HOW TO SOLVE THE PRICE PUZZLE? A META-ANALYSIS

Marek Rusnák
Tomáš Havránek
Roman Horváth

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

446

Working Paper Series 446
(ISSN 1211-3298)

How to Solve the Price Puzzle? A Meta-Analysis

Marek Rusnák
Tomáš Havránek
Roman Horváth

CERGE-EI
Prague, September 2011

EVROPSKÝ SOCIÁLNÍ FOND
PRAHA & EU: INVESTUJEME DO VAŠÍ BUDOUCNOSTI

ISBN 978-80-7343-249-2 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-241-5 (Národohospodářský ústav AV ČR, v.v.i.)

How to Solve the Price Puzzle? A Meta-Analysis*

Marek Rusnák[a,b], Tomáš Havránek[b,c], and Roman Horváth[c]

[a]CERGE-EI, Prague†
[b]Czech National Bank
[c]Charles University, Prague

Abstract

The short-run increase in prices following an unexpected tightening of monetary policy represents a frequently reported puzzle. Yet the puzzle is easy to explain away when all published models are quantitatively reviewed. We collect and examine about 1,000 point estimates of impulse responses from 70 articles using vector autoregressive models to study monetary transmission in various countries. We find some evidence of publication selection against the price puzzle in the literature, but our results also suggest that the reported puzzle is mostly caused by model misspecifications. Finally, the long-run response of prices to monetary policy shocks depends on the characteristics of the economy.

Keywords: Monetary policy transmission; Price puzzle; Meta-analysis; Publication selection bias

JEL Codes: C83; E52

*We are grateful to Adam Elbourne, Bill Gavin, Jakob de Haan, Tomas Holub, Sergey Slobodyan, Tom Stanley, and Harald Uhlig for comments or providing data. The paper benefited from discussions at the Meta-Analysis of Economics Research Colloquium 2010, Conway; the Scottish Economics Society Annual Conference 2011, Perth; the International Conference on Macroeconomic Analysis and International Finance 2011, Rethymno; and the Annual Congress of the European Economic Association 2011, Oslo. We are thankful to Richard Stock for English editing. All remaining errors are ours. We acknowledge financial support from the Grant Agency of the Czech Republic (grant #P402/11/1487), and the Grant Agency of Charles University (grant #267011). The views expressed here are those of the authors and not necessarily those of the Czech National Bank. Corresponding author: Marek Rusnak, `marek.rusnak@cnb.cz`. An online appendix is available at `meta-analysis.cz/price_puzzle`.

†CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic. Correspondence address: CERGE-EI, P.O.Box 882, Politických vězňů 7, Prague, Czech Republic.
The work was supported by the grant SVV–2011–263 801.

Abstrakt

Jednou z nejdůležitejších otázek monetární ekonomie je tzv. cenová hádanka - zjištění mnoha empirických studií, že cenová hladina v krátkém období po měnové restrikci roste. V tomto článku sbíráme výsledky 70 studií používajících vektorové autoregrese ke zkoumání transmise měnové politiky a analyzujeme, co tyto výsledky ovlivňuje. Naše meta-analýza naznačuje, že výsledky jsou systematicky ovlivněny metodologií studie. Zdánlivá cenová hádanka je způsobena chybami v identifikaci vektorových autoregresí. Reakce cenové hladiny na monetární politiku v dlouhém období závisí systematicky na strukturálních charakteristikách ekonomiky.

Non-technical Summary

One of the major peculiarities of vector autoregressions, the dominant framework for the empirical analysis of monetary policy, is the counterintuitive rise in prices often reported in these models following a monetary contraction. The so-called price puzzle is encountered by about half of all empirical studies, and in many of them the puzzle is even statistically significant. In this paper we collect 70 published studies using vector autoregressions to examine the effects of monetary policy. Employing meta-regression analysis, a quantitative method of research synthesis, we investigate which aspects of methodology systematically contribute to reporting the price puzzle. The meta-regression analysis also shows how the characteristics of the countries examined influence the reported shape of the impulse responses and thus help explain the cross-country heterogeneity in monetary transmission.

