per study. Standard errors are clustered at the study level; 95% confidence intervals from wild bootstrap clustering are reported in square brackets, if applicable. Panel B presents the mean effect corrected for publication bias using non-linear techniques. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The formal linear tests assume a robust linear relationship between the reported estimates and the standard errors. However, several studies argue that this relationship is not necessarily linear (Andrews and Kasy, 2019). Relaxing this assumption, I use four non-linear techniques to investigate publication bias. These methods usually assume that the linear correlation between the effect and its standard error is distorted by crossing different precision thresholds. Panel B in Table 1 reports the results of the different non-linear techniques. The results are consistent with linear regressions, as they yield a mean beyond bias (between 0.76 and 0.82) larger than the mean reported estimate in the literature. Similar to these results, Meenagh et al. (2022) show that in the case of Bayesian methods, estimates of price rigidity are biased toward the adopted priors, which are usually close to the mean (common) estimate in the literature. The online appendix provides details on non-linear methods and additional results from different subsamples for both linear and non-linear techniques.

4 Heterogeneity

The first set of results indicates the effect of publication bias on estimates. However, publication bias may be the product of heterogeneity among estimates. To address heterogeneity, 26 additional explanatory variables are used that reflect various aspects of studies in which estimates are reported. The online appendix provides more details about explanatory variables.

Figure 3: Model inclusion in Bayesian model averaging



Notes: The response variable is the Calvo parameter estimated within the NKPC. The columns denote individual models; variables are sorted by posterior inclusion probability in descending order. The horizontal axis denotes the cumulative posterior model probabilities. The estimate is based on the unit information prior (UIP) recommended by Eicher et al. (2011) and the dilution prior suggested by George (2010), which takes into account collinearity. Black (darker in grayscale) = the variable has a positive estimated sign. Red (lighter in grayscale) = the variable has a negative estimated sign. No color = the variable is excluded from the given model. Table 3 presents a detailed description of all variables. The numerical results are reported in Table 2.

The first option for investigating heterogeneity is a simple OLS test to regress the reported estimates on the set of explanatory variables. This simple regression, however, does not address the issue of model uncertainty. To this end, I use Bayesian model averaging (BMA) to capture model uncertainty. Using various subsets of explanatory variables, BMA runs multiple regressions and ranks models' relative performance by their posterior model probabilities (PMP). Moreover, posterior inclusion probability (PIP) for each variable indicates the sum of posterior model probabilities (PMP) of the models in which the variable is included. Since, in our case, the number of models visited by BMA is significantly large (2^{27}), I apply the birth-and-death Markov chain Monte Carlo (BDMCMC) algorithm proposed by Stephens (2000), which includes models with the highest PMP. I use the `bms` package developed by Zeugner and Feldkircher (2015). Considering the collinearity of the variables included in each model, I employ the dilution prior, suggested by George (2010), in the benchmark specification. The online appendix provides more details on BMA methods. Moreover, based on the benchmark BMA results, I run a frequentist OLS check, including only variables with PIPs higher than 0.5 obtained from the benchmark BMA specification. Lastly, using Mallow's weights (Hansen, 2007) and the orthogonalization of covariate space (Amini and Parmeter, 2012), I apply frequentist model averaging (FMA), which assumes that explanatory variables are fixed and does not rely on probabilistic information based on prior knowledge. In a comprehensive and insightful study, Steel (2020) discusses the BMA and FMA methods in greater detail.

Figure 3 illustrates the results of the benchmark BMA specification. In addition, the left-hand panel in Table 2 reports the corresponding numerical result. Based on these results, fourteen variables with a PIP larger than 0.5 systematically affect the size of the estimates. Among data characteristics, not surprisingly, the region in which data is taken significantly affects the size of estimated parameters. In addition, studies conducted in countries with an inflation-targeting monetary policy tend to report smaller parameter values. As an intuitive result, model specifications significantly impact reported estimates. Although the choice of inflation measure seems not to have a systematic effect on the estimated parameters, the magnitude of estimates is sensitive to the choice of forcing variable.

Unemployment and output gaps are associated with lower values of the Calvo parameter. Furthermore, BMA results indicate that using the GMM estimator to account for endogeneity could systematically result in smaller estimates. Similar to the forcing variable, the output gap among instruments is systematically associated with smaller estimates. The results also suggest that higher citations are associated with larger estimates. Likewise, studies with at least one author affiliated with a central bank tend to report larger estimates. Finally, the results of frequentist OLS and FMA checks are generally consistent with the benchmark BMA findings. In addition to crucial variables highlighted in the Bayesian setting, FMA results suggest that the estimates are sensitive to all regional data as well as to the OLS method. More details and robustness checks are provided in the online appendix.

Table 2: Explaining heterogeneity

	BMA			OLS			FMA		
Variable	Post. Mean	Post. SD	PIP	Coeff.	S.E.	P-val.	Coeff.	S.E.	P-val.
Constant	0.989	N.A.	1.000	0.941	0.078	0.000	1.127	0.098	0.000
Standard error	0.143	0.141	0.595	0.217	0.269	0.426	0.294	0.102	0.004
Data characteristics									
Time span	0.017	0.027	0.365				0.035	0.024	0.156
Midpoint	-0.002	0.006	0.171				-0.013	0.011	0.218
Quarterly	-0.205	0.063	0.988	-0.197	0.078	0.015	-0.233	0.059	0.000
Inflation targeting	-0.045	0.053	0.511	-0.060	0.047	0.215	-0.101	0.038	0.007
US	-0.075	0.100	0.444				-0.183	0.053	0.001
Europe	-0.079	0.102	0.467				-0.193	0.054	0.000
Oceania	0.143	0.085	0.822	0.233	0.069	0.002	0.095	0.055	0.088
Asia	-0.114	0.109	0.652	-0.057	0.086	0.509	-0.184	0.063	0.003
Specifications									
Hybrid	-0.065	0.020	0.984	-0.074	0.031	0.022	-0.056	0.018	0.002
Open economy	0.068	0.043	0.806	0.098	0.042	0.027	0.076	0.033	0.021
Model	-0.177	0.033	1.000	-0.159	0.079	0.052	-0.178	0.035	0.000
Augmented	0.021	0.033	0.382				0.059	0.026	0.023
CPI	-0.003	0.011	0.166				-0.015	0.020	0.450
Labor share	0.000	0.020	0.163				-0.019	0.043	0.661
Unemployment	-0.361	0.065	1.000	-0.346	0.119	0.006	-0.386	0.068	0.000
Output gap	-0.045	0.053	0.520	-0.058	0.081	0.476	-0.053	0.057	0.355
Estimation techniques									
OLS	-0.023	0.045	0.293				-0.091	0.049	0.063
GMM	-0.092	0.032	0.967	-0.097	0.041	0.023	-0.089	0.030	0.003
Inflation lags included	-0.021	0.047	0.254				-0.100	0.049	0.041
Labor share included	-0.095	0.033	0.957	-0.108	0.032	0.002	-0.072	0.028	0.010
Output gap included	0.120	0.027	0.999	0.136	0.037	0.001	0.113	0.026	0.000
Interest rate included	0.010	0.024	0.243				0.017	0.031	0.586
Wage inflation included	0.002	0.011	0.142				0.009	0.024	0.713
Publication characteristics									
Publication year	-0.004	0.012	0.177				-0.003	0.019	0.891
Central bank affiliation	0.096	0.026	0.997	0.091	0.041	0.032	0.090	0.029	0.002
Citations	0.028	0.010	0.946	0.028	0.014	0.046	0.034	0.011	0.002
Observations	509			509			509		
Studies	40			40			40		

Notes: The response variable is the Calvo parameter estimated within the NKPC. SD = standard deviation, PIP = posterior inclusion probability, SE = standard error. The left panel applies BMA based on the UIP g prior and the dilution prior (Eicher et al. 2011; George 2010). The middle panel reports a frequentist check using OLS, which includes variables with PIPs greater than 0.50 in the benchmark BMA. Standard errors in the frequentist check are clustered at the study level. To conduct the frequentist model averaging, reported on the right panel, we use Mallow's weights by Hansen (2007) and the orthogonalization of the covariate space suggested by Amini and Parmeter (2012). Table C1 in the online appendix presents a detailed description of all variables.