We evaluate the reported graphs of impulse responses at five time horizons (representing the short, medium, and long run) and for each horizon extract the numerical value of the impulse response. In this way we collect more than 1,000 estimates, 210 on average for each horizon; the estimates summarize evidence from 31 countries and were produced by 103 researchers. We present a method of research synthesis suitable for graphical results such as impulse responses and employ modern meta-analysis methods to examine the extent of publication selection bias (the preference of authors, editors, or referees for some particular results based on significance or consistency with theory).

Our results indicate some evidence of publication selection against the price puzzle, and the selection seems to strengthen for responses with longer horizons after monetary tightening. The finding is in line with Doucouliagos & Stanley (2008), who suggest that publication selection is likely to be stronger for research areas with less theory competition. In macroeconomics, agreement exists about the effects of monetary policy on prices in the long run: prices should eventually decrease after a contraction. On the other hand, a smaller consensus arises regarding the exact effects of monetary policy in the short run because of the uncertainty caused by transmission lags. Published results often exhibit the price puzzle for the short run; on the contrary, results showing the price puzzle for the long run would be difficult to publish.

The reported impulse responses are systematically affected by study design and country-specific characteristics. Study design is important in particular for the short run: the reported short-run increase in prices after a tightening is well explained by the effects of commonly questioned aspects of methodology, such as the omission of commodity prices, the omission of potential output, or the use of recursive identification. When these aspects of methodology are filtered out, the average impulse-response function inferred from the entire literature becomes hump-shaped with no evidence of the price puzzle. Based on such "best-practice" impulse response the maximum decrease in prices following a one percentage-point increase in the interest rate is 0.33% and occurs already half a year after the tightening.

Our results suggest that heterogeneity between countries is important for the long-run response of prices to monetary policy action. Structural characteristics such as GDP growth, average inflation, and openness, as well as institutional characteristics such as financial development and central bank independence, determine the strength of transmission.

1 Introduction

How does monetary policy affect the price level? This fundamental question of monetary economics still ranks among the most controversial when it comes to empirical evidence. Although intuition and stylized macro models suggest that prices should decrease following a surprise increase in interest rates, empirical findings often challenge the theory. About 50% of modern studies using vector autoregressions (VARs) to investigate the effects of monetary policy report that after a tightening prices actually increase, at least in the short run. Beginning with Sims (1992), many different solutions to the "price puzzle" have been proposed, varying from alleged misspecifications of VARs (Bernanke *et al.*, 2005) to theoretical models that try to justify the observed rise in prices (Rabanal, 2007).

Depending on the point of view, the price puzzle casts serious doubt on either the ability of VAR models to correctly identify monetary policy shocks or on the ability of central banks to control inflation in the short run, or both. Since macroeconomists have produced a plethora of empirical research on the topic, it seems natural to ask what general effect the literature implies. The method designed to answer such questions is meta-analysis, a quantitative method of research synthesis commonly used in economics (Smith & Huang, 1995; Stanley, 2001; Disdier & Head, 2008; Card *et al.*, 2010), which can provide a unifying framework for this stream of literature. In contrast to narrative literature surveys, meta-analysis takes into account possible publication selection, the preference of authors, editors, or referees for results that are statistically significant or consistent with the theory, a bias that has become a great concern in empirical economic research (DeLong & Lang, 1992; Card & Krueger, 1995; Ashenfelter & Greenstone, 2004; Stanley, 2008).

Meta-analysis enables researchers to examine the systematic dependencies of reported results on study design and to separate the wheat from the chaff by filtering out the effects of misspecifications. Meta-analysis can create a synthetic study with ideal parameters, such as the maximum breadth of data or a consensus best-practice methodology, and, in our case, estimates the underlying effect of monetary policy corrected for potential misspecification and publication biases. Furthermore, meta-analysis makes it possible to investigate how the strength of monetary transmission depends on the characteristics of the countries examined. In this paper we attempt to collect all published studies examining monetary transmission within a VAR framework and extract point estimates of impulse responses together with the corresponding confidence bounds. We investigate the degree of publication selection, the role of model misspecification for the occurrence of the price puzzle, and the factors underlying the heterogeneity of price responses to monetary shocks across countries and over time.

Based on the mixed-effects multilevel model we illustrate how meta-analysis is able to disentangle various factors causing researchers to encounter the price puzzle. We show that when best practice is followed, the researcher is likely to find that prices decrease significantly soon after a monetary tightening. Our results thus suggest that the puzzle stems from model misspecification rather than from the real behavior of prices. In addition, the results indicate publication selection in favor of the negative responses of prices to a monetary contraction. Finally, the analysis of the determinants of transmission heterogeneity suggests that monetary policy has a stronger effect on prices in more open economies, in countries with a more independent central bank, and during economic downturns.

The remainder of the paper has the following structure. Section 2 describes how we collected the estimates from VAR models. Section 3 reviews the suggested solutions to the price puzzle. Section 4 tests for publication selection bias and for the underlying effect of monetary tightening on prices. Section 5 models the method and structural heterogeneity among impulse responses. Section 6 concludes. Appendix A provides additional robustness checks, and Appendix B lists all the studies used to construct the data set.

2 The Impulse Responses Data Set

Ever since the seminal contribution of Sims (1980), VARs have been the dominant tool for investigating monetary transmission. Researchers using VARs to examine the impact of monetary policy usually assume that the economy can be described by the following dynamic model:

$$AY_t = B(L)Y_{t-1} + \varepsilon_t, \tag{1}$$

where Y_t is a vector of endogenous variables typically containing a measure of output, prices, interest rates, and, in the case of a small open economy, the exchange rate. Matrix A describes contemporaneous relationships between endogenous variables, $B(L)$ is a matrix lag polynomial, and ε_t is a vector of structural shocks with the variance-covariance matrix $E(\varepsilon_t \varepsilon_t') = I$. The system is called the structural-form VAR. In order to estimate it, researchers rewrite the system to its reduced form:

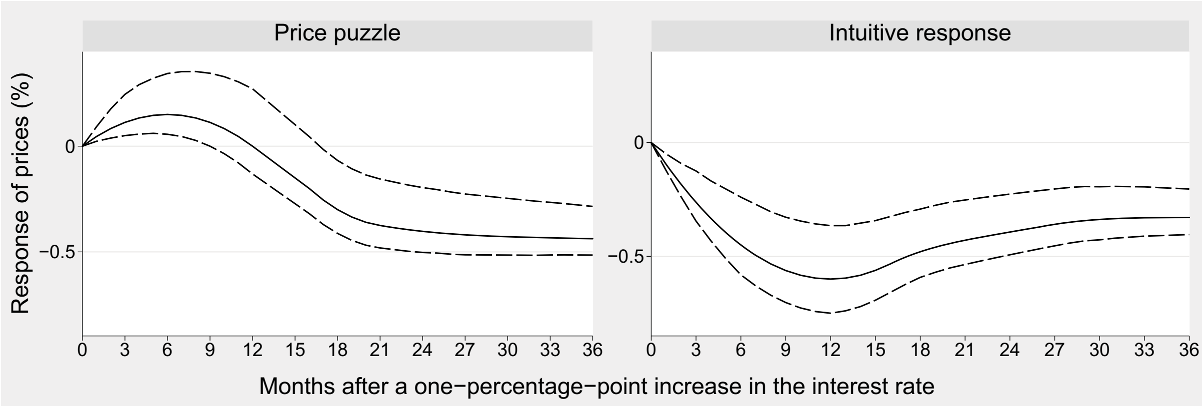
$$Y_t = C(L)Y_{t-1} + u_t, \tag{2}$$

where the elements of matrix $C(L)$ are the convolutions of the elements of matrices A and B, and u_t is a vector of reduced-form shocks with the variance-covariance matrix $E(u_t u_t') = \Sigma$; the relationship between structural shocks and reduced-form residuals is $\varepsilon_t = Au_t$. The dynamic responses of endogenous variables to structural shocks can be studied by impulse-response functions.