5 Conclusion

This paper provides a meta-analysis of the literature on Calvo-based NKPC estimates. The results based on a dataset of 509 reported estimates from 40 studies suggest that publication bias is present in the literature, distorting the reported estimates towards more orthodox values of the Calvo parameter. The linear and non-linear techniques suggest that the implied average price durations, after correcting for publication bias, exhibit some discrepancies with microeconomic data evidence.

Moreover, the benchmark Bayesian model averaging results show that model specifications, in particular the choice of forcing variables in the NKPC, play a significant role in determining the Calvo parameter value. Surprisingly, no evidence indicates that the inflation measure is systematically correlated with the magnitude of estimates. Similarly, the estimated parameters are sensitive to instrument selection. Finally, in addition to using quarterly data, the central bank's inflation targeting strategy tends to shrink the value of estimated parameters. On the other hand, the results indicate that central bank affiliation is positively associated with larger estimates of the Calvo parameter. Robustness checks are also in line with the findings of the benchmark BMA setting. These results provide a complementary set of helpful information to calibrate and estimate the Calvo parameter. Researchers may use an unbiased Calvo parameter to calibrate within the empirical Calvo-based NKPC, based on the context of their research (e.g., the choice of proxy for marginal costs and the region where data are obtained). Similarly, the results are helpful for comparative analyzes of the estimated NKPC. Further studies can extend the framework of this paper by investigating other aspects of research design to estimate the Calvo parameter absent from this paper.

Acknowldgments

Elminejad acknowledges support from the Czech Science Foundation grant "Spurious Precision in Meta-Analysis of Social Science Research" (#23-05227M).

References

Alvarez, L. J. and Burriel, P. (2010). Is a Calvo price setting model consistent with individual price data? *The BE Journal of Macroeconomics*, 10(1).

Alvarez, L. J., Dhyne, E., Hoeberichts, M., Kwapil, C., Le Bihan, H., Lünnemann, P., Martins, F., Sabbatini, R., Stahl, H., Vermeulen, P., et al. (2006). Sticky prices in the Euro area: a summary of new micro-evidence. *Journal of the European Economic Association*, 4(2-3):575–584.

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

Andrews, I. and Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, 109(8):2766–94.

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

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2013). Does indivisible labor explain the difference between micro and macro elasticities? a meta-analysis of extensive margin elasticities. *NBER Macroeconomics Annual*, 27(1):1–56.

Cravino, J., Lan, T., and Levchenko, A. A. (2020). Price stickiness along the income distribution and the effects of monetary policy. *Journal of Monetary Economics*, 110:19–32.

Dufour, J.-M., Khalaf, L., and Kichian, M. (2010). On the precision of Calvo parameter estimates in structural NKPC models. *Journal of Economic Dynamics and Control*, 34(9):1582–1595.

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

Furukawa, C. (2021). Publication bias under aggregation frictions: from communication model to new correction method. Working paper, MIT, mimeo.

Galí, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222.

Galí, J., Gertler, M., and Lopez-Salido, J. D. (2001). European inflation dynamics. *European Economic Review*, 45(7):1237–1270.

Gautier, E., Conflitti, C., Faber, R. P., Fabo, B., Fadejeva, L., Jouvanceau, V., Menz, J.-O., Messner, T., Petroulas, P., Roldan-Blanco, P., et al. (2022). New facts on consumer price rigidity in the euro area.

Gechert, S., Havranek, T., Irsova, Z., and Kolcunova, D. (2022). Measuring capital-labor substitution: The importance of method choices and publication bias. *Review of Economic Dynamics*, 45:55–82.

George, E. I. (2010). Dilution priors: Compensating for model space redundancy. In *Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown*, pages 158–165. Institute of Mathematical Statistics.

Hansen, B. E. (2007). Least squares model averaging. *Econometrica*, 75(4):1175–1189.

Havranek, T., Irsova, Z., Laslopova, L., and Zeynalova, O. (2022). Publication and attenuation biases in measuring skill substitution. *The Review of Economics and Statistics*, 1(1):1–37.

Havránek, T., Stanley, T., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., and Van Aert, R. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, 34(3):469–475.

Ioannidis, J. P., Stanley, T. D., and Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605):F236–F265.

Meenagh, D., Minford, P., and Wickens, M. R. (2022). The macroeconomic controversy over price rigidity—how to resolve it and how bayesian estimation has led us astray. *Open Economies Review*, pages 1–14.

Nakamura, E. and Steinsson, J. (2013). Price rigidity: Microeconomic evidence and macroeconomic implications. *Annual Review of Economics*, 5(1):133–163.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European economic association*, 1(5):1123–1175.

Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature*, 58(3):644–719.

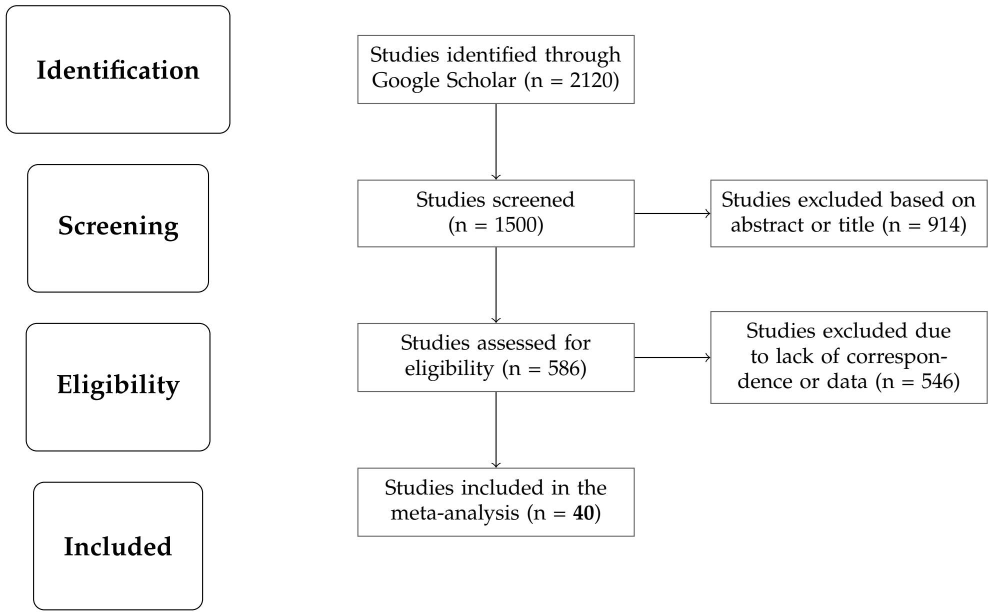
Stephens, M. (2000). Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. *Annals of Statistics*, 28(1):40–74.

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

Appendices (for online publication)

A Literature Search

Figure A1: PRISMA flow diagram

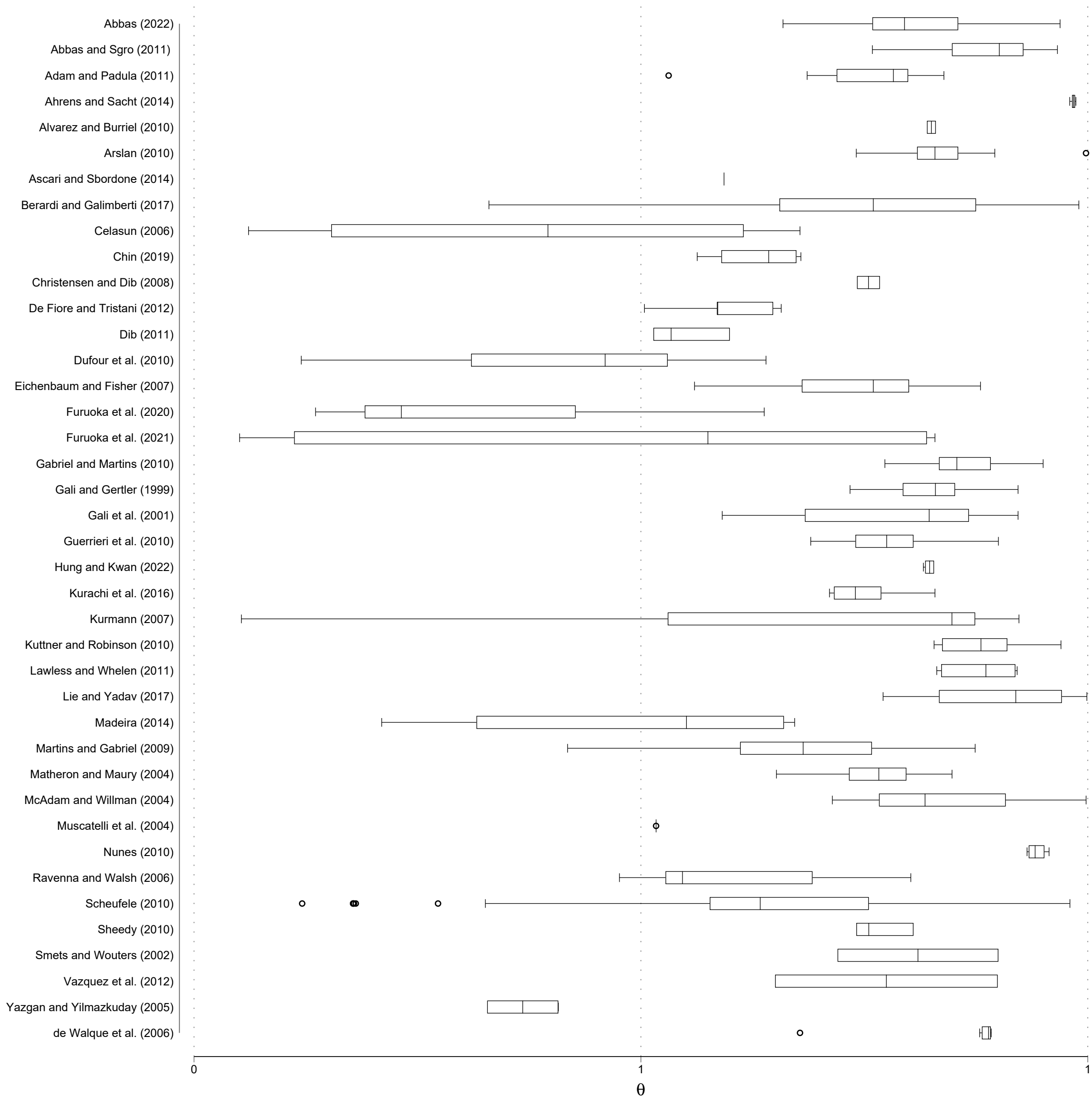


Notes: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is an evidence-based set of items for reporting in systematic reviews and meta-analyses. More details on PRISMA and reporting standards of meta-analysis in general are provided by Havránek et al. (2020).

Table A1: Studies used in the meta-analysis

Abbas (2022)	Galí et al. (2001)
Abbas and Sgro (2011)	Guerrieri et al. (2010)
Adam and Padula (2011)	Hung and Kwan (2022)
Ahrens and Sacht (2014)	Kurachi et al. (2016)
Alvarez and Burriel (2010)	Kurmann (2007)
Arslan (2010)	Kuttner and Robinson (2010)
Ascari and Sbordone (2014)	Lawless and Whelan (2011)
Berardi and Galimberti (2017)	Lie and Yadav (2017)
Celasun (2006)	Madeira (2014)
Chin (2019)	Martins and Gabriel (2009)
Christensen and Dib (2008)	Matheron and Maury (2004)
Fiore and Tristani (2013)	McAdam and Willman (2004)
Walque et al. (2006)	Muscatelli et al. (2004)
Dib (2011)	Nunes (2010)
Dufour et al. (2010)	Ravenna and Walsh (2006)
Eichenbaum and Fisher (2007)	Scheufele (2010)
Furuoka et al. (2020)	Sheedy (2010)
Furuoka et al. (2021)	Smets and Wouters (2002)
Gabriel and Martins (2010)	Vázquez et al. (2012)
Galí and Gertler (1999)	Yazgan and Yilmazkuday (2005)

Figure A2: Variation of the estimates within and between studies



B Additional results for publication bias

B.1 Linear tests

Table B1: Linear funnel asymmetry tests: GDP deflator and CPI

Panel A: GDP deflator	WLS	FE	BE	Study
Standard error	-1.572***	-0.023	-3.356***	-2.161***
(*publication bias*)	(0.566)	(1.018)	(1.126)	(0.630)
	[-2.893, -0.216]			[-3.512, -0.839]
Constant	0.829***	0.771***	0.878***	0.836***
(*mean beyond bias*)	(0.020)	(0.038)	(0.033)	(0.027)
	[0.775, 0.868]			[0.768, 0.897]
Observations	353	353	353	353
Studies	28	28	28	28
Panel B: CPI	WLS	FE	BE	Study
Standard error	-2.395**	2.406	-2.868	-2.214**
(*publication bias*)	(1.054)	(1.977)	(1.630)	(1.059)
	[-6.172, 0.141]			[-4.738, 0.362]
Constant	0.886***	0.647***	0.876***	0.852***
(*mean beyond bias*)	(0.077)	(0.099)	(0.043)	(0.023)
	[0.625, 1.08]			[0.791, 0.974]
Observations	156	156	156	156
Studies	13	13	13	13

Notes: Panel A presents the results of funnel asymmetry test for the subset of estimates with GDP deflator as the measure of inflation and Panel B presents the results of the same test when CPI is used. WLS = weighted least squares. FE = study fixed effects. Study = estimates are weighted by the inverse of the number of estimates reported per study. Standard errors are clustered at the study level; 95% confidence intervals from wild bootstrap clustering are reported in square brackets, if applicable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Linear funnel asymmetry tests: labor share and output gap

Panel A: Labor share	WLS	FE	BE	Study
Standard error	-1.331***	-0.051	-3.090***	-1.872***
(*publication bias*)	(0.462)	(0.865)	(1.004)	(0.606)
	[-2.485, -0.336]			[-3.167, -0.540]
Constant	0.837***	0.783***	0.890***	0.851***
(*mean beyond bias*)	(0.0178)	(0.037)	(0.032)	(0.028)
	[0.790, 0.883]			[0.769, 0.910]
Observations	403	403	403	403
Studies	29	29	29	29
Panel B: Output gap	WLS	FE	BE	Study
Standard error	-3.791***	0.387	-3.905**	-3.683***
(*publication bias*)	(1.100)	(2.450)	(1.618)	(1.108)
	[-7.305, 0.014]			[-6.547, -0.099]
Constant	0.844***	0.709***	0.865***	0.849***
(*mean beyond bias*)	(0.022)	(0.079)	(0.036)	(0.021)
	[0.746, 0.911]			[0.780, 0.897]
Observations	45	45	45	45
Studies	12	12	12	12

Notes: Panel A presents the results of funnel asymmetry test for the subset of estimates when the forcing variable is labor share and Panel B presents the results when the output gap is the forcing variable. See Table B1 for details.