Figure 1 presents two stylized types of the price level's impulse responses to a monetary tightening. The left panel demonstrates the price puzzle: prices increase significantly in the short run, while in contrast, the right panel shows a response which corresponds with the mainstream prior: the price level declines soon after a tightening.

The first step of meta-analysis is to select the studies to be included. While some meta-analysts use both published and unpublished studies, others confine their sample to journal articles (for instance, Abreu *et al.*, 2005). Including working papers and mimeographs in meta-analysis does not help alleviate publication bias: if journals systematically prefer certain results, rational authors will already adopt the same preference in the earlier stages of research as they prepare for journal submission. Indeed, empirical evidence suggests no difference in the magni-

Figure 1: Stylized impulse responses



tude of publication bias between published and unpublished studies (see the meta-analysis of 65 meta-analyses by Doucouliagos & Stanley, 2008). Even if there was a difference, modern meta-regression methods not only identify but also filter out the bias. Therefore, as a preliminary and simple criterion of quality, we only consider articles published in peer-reviewed journals or in handbooks (such as the Handbook of Macroeconomics).

The following literature search strategy was employed. First, we examined two literature surveys (Stock & Watson, 2001; Egert & MacDonald, 2009) and set up a search query able to capture most of the relevant studies; we searched both the EconLit and RePEc databases. Next, we checked the references of studies published in 2010 and the citations of the most widely cited study in the VAR literature, Christiano *et al.* (1999). After going through the abstracts of all the identified studies, we selected 195 that showed any promise of containing empirical estimates of impulse responses and examined them in detail. The search was terminated on September 15, 2010.

To be able to use meta-analysis methods fully, we exclude the studies that omit to report confidence intervals around impulse responses. Unfortunately, we thus have to exclude some seminal articles such as Sims (1992) or a few recent studies that estimate time-varying-parameter VARs. To obtain a more homogeneous sample we only focus on studies that define a monetary policy shock as a shock in the interest rate. A number of studies investigate the change in the monetary base; since Bernanke & Blinder (1992) and Sims (1992), however, the majority of the literature investigates interest rate shocks because most central banks now use the interest rate as their main policy instrument. We only include studies examining the response of the price level; a minority of studies examine the responses of the inflation rate. These incorporation criteria leave 70 studies in our database. The full list of studies included in the data set can be found in Appendix B, and the list of excluded studies is presented in the online appendix at `meta-analysis.cz/price_puzzle`.

Considering the richness and heterogeneity of the empirical evidence on the effects of monetary policy, it is surprising there has been no quantitative synthesis using modern meta-regression methods.[1] One reason is that the results are typically presented in the form of graphs instead of numerical values, and the graphs contain estimates for many time horizons following the monetary policy shock. Researchers usually investigate up to 36- or 48-month horizons when using monthly data and up to 20 quarters when using quarterly data; it is unclear which horizon should be chosen to summarize the effect.

Our meta-analysis is designed in the following way. We extract responses at 3- and 6-month horizons to capture the short-run effect, at 12- and 18-month horizons to capture the medium-run effect, and at the 36-month horizon to capture the long-run effect. We enlarge the graphs of impulse responses and using pixel coordinates we measure the response and its confidence bounds. The graphs of all impulse responses as well as the extracted values are available in the online appendix. The resulting measurement error is random, similar to the rounding error in numerical outcomes, and thus inevitable in a meta-analysis.

The extracted values must be transformed into a common metric to ensure that the estimates are comparable. To standardize the estimates to represent the effect of a one-percentage-point increase in the interest rate, we divide the responses by the magnitude of the monetary policy shock used in the study.[2] In the case of factor-augmented VAR (FAVAR) studies, where the responses are usually given in standard-deviation units, we normalize the responses by the standard deviation of the particular time series.