Table B3: Linear funnel asymmetry tests: GMM vs other estimators

Panel A: GMM	WLS	FE	BE	Study
Standard error	-1.807***	0.441	-3.958***	-2.186***
(*publication bias*)	(0.576)	(1.062)	(1.184)	(0.607)
	[-3.399, -0.644]			[-3.438, -0.921]
Constant	0.848***	0.754***	0.915***	0.849***
(*mean beyond bias*)	(0.026)	(0.044)	(0.040)	(0.031)
	[0.793, 0.924]			[0.770, 0.915]
Observations	416	416	416	416
Studies	28	28	28	28
Panel B: Other estimators	WLS	FE	BE	Study
Standard error	-1.869**	0.121	-2.802*	-2.573**
(*publication bias*)	(0.936)	(2.122)	(1.416)	(1.067)
	[-4.826, 0.209]			[-5.203, 0.116]
Constant	0.824***	0.754***	0.854***	0.842***
(*mean beyond bias*)	(0.044)	(0.074)	(0.032)	(0.025)
	[0.694, 0.921]			[0.761, 0.900]
Observations	93	93	93	93
Studies	15	15	15	15

Notes: Panel A presents the results of the formal funnel asymmetry test for the subset of parameters estimated by GMM and Panel B presents the results of the other estimators. See Table B1 for details.

Table B4: Linear funnel asymmetry tests: countries

Panel A: US	WLS	FE	BE	Study
Standard error	-0.967***	0.225	-2.146	-1.153**
(*publication bias*)	(0.370)	(0.878)	(1.401)	(0.535)
	[-2.034, -0.150]			[-2.239, -0.073]
Constant	0.798***	0.747***	0.840***	0.785***
(*mean beyond bias*)	(0.027)	(0.038)	(0.047)	(0.026)
	[0.695, 0.841]			[0.724, 0.846]
Implied duration (quarters)	4.950	3.953	6.250	4.651
Observations	303	303	303	303
Studies	25	25	25	25
Panel B: Europe	WLS	FE	BE	Study
Standard error	-3.186***	0.038	-3.895	-2.867***
(*publication bias*)	(0.700)	(3.807)	(2.234)	(0.873)
	[-6.767, -0.775]			[-6.354, -0.758]
Constant	0.891***	0.783***	0.894***	0.880***
(*mean beyond bias*)	(0.026)	(0.127)	(0.044)	(0.022)
	[0.791, 0.952]			[0.791, 0.952]
Implied duration (quarters)	9.174	4.608	9.434	8.333
Observations	93	93	93	93
Studies	10	10	10	10
Panel C: Oceania	WLS	FE	BE	Study
Standard error	2.121***	1.945	6.170	2.783***
(*publication bias*)	(0.514)	(0.881)	(2.620)	(0.959)
	[1.681, 6.653]			[1.376, 6.161]
Constant	0.780***	0.786***	0.599*	0.740***
(*mean beyond bias*)	(0.033)	(0.033)	(0.091)	(0.059)
	[0.614, 1.285]			[0.621, 0.876]
Implied duration (quarters)	4.545	4.673	2.494	3.846
Observations	48	48	48	48
Studies	3	3	3	3
Panel D: Asia	WLS	FE	BE	Study
Standard error	-4.311***	5.185*	-7.580*	-4.129***
(*publication bias*)	(1.625)	(1.921)	(2.760)	(1.509)
	[-16.140, 1.049]			[-13.510, 0.304]
Constant	0.773***	0.263*	0.956***	0.839***
(*mean beyond bias*)	(0.144)	(0.103)	(0.099)	(0.0727)
	[-37.460, 2.210]			[-2.581, 1.595]
Implied duration (quarters)	4.405	1.357	22.727	6.211
Observations	42	42	42	42
Studies	5	5	5	5

Notes: This table reports the results for different regions. See Table B1 for details.

Table B5: Linear funnel asymmetry tests: significant explanatory variables

Panel A: Hybrid NKPC	WLS	FE	BE	Study
Standard error	-1.844***	0.419	-2.933***	-1.986***
(*publication bias*)	(0.549)	(1.256)	(0.937)	(0.649)
	[-3.164, -0.634]			[-3.537, -0.581]
Constant	0.817***	0.722***	0.868***	0.834***
(*mean beyond bias*)	(0.028)	(0.053)	(0.028)	(0.024)
	[0.725, 0.875]			[0.764, 0.897]
Observations	284	284	284	284
Studies	27	27	27	27
Panel B: CB affiliation	WLS	FE	BE	Study
Standard error	-1.852**	-0.837	-3.680**	-2.718***
(*publication bias*)	(0.727)	(1.623)	(1.588)	(0.989)
	[-4.415, -0.332]			[-4.893, -0.635]
Constant	0.844***	0.805***	0.875***	0.848***
(*mean beyond bias*)	(0.035)	(0.062)	(0.039)	(0.027)
	[0.740, 0.922]			[0.752, 0.915]
Observations	221	221	221	221
Studies	21	21	21	21
Panel C: Open economy	WLS	FE	BE	Study
Standard error	-2.252	0.890	-2.619	-1.781
(*publication bias*)	(1.766)	(1.858)	(2.828)	(1.682)
	[-9.362, 1.774]			[-6.820, 1.920]
Constant	0.918***	0.814***	0.883***	0.857***
(*mean beyond bias*)	(0.074)	(0.061)	(0.073)	(0.044)
	[0.752, 1.15]			[0.746, 1.137]
Observations	84	84	84	84
Studies	8	8	8	8
Panel D: Inflation target	WLS	FE	BE	Study
Standard error	-1.126	1.094	-10.62***	-2.245
(*publication bias*)	(1.784)	(1.115)	(2.193)	(1.719)
	[-9.956, 2.098]			[-8.697, 2.142]
Constant	0.850***	0.759***	1.228***	0.834***
(*mean beyond bias*)	(0.050)	(0.046)	(0.093)	(0.078)
	[0.616, 1.188]			[0.501, 1.197]
Observations	82	82	82	82
Studies	8	8	8	8
Panel E: Model estimated	WLS	FE	BE	Study
Standard error	-7.688***	-0.278	-6.643**	-5.287***
(*publication bias*)	(1.340)	(4.547)	(2.148)	(1.307)
	[-12.470, -2.803]			[-9.087, -2.287]
Constant	0.921***	0.729***	0.904***	0.869***
(*mean beyond bias*)	(0.056)	(0.118)	(0.047)	(0.041)
	[0.657, 1.027]			[0.745, 0.996]
Observations	39	39	39	39
Studies	9	9	9	9

Notes: This table reports the results for different regions. See Table B1 for details.

B.2 Non-linear tests

Panel B of Table 1 in the paper and also Tables B6-B9 present the results obtained from non-linear techniques. Ioannidis et al. (2017) propose the Weighted Average of Adequately Powered (WAAP) technique, which considers the estimates when their statistical power is above an 80% threshold. In other words, by using the WAAP technique, we assign a weight to each estimate with adequate power to compute a weighted mean corrected for bias. Furthermore, Andrews and Kasy (2019) suggest the second non-linear method used in this paper. This technique assumes that publication probability changes after crossing conventional t-statistic thresholds. This technique re-weights estimates in the vicinity of the threshold based on how they are present in the literature.

The Endogenous Kink (EK) technique proposed by Bom and Rachinger (2019), is the third non-linear method used in the meta-analysis. This method extends the linear funnel asymmetry test by assuming that the selection of estimates for publication is constrained with particular precision cut-offs in each literature. Finally, Furukawa (2021) develops a stem-based method that considers only the most precise estimates (i.e., the stem of the funnel plot). The method considers both efficiency (increasing in the number of included estimates) and bias (decreasing in the number of included precise estimates) and optimizes the trade-off between them.

Table B6: Non-linear funnel asymmetry tests: GDP deflator and CPI

Panel A: GDP deflator	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.795***	0.792***	0.795***	0.848***
	(0.020)	(0.010)	(0.007)	(0.052)
Observations	353	353	353	353
Studies	28	28	28	28
Panel B: CPI	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.820***	0.749***	0.820***	0.820***
	(0.072)	(0.016)	(0.016)	(0.048)
Observations	156	156	156	156
Studies	13	13	13	13

Notes: This table reports the results of non-linear techniques regarding different inflation measures. Standard errors are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B7: Non-linear funnel asymmetry tests: labor share and output gap

Panel A: Labor share	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.805***	0.789***	0.805***	0.794***
	(0.018)	(0.009)	(0.006)	(0.051)
Observations	403	403	403	403
Studies	29	29	29	29
Panel B: Output gap	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.776***	0.639***	0.776***	0.770***
	(0.025)	(0.034)	(0.020)	(0.035)
Observations	45	45	45	45
Studies	12	12	12	12

Notes: This table reports the results of non-linear techniques regarding different proxies of marginal costs. Standard errors are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8: Non-linear funnel asymmetry tests: GMM vs other estimators

Panel A: GMM	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.804***	0.792***	0.804***	0.823***
	(0.026)	(0.009)	(0.004)	(0.054)
Observations	416	416	416	416
Studies	28	28	28	28
Panel B: Other estimators	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.789***	0.753***	0.789***	0.777***
	(0.042)	(0.023)	(0.014)	(0.032)
Observations	93	93	93	93
Studies	15	15	15	15

Notes: This table reports the results of non-linear techniques regarding GMM and other estimators. Standard errors are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B9: Non-linear funnel asymmetry tests: countries

Panel A: US	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.774***	0.756***	0.774***	0.791***
	(0.025)	(0.010)	(0.007)	(0.063)
Implied duration (quarters)	4.425	4.098	4.425	4.785
Observations	303	303	303	303
Studies	25	25	25	25
Panel B: Europe	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.832***	0.821***	0.831***	0.783***
	(0.035)	(0.019)	(0.016)	(0.036)
Implied duration (quarters)	5.952	5.587	5.917	4.608
Observations	93	93	93	93
Studies	10	10	10	10
Panel C: Oceania	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.826***	0.879***	0.826***	0.846***
	(0.028)	(0.011)	(0.014)	(0.047)
Implied duration (quarters)	5.747	8.264	5.747	6.494
Observations	48	48	48	48
Studies	3	3	3	3
Panel D: Asia	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.654**	0.446***	0.657***	0.588*
	(0.163)	(0.001)	(0.043)	(0.304)
Implied duration (quarters)	2.890	1.805	2.915	2.427
Observations	42	42	42	42
Studies	5	5	5	5

Notes: This table reports the results of non-linear techniques for reported estimates based on different regions. Standard errors are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B10: Non-linear funnel asymmetry tests: significant explanatory variables

Panel A: Hybrid NKPC	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.775***	0.764***	0.775***	0.813***
	(0.028)	(0.017)	(0.009)	(0.083)
Observations	284	284	284	284
Studies	27	27	27	27
Panel B: CB affiliation	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.800***	0.778***	0.800***	0.791***
	(0.029)	(0.012)	(.009)	(0.036)
Observations	221	221	221	221
Studies	21	21	21	21
Panel C: Open economy	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.865***	0.835***	0.865***	0.796***
	(0.054)	(0.017)	(0.154)	(0.065)
Observations	84	84	84	84
Studies	8	8	8	8
Panel D: Inflation targeting	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.823***	0.889***	0.823***	0.834***
	(0.044)	(0.022)	(0.189)	(0.144)
Observations	82	82	82	82
Studies	8	8	8	8
Panel E: Model estimated	Ioannidis et al. (2017)	Andrews and Kasy (2019)	Bom and Rachinger (2019)	Furukawa (2021)
Effect beyond bias	0.783***	0.640***	0.783***	0.809***
	(0.061)	(0.001)	(0.264)	(0.064)
Observations	39	39	39	39
Studies	9	9	9	9

Notes: This table reports the results of non-linear techniques for subgroups of reported estimates based on the most decisive variables obtained from BMA. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Explanatory variables, summary statistics, and additional BMA results

Bayesian model averaging (BMA) method is a natural solution to model uncertainty within the Bayesian setting. Using all possible subsets of explanatory variables, BMA runs numerous regression models and forms a weighted average over all of them. If the set of explanatory variables contains n variables, there will be combinations of 2^n variables and 2^n models.

Defining $\mathcal{P}(M_i)$, $\mathcal{P}(y \mid M_i, X_i)$, and $\mathcal{P}(y \mid X_i)$ as the model prior, the marginal likelihood, and the integrated likelihood, respectively, posterior model probabilities (PMP) are obtained as follows:

$$\mathcal{P}(M_i \mid y, X) = \frac{\mathcal{P}(y \mid M_i, X)\mathcal{P}(M_i)}{\mathcal{P}(y \mid X_n)} \equiv \frac{\mathcal{P}(y \mid M_i, X)\mathcal{P}(M_i)}{\sum_{s=1}^{2N}\mathcal{P}(y \mid M_s, X_s)\mathcal{P}(M_s)},$$