Since the confidence intervals around the impulse response estimates are often asymmetrical (confidence intervals are usually computed by Bayesian Monte Carlo integration method; see Sims & Zha, 1999), the standard errors of the estimates cannot be obtained directly. In this case we approximate the standard error by the distance from the point estimate to the confidence bound closer to zero; that is, we take the lower confidence bound for positive responses and the upper bound for negative responses. This bound determines significance and would be associated with potential publication selection. Should we use the average of the distance to both confidence bounds, the inference would remain similar; these additional results are available in the online appendix. When the reported confidence interval is presented in standard-deviation units (for example, +/− two standard deviations), we can immediately approximate the standard error. Otherwise, we proceed as if the estimates were symmetrically distributed and assume that, for example, the 68% confidence interval represents an interval of one standard error around the mean.

Following the recent trend in meta-analysis (Disdier & Head, 2008; Doucouliagos & Stanley, 2009), we use all reported estimates from the 70 primary studies. Arbitrarily selecting the "best" estimate or using the average reported estimate would discard a great deal of useful information about the differences in methods within one study. We do not clean the data set: rather, for all regressions we evaluate the stability of coefficients employing the random sample method, which replicates the regression 1,000 times with a subset of 80% of the original data set.

[1] To our knowledge, there has been one unpublished meta-analysis on the impact of monetary policy on prices (de Grauwe & Storti, 2004) and it focused solely on heterogeneity in the reported estimates; that is, it did not filter out publication bias and misspecifications to estimate the underlying impulse response. We also use four times more point estimates of impulse responses and three times more variables that explain heterogeneity.

[2] When we were uncertain about the magnitude of shock used in the primary study, we contacted the authors.

This sensitivity analysis, presented in the online appendix, indicates that our results are robust to outliers. In addition the method of meta-analysis includes a built-in robustness check—for instance the current version of the manuscript covers 10% more studies than the first draft, yet the results remain qualitatively identical.

The number of impulse responses collected for each of the horizons is approximately 210, which in total amounts to more than 1,000 point estimates. More specifically, we collect 208 estimates for the 3-month horizon, 215 for the 6- and 12-month horizons, 217 for the 18-month horizon, and 205 for the 36-month horizon. For comparison, consider Nelson & Kennedy (2009), who review 140 economic meta-analyses and report that the median analysis only uses 92 point estimates from 33 primary studies. The oldest study in our sample was published in 1992 and the median study was published in 2006, the data set covers evidence from 31 countries, and we build upon the work of 103 researchers. The median time span of the data used by the primary studies is 1980–2002. All studies in the sample combined receive approximately 800 citations in Google Scholar per year, indicating the influence of VARs in monetary economics.

3 Collecting the Pieces of the Puzzle

To motivate the selection of explanatory variables in the multivariate meta-regression analysis (Section 5), we now briefly review the solutions to the price puzzle that have been proposed in the literature. Most of these remedies have proven to alleviate the puzzle in some cases; none of them, though, has been fully successful in solving it. Table 1 demonstrates that from the 208 estimates collected for the 3-month horizon, exactly half exhibit the price puzzle, and in 15% of the estimates the puzzle is even statistically significant at the 5% level. The table summarizes the effectiveness of the different solutions to the puzzle. Even in the case of the most effective solution, 24% of specifications still exhibit the puzzle (except for sign restrictions, which in some cases, however, represent a tautological solution).

Table 1: Effectiveness of the suggested solutions to the price puzzle

	Methodology used in the estimation						
	All	Commodity	Trend/Gap	Single	FAVAR	SVAR	Sign
Responses estimated	208	125	33	64	11	60	31
Price puzzle present	104	61	8	24	8	20	3
Price puzzle significant	32	16	1	5	3	6	0

Note: Commodity = Commodity prices are included in the VAR, Trend/Gap = time trend or output gap is included, Single = the VAR is estimated on the sample containing a single monetary policy regime, FAVAR = a factor-augmented VAR is estimated, SVAR = non-recursive identification is used, Sign = shocks are identified by imposing sign restrictions, not necessarily on prices.