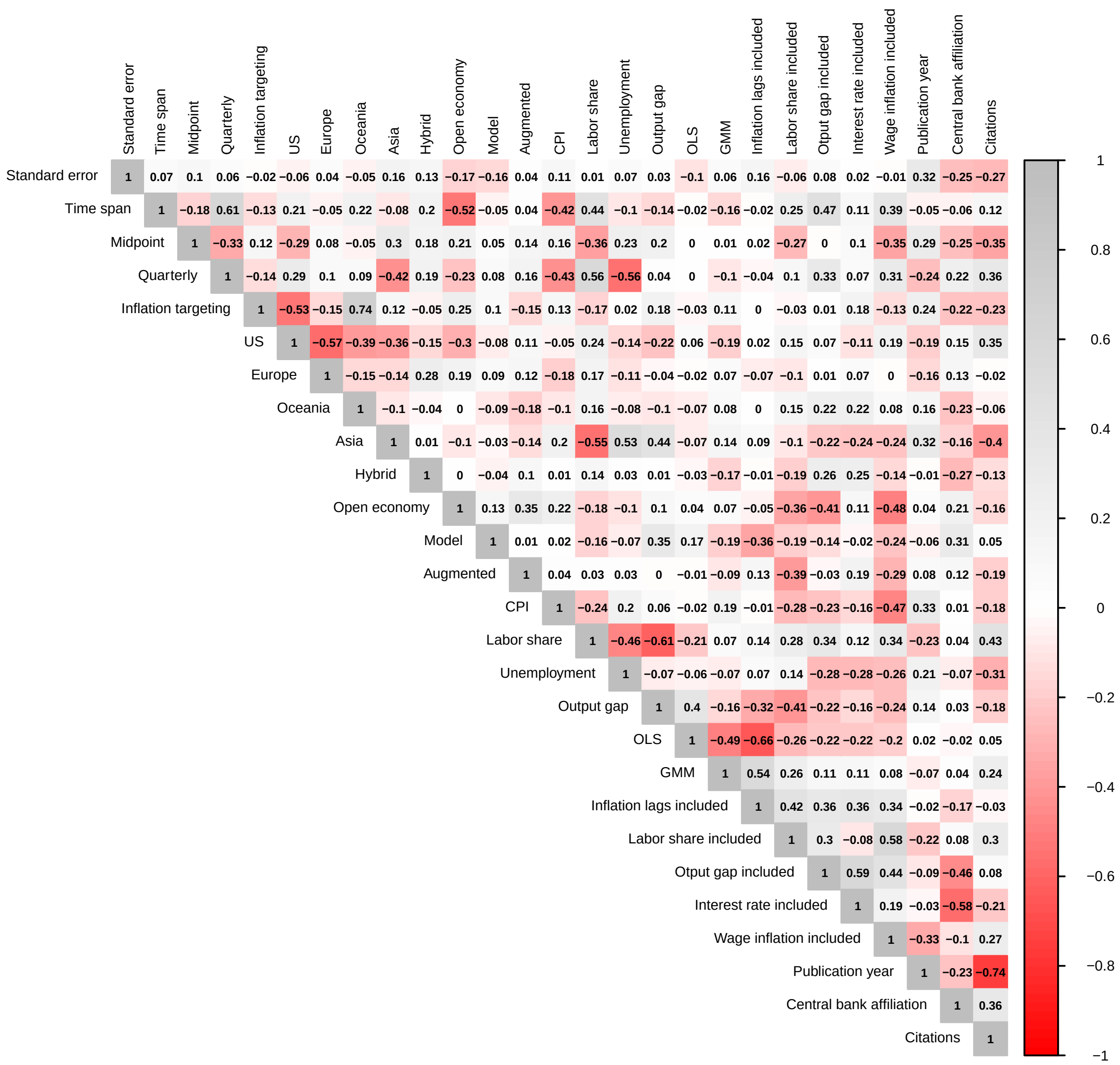
The model weighted posterior distribution for the coefficient θ can be written as:

$$\mathcal{P}(\theta \mid y, X) = \sum_{i=1}^{2^n}\mathcal{P}(\theta \mid M_i, y, X)\mathcal{P}(M_i \mid y, X).$$

The model prior is a key factor in conducting BMA since it reflects the prior beliefs about the model. The benchmark prior is dilution prior suggested by George (2010), which takes into account the collinearity of variables in each model by assigning higher weights to models that exhibit lower collinearity. Additionally, for robustness checks, BRIC g-prior and HQ g-prior are used. The former is the benchmark g-prior for parameters with the beta-binomial model prior, while the latter asymptotically mimics the Hannan-Quinn criterion. Steel (2020) provides a comprehensive and insightful summary of model averaging in economics.

C.1 Explanatory variables

Figure C1: Correlation matrix



Notes: The figure shows Pearson correlation coefficients for the explanatory variables described in Table C1.

Data characteristics. I control for the time span in which the Calvo parameter is estimated. I also control for the frequency of data by including dummy variables indicating whether quarterly data is used. There are dummy variables reflecting the region of the data source used in the estimation: the US, Europe, Oceania, and Asia.

Specifications. Controlling for model specifications, I codify a dummy variable capturing if the estimate is obtained within the hybrid NKPC or a purely forward-looking NKPC setting. Besides, two other dummy variables indicate if the reported estimate is obtained within an open economy setting or an augmented NKPC setting (i.e., the NKPC includes other terms in addition to expected inflation and economic activity). I also codify a dummy variable reflecting if the Calvo parameter is estimated within a model. Estimating the NKPC and, in particular, the Calvo parameter is sensitive to the choice of inflation measurement; see, e.g., (Mavroeidis et al., 2014). Hence, I introduce a dummy variable accounting for CPI and GDP deflator as inflation measurements. As discussed by Galí and Gertler (1999), the choice of a valid proxy for marginal cost can affect the estimated parameters within the NKPC. Three dummy variables control for marginal costs proxies: labor share, unemployment, and the output gap.

Estimation techniques. There are seven dummy variables defined to capture different aspects of estimation methods. Two dummy variables denote the ordinary least squares (OLS) and the generalized method of moments (GMM) methods used in estimating the parameter, which are used for 87% of the estimates in the sample. Moreover, I include five dummy variables reflecting the instruments used in estimating the parameter.

Publication characteristics. There is a variable for the publication year to capture the fact that a recent study is more likely to provide more accurate results since it employs newer theoretical and empirical methods. As a proxy accounting for the ex-post quality of the study, there is an explanatory variable denoting the number of citations of each study. Finally, I codify a dummy variable indicating if at least one of the authors is affiliated with a central bank. This variable helps capture possible workplace bias.

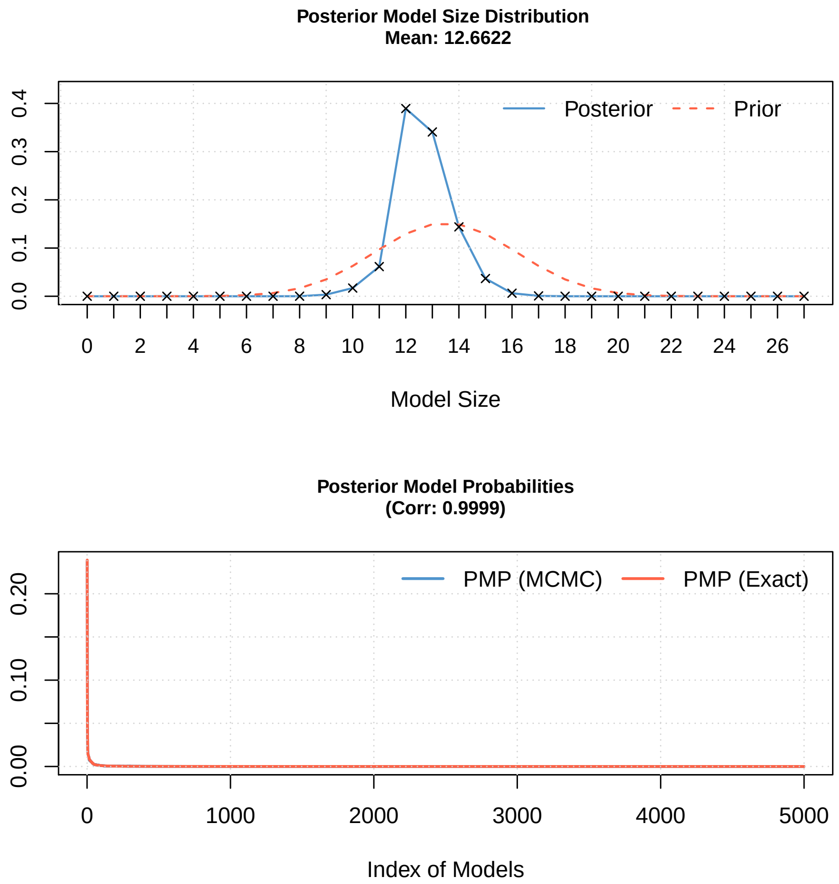
Table C1: Definition and summary statistics of explanatory variables