3.1 Omitted Variables

Commodity Prices According to Sims (1992) the price puzzle occurs because central banks are forward-looking and react to the anticipated future movements of inflation by raising the interest rate. When a VAR system omits information about future inflation, the examined shocks become combinations of true monetary policy shocks and endogenous reactions to expected inflation. If the central bank does not fully accommodate the expected inflation, the data might show that an increase in the interest rate, mistakenly recognized as a monetary policy shock, is followed by an increase in the price level. Sims (1992) finds that including commodity prices into the VAR mitigates the price puzzle. Nevertheless, as follows from Table 1, the inclusion of commodity prices does not solve the puzzle automatically—in fact, it seems to help little.

Output Gap Giordani (2004) argues that the use of GDP in the VAR system without controlling for the potential of the economy can bias the estimates and cause the price puzzle. He claims that commodity prices alleviate the puzzle mostly because they contain useful information about the output gap, not just because they are a good predictor of future inflation. In a similar vein, Hanson (2004) examines a battery of other indicators and finds little correlation between the ability to solve the price puzzle and the ability to forecast inflation. Approximately 16% of the specifications in our sample use the output gap (or add a time trend), but some of them still encounter the puzzle.

Factor-augmented VAR To address the major shortcomings of standard small-scale VARs, Bernanke *et al.* (2005) introduce the factor-augmented VAR approach. They argue that, in practice, policymakers take into account hundreds of variables when deciding about monetary policy. Standard VAR models with typically three to six variables may therefore suffer from omitted-variable bias; the FAVAR approach, on the other hand, makes use of additional information by extracting principal components from many time series. Nonetheless, simple summary statistics in Table 1 indicate that FAVAR is not particularly effective in explaining the puzzle away.

3.2 Identification

While some researchers stress the role of omitted variables, others argue that the puzzle arises from implausible identification of monetary policy shocks. The usual recursive identification, which assumes that monetary policy affects output and prices only with a lag, is, for example, not consistent with the New-Keynesian class of theoretical models (Carlstrom *et al.*, 2009).[3]

Non-recursive Identification Kim (1999) and Kim & Roubini (2000) introduced and applied a non-recursive method for the identification of shocks. The main idea, going back to Bernanke (1986) and Blanchard & Watson (1986), says that the matrix contemporaneously linking structural shocks and reduced-form residuals is no longer lower triangular, but that it assumes a general form indicated by theory: the rows of the matrix have a structural interpre-

[3]Romer & Romer (2004) point out that the traditional indicators (money supply and interest rates) contain anticipatory movements that might contaminate estimated monetary policy shocks. By using quantitative and narrative records from the Federal Open Market Committee meetings they produce a measure of monetary policy shocks based on changes in the intended federal funds rate and the Fed's expectations on future inflation and output.

tation. The restrictions presented by Kim & Roubini (2000) are elicited from the structural stochastic equilibrium model developed by Sims & Zha (1998).

Alternatively, researchers may impose a long-run restriction in order to identify monetary policy shocks; this approach is pursued by Blanchard & Quah (1989) and Clarida & Gali (1994), who only allow technological shocks to have a permanent effect on economic activity. Recently Bjornland & Leitemo (2009) combine short-run and long-run restrictions. Although non-recursive identification is appealing, in almost 33% of the responses computed using this strategy the price puzzle still occurs.

Sign Restrictions Faust (1998), Canova & Nicolo (2002), and Uhlig (2005) present a novel identification approach that assigns a structural interpretation to orthogonal innovations by imposing sign restriction on the responses to shocks. The method is attractive since sign restrictions can be derived from the theory (e.g. from the dynamic stochastic general equilibrium (DSGE) model). The identifying assumptions are clearly stated and the shocks can be given the structural interpretation without imposing zero restrictions. The use of sign restrictions in VARs has recently been criticized by Fry & Pagan (2011): because median impulse responses that are typically presented do not come from a common model, the shocks may not be orthogonal. They instead recommend presenting the orthogonal responses that are as close to the median as possible. Nevertheless, as Table 1 documents, VARs estimated with sign restrictions rarely encounter the price puzzle.