Variable	Description	Mean	SD	No. papers
θ	The estimated Calvo parameter in the NKPC equation.	0.72	0.25	-
Standard error	The standard error of the estimated coefficient of inflation expectations.	0.31	2.69	-
Data characteristics				
Time span	The logarithm of the data time span used to estimate θ.	3.37	0.52	-
Midpoint	The logarithm of the median year of the data used minus the earliest median year in primary studies.	2.80	0.75	-
Quarterly	= 1 if the data frequency is annual (reference category: monthly/annual).	0.92	0.26	37
Inflation targeting	=1 if the central bank employs an inflation targeting regime during at least half of the estimation period.	0.16	0.37	8
US	= 1 if the estimate is for the U.S. (reference category: other countries).	0.99	0.49	25
Europe	= 1 if the estimate is for European countries (reference category: other countries).	0.18	0.39	10
Oceania	= 1 if the estimate is for Australia and New Zealand countries (reference category: other countries).	0.10	0.29	3
Asia	= 1 if the estimate is for Asian countries (reference category: other countries).	0.08	0.27	5
Specifications				
Hybrid	= 1 if the estimate is from a hybrid NKPC setting (reference category: purely forward-looking NKPC).	0.56	0.50	27
Open economy	= 1 if the estimate is from an open economy specification (reference category: closed economy).	0.16	0.37	8
Model	= 1 if θ is estimated within a model.	0.08	0.27	9
Augmented	= 1 if the NKPC includes other terms in addition to expected inflation and economic activity.	0.24	0.43	10
CPI	= 1 if CPI is the measure of inflation (reference category: GDP deflator).	0.31	0.46	13
Labor share	= 1 if the labor income share (unit labor costs) is a proxy for marginal costs (reference category: other proxies).	0.79	0.41	29
Unemployment gap	= 1 if unemployment is a proxy for marginal costs (reference category: other proxies).	0.05	0.22	2
Output gap	= 1 if output gap is a proxy for marginal costs (reference category: other proxies).	0.09	0.28	12
Estimation techniques				
OLS	= 1 if the ordinary least square (OLS) method is used for the estimation (reference category: other methods).	0.05	0.22	8
GMM	= 1 if the generalized method of moments (GMM) is used for the estimation (reference category: other methods).	0.82	0.39	28
inflation lags included	= 1 if inflation lags are among instruments (reference category: inflation lags not among instruments).	0.91	0.28	30
Labor share included	= 1 if labor income share is among instruments (reference category: labor share not among instruments).	0.65	0.48	20
Output gap included	= 1 if the output gap is among instruments (reference category: Output gap not among instruments).	0.57	0.49	20
Interest rate included	= 1 if the interest rate is among instruments (reference category: interest rate not among instruments).	0.58	0.49	18
Wage inflation included	= 1 if wage inflation is among instruments (reference category: Wage inflation not among instruments).	0.54	0.50	16
Publication characteristics				
Publication year	The logarithm of the publication year of the study minus the publication year of the first primary study.	2.32	0.68	-
Central bank affiliation	= 1 if at least one of the authors is affiliated with a central bank.	0.43	0.50	21
Citations	The logarithm of the number of per-year citations of the study, according to Google Scholar.	1.40	1.51	-

Notes: SD = standard deviation No. papers = the number of papers that capture the dummy variable. The table excludes the definition and summary statistics of the reference categories, which are omitted from the regressions.

C.2 Robustness checks

Figure C2: Model size and convergence for the benchmark BMA model



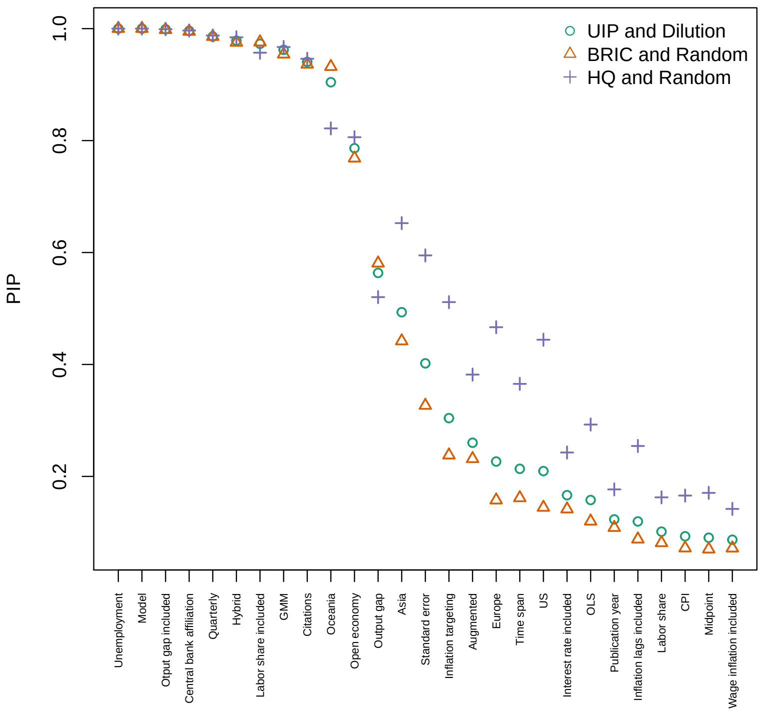
Notes: The figure illustrates the posterior model size distribution and the posterior model probabilities of the weighted BMA exercise reported in Table 2.

Table C2: Alternative BMA priors

	BRIC g-prior			HQ g-prior		
Variable	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	PIP
Constant	0.959	N.A.	1.000	0.957	N.A.	1.000
Standard error	0.074	0.121	0.327	0.090	0.129	0.394
Data characteristics						
Time span	0.007	0.019	0.162	0.009	0.021	0.200
Midpoint	-0.001	0.004	0.070	-0.001	0.004	0.077
Quarterly	-0.198	0.061	0.986	-0.200	0.060	0.989
Inflation targeting	-0.018	0.038	0.238	-0.022	0.041	0.285
US	-0.020	0.063	0.145	-0.026	0.070	0.176
Europe	-0.022	0.064	0.158	-0.028	0.071	0.189
Oceania	0.174	0.061	0.932	0.171	0.065	0.920
Asia	-0.062	0.087	0.442	-0.067	0.090	0.471
Specifications						
Hybrid	-0.067	0.020	0.975	-0.068	0.019	0.989
Open economy	0.066	0.044	0.768	0.069	0.042	0.806
Model	-0.175	0.032	1.000	-0.175	0.032	1.000
Augmented	0.013	0.027	0.232	0.012	0.026	0.238
CPI	-0.001	0.007	0.072	-0.001	0.007	0.074
Labor share	-0.001	0.015	0.081	-0.001	0.015	0.083
Unemployment	-0.366	0.065	1.000	-0.364	0.062	1.000
Output gap	-0.056	0.055	0.581	-0.055	0.054	0.577
Estimation techniques						
OLS	-0.008	0.028	0.120	-0.010	0.031	0.142
GMM	-0.093	0.032	0.954	-0.095	0.030	0.974
Inflation lags included	-0.006	0.026	0.088	-0.006	0.027	0.093
Labor share included	-0.104	0.028	0.976	-0.103	0.027	0.980
Otput gap included	0.124	0.027	0.998	0.125	0.026	1.000
Interest rate included	0.007	0.020	0.142	0.007	0.020	0.153
Wage inflation included	0.002	0.008	0.072	0.002	0.008	0.074
Publication characteristics						
Publication year	-0.003	0.011	0.109	-0.002	0.010	0.100
Central bank affiliation	0.095	0.025	0.995	0.095	0.025	0.998
Citations	0.027	0.010	0.936	0.027	0.009	0.954
Observations	509			509		
Studies	40			40		

Notes: The response variable is the estimated Calvo parameter. SD = standard deviation, PIP = Posterior inclusion probability. The left panel applies BMA based on BRIC g-prior (the benchmark g-prior for parameters with the beta-binomial model prior). The right panel reports the results of BMA based on HQ g-prior, which asymptotically mimics the Hannan-Quinn criterion. Table C1 presents a detailed description of all the variables.

Figure C3: Posterior inclusion probabilities across different prior settings



Notes: UIP and Dilution = priors according to Eicher et al. (2011) and George (2010); BRIC and Random = the benchmark g-prior for parameters with the beta-binomial model prior. The HQ prior asymptotically mimics the Hannan-Quinn criterion.

References

Abbas, S. K. (2022). Asymmetry in the regimes of inflation and business cycles: the new keynesian phillips curve. *Applied Economics*, pages 1–14.

Abbas, S. K. and Sgro, P. M. (2011). New keynesian phillips curve and inflation dynamics in australia. *Economic Modelling*, 28(4):2022–2033.

Adam, K. and Padula, M. (2011). Inflation dynamics and subjective expectations in the united states. *Economic Inquiry*, 49(1):13–25.

Ahrens, S. and Sacht, S. (2014). Estimating a high-frequency new-keynesian phillips curve. *Empirical Economics*, 46(2):607–628.

Alvarez, L. J. and Burriel, P. (2010). Is a Calvo price setting model consistent with individual price data? *The BE Journal of Macroeconomics*, 10(1).

Andrews, I. and Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, 109(8):2766–94.

Arslan, M. M. (2010). Relative importance of sticky prices and sticky information in price setting. *Economic Modelling*, 27(5):1124–1135.

Ascari, G. and Sbordone, A. M. (2014). The macroeconomics of trend inflation. *Journal of Economic Literature*, 52(3):679–739.

Berardi, M. and Galimberti, J. K. (2017). On the initialization of adaptive learning in macroeconomic models. *Journal of Economic Dynamics and Control*, 78:26–53.

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

Celasun, O. (2006). Sticky inflation and the real effects of exchange rate-based stabilization. *Journal of International Economics*, 70(1):115–139.

Chin, K.-H. (2019). New keynesian phillips curve with time-varying parameters. *Empirical Economics*, 57(6):1869–1889.

Christensen, I. and Dib, A. (2008). The financial accelerator in an estimated new keynesian model. *Review of economic dynamics*, 11(1):155–178.

Dib, A. (2011). Monetary policy in estimated models of small open and closed economies. *Open Economies Review*, 22(5):769–796.

Dufour, J.-M., Khalaf, L., and Kichian, M. (2010). On the precision of Calvo parameter estimates in structural NKPC models. *Journal of Economic Dynamics and Control*, 34(9):1582–1595.

Eichenbaum, M. and Fisher, J. D. (2007). Estimating the frequency of price re-optimization in calvo-style models. *Journal of monetary Economics*, 54(7):2032–2047.

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

Fiore, F. D. and Tristani, O. (2013). Optimal monetary policy in a model of the credit channel. *The Economic Journal*, 123(571):906–931.

Furukawa, C. (2021). Publication bias under aggregation frictions: from communication model to new correction method. Working paper, MIT, mimeo.

Furuoka, F., Ling, P. K., Chomar, M. T., and Nikitina, L. (2020). Trade openness and the phillips curve: Evidence from asean countries. *The Singapore Economic Review*, pages 1–25.

Furuoka, F., Pui, K. L., Chomar, M. T., and Nikitina, L. (2021). Is the phillips curve disappearing? evidence from a new test procedure. *Applied Economics Letters*, 28(6):493–500.

Gabriel, V. J. and Martins, L. F. (2010). The cost channel reconsidered: A comment using an identification-robust approach. *Journal of Money, Credit and Banking*, 42(8):1703–1712.

Galí, J. and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222.

Galí, J., Gertler, M., and Lopez-Salido, J. D. (2001). European inflation dynamics. *European Economic Review*, 45(7):1237–1270.

George, E. I. (2010). Dilution priors: Compensating for model space redundancy. In *Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown*, pages 158–165. Institute of Mathematical Statistics.

Guerrieri, L., Gust, C., and López-Salido, J. D. (2010). International competition and inflation: a new keynesian perspective. *American Economic Journal: Macroeconomics*, 2(4):247–80.

Havránek, T., Stanley, T., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., and Van Aert, R. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, 34(3):469–475.

Hung, T. H. and Kwan, Y. K. (2022). Hong kong's new keynesian phillips curve: Sticky information or sticky price? *Pacific Economic Review*, 27(1):42–55.

Ioannidis, J. P., Stanley, T. D., and Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605):F236–F265.

Kurachi, Y., Hiraki, K., Nishioka, S., et al. (2016). Does a higher frequency of micro-level price changes matter for macro price stickiness?: Assessing the impact of temporary price changes. Technical report, Bank of Japan.

Kurmann, A. (2007). Var-based estimation of euler equations with an application to new keynesian pricing. *Journal of Economic Dynamics and Control*, 31(3):767–796.

Kuttner, K. and Robinson, T. (2010). Understanding the flattening phillips curve. *The North American Journal of Economics and Finance*, 21(2):110–125.

Lawless, M. and Whelan, K. T. (2011). Understanding the dynamics of labor shares and inflation. *Journal of Macroeconomics*, 33(2):121–136.

Lie, D. and Yadav, A. S. (2017). Time-varying trend inflation and the new keynesian phillips curve in australia. *Economic Record*, 93(300):42–66.

Madeira, J. (2014). Overtime labor, employment frictions, and the new keynesian phillips curve. *Review of Economics and Statistics*, 96(4):767–778.

Martins, L. F. and Gabriel, V. J. (2009). New keynesian phillips curves and potential identification failures: A generalized empirical likelihood analysis. *Journal of Macroeconomics*, 31(4):561–571.

Matheron, J. and Maury, T.-P. (2004). Supply-side refinements and the new keynesian phillips curve. *Economics Letters*, 82(3):391–396.

Mavroeidis, S., Plagborg-Møller, M., and Stock, J. H. (2014). Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature*, 52(1):124–88.

McAdam, P. and Willman, A. (2004). Supply, factor shares and inflation persistence: Re-examining euro-area new-keynesian phillips curves. *Oxford Bulletin of Economics and Statistics*, 66:637–670.

Muscatelli, V. A., Tirelli, P., and Trecroci, C. (2004). Fiscal and monetary policy interactions: Empirical evidence and optimal policy using a structural new-keynesian model. *Journal of Macroeconomics*, 26(2):257–280.

Nunes, R. (2010). Inflation dynamics: the role of expectations. *Journal of Money, Credit and Banking*, 42(6):1161–1172.

Ravenna, F. and Walsh, C. E. (2006). Optimal monetary policy with the cost channel. *Journal of Monetary Economics*, 53(2):199–216.

Scheufele, R. (2010). Evaluating the german (new keynesian) phillips curve. *The North American Journal of Economics and Finance*, 21(2):145–164.

Sheedy, K. D. (2010). Intrinsic inflation persistence. *Journal of Monetary Economics*, 57(8):1049–1061.

Smets, F. and Wouters, R. (2002). Openness, imperfect exchange rate pass-through and monetary policy. *Journal of Monetary Economics*, 49(5):947–981.

Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature*, 58(3):644–719.

Vázquez, J., María-Dolores, R., and Londoño, J. M. (2012). The effect of data revisions on the basic new keynesian model. *International Review of Economics & Finance*, 24:235–249.

Walque, G. d., Smets, F., and Wouters, R. (2006). Price shocks in general equilibrium: Alternative specifications. *CESifo Economic Studies*, 52(1):153–176.

Yazgan, M. E. and Yilmazkuday, H. (2005). Inflation dynamics of turkey: a structural estimation. *Studies in Nonlinear Dynamics & Econometrics*, 9(1).