3.3 Monetary Policy Regime

Another stream of literature suggests that the puzzle is historically limited to periods of passive monetary policy[4] or that it emerges when the data mix different monetary regimes (Elbourne & de Haan, 2006; Borys *et al.*, 2009). For example, if a researcher assumes that the central bank uses the interest rate to target inflation, although for some part of the sample monetary or exchange rate targeting was in place, monetary policy shocks in the VAR system become incorrectly identified.

Hanson (2004) shows that neither commodity prices nor other indicators are able to solve the price puzzle in the 1959–1979 period, suggesting that the puzzle is associated with that period. Similar results are reported by Castelnuovo & Surico (2010), who find the price puzzle in a pre-1979 sample even after controlling for the output gap. This finding has been reported mainly for the United States, but Benati (2008) presents similar evidence for the United Kingdom.

3.4 Cost Channel

A decrease in the price level following a tightening of monetary policy is predicted by stylized theoretical models stressing the importance of the demand channel of transmission. On the

[4]Monetary policy is considered passive when it violates the so-called Taylor principle. The Taylor principle requires the central bank to sufficiently increase the interest rate after a positive shock to inflation expectations, so that the real interest rate also increases (Clarida *et al.*, 2000).

other hand, Barth & Ramey (2002) emphasize the supply-side effects and present evidence for the so-called cost channel. Since firms depend on credit to finance production, their costs rise when the central bank increases the interest rate, and they may increase prices. In this view the price puzzle does not stem from methodological problems in VARs, but represents a genuine response to monetary tightening when the cost channel dominates the demand channel.

For the United States, Christiano *et al.* (2005) build a DSGE model incorporating the cost channel, but only find a minor role for it. In a similar vein the results of Rabanal (2007) suggest that the demand-side effects of monetary policy dominate the supply-side effects, thus leaving the cost channel relatively unimportant. Henzel *et al.* (2009) come to similar conclusions for the euro area.

4 Consequences of Publication Selection

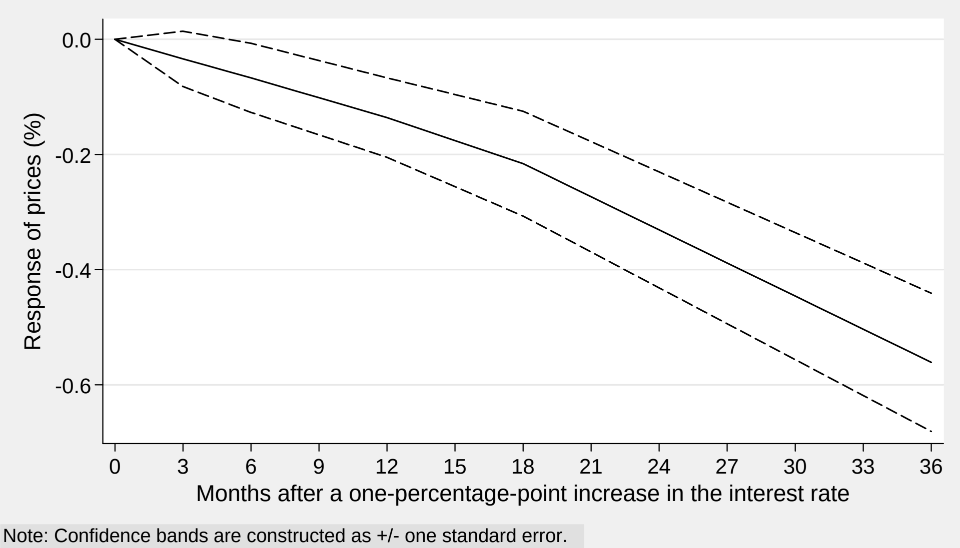
After we have collected about 1,000 estimates of the response of prices to monetary tightening, a natural question arises: what general impulse response does the literature suggest? Meta-analysis was originally developed in medicine to combine many small studies into one large one, and therefore to boost the number of degrees of freedom. Clinical trials are costly, and meta-analysis thus became the dominant method of taking stock of medical research. Estimating a VAR model may be less expensive, but the degrees of freedom in macroeconomics are limited. Hence, the original purpose of meta-analysis is useful even here since it combines information from many countries and time periods: when recomputed into quarters the primary studies in our sample taken together use 2,452 unique observations.

Taking a simple mean of all point estimates for each of the five horizons implies the impulse-response function depicted in Figure 2. This average impulse response shows a relatively intuitive short-run reaction of prices to a one-percentage-point increase in the interest rate: prices decline already in the short run, the decrease becomes significant in the medium run and reaches 0.56% after 36 months. The response shows no sign of bottoming out.

Simply averaging the collected impulse responses has two major shortcomings. First, it ignores possible publication selection. If some results are more likely to get published than others, the average becomes a biased estimator of the underlying impulse response. Second, it ignores heterogeneity in the results of the primary studies. Since different researchers use different data and methods, and the studies are of different quality, it is unrealistic to assume that all estimates are drawn from the same population. In addition, as discussed in Section 3, some VAR models are misspecified, and if misspecifications have a systematic influence on the results, it may be possible to improve upon the average response by filtering out the misspecifications. We address publication selection in this section and heterogeneity and misspecification issues in Section 5.

Stanley (2008), among others, points out that publication selection is of major concern for empirical research in economics. When there is little theory competition for what sign the underlying effect should have, estimates inconsistent with the predominant theory will be treated with suspicion or even be discarded. An illustrative example can be found in the literature on

Figure 2: Average impulse response implied by the literature: slow transmission



the effect of a common currency on trade (Rose & Stanley, 2005): it is hard to defend negative estimates of the trade effect of currency unions. The negative estimates most likely result from misspecification, and researchers may be correct in not stressing them. On the other hand, it is far more difficult to identify excessively large estimates of the same effect that also arise from misspecifications. No specific threshold exists above which the estimate would become suspicious. If researchers include the large positive estimates but omit the negative ones, the inference will be on average biased toward a stronger effect.

A similar selection, perhaps of lower intensity, may be taking place in the VAR literature on monetary transmission as well (Uhlig, 2010, p. 17, provides anecdotal evidence). Some researchers treat the price puzzle as a clear indication of a misspecification error and try to find an intuitive impulse response for interpretation. Statistical significance is also important. Significant impulse responses are more convenient for interpretation, and especially researchers in central banks may be interested in reporting a well-functioning monetary transmission with a significant reaction of prices to a change in monetary policy. The selection for significance does not distort the average estimate from the literature if the true underlying effect equals zero, but otherwise it creates a bias, again in favor of a stronger effect, since estimates with the wrong sign are less likely to be significant.

A common way to detect publication selection is an informal examination of a so-called funnel plot (Stanley & Doucouliagos, 2010). The funnel plot depicts the estimates on the horizontal axis against their precision (the inverse of the standard error) on the vertical axis. If there is no heterogeneity or misspecification, the estimates with the highest precision will be close to the true underlying effect. In the absence of publication selection the funnel is symmetrical: the reported estimates are dispersed randomly around the true effect. The asymmetry of the funnel plot suggests publication bias; for example, if estimates with a positive sign are less likely to be selected for publication, estimates on the right side of the funnel will be underrepresented.

The funnel plots for all five horizons are depicted in Figure 3.[5] The plots resemble funnels commonly reported in economic meta-analyses, which indicates that the employed approximation of standard errors is plausible. As expected, the left part of all funnels is clearly heavier, suggesting publication selection against the price puzzle and in favor of the more negative (that is, stronger) effects of monetary tightening on prices. Nevertheless, the interpretation of funnel plots is subjective, and a more formal test of publication bias is required.

Given small samples, authors wishing to obtain significant results may be tempted to try different specifications until they find estimates large enough to offset the standard errors. In contrast, with large samples even tiny estimates might be statistically significant, and authors therefore have fewer incentives to conduct a specification search. If publication selection is present, we should observe a relationship between an estimate and its standard error (or the square root of the number of observations). The following regression formalizes the idea (Card & Krueger, 1995):

$$\hat{\beta}_j = \beta + \beta_0 SE_j + e_j, \tag{3}$$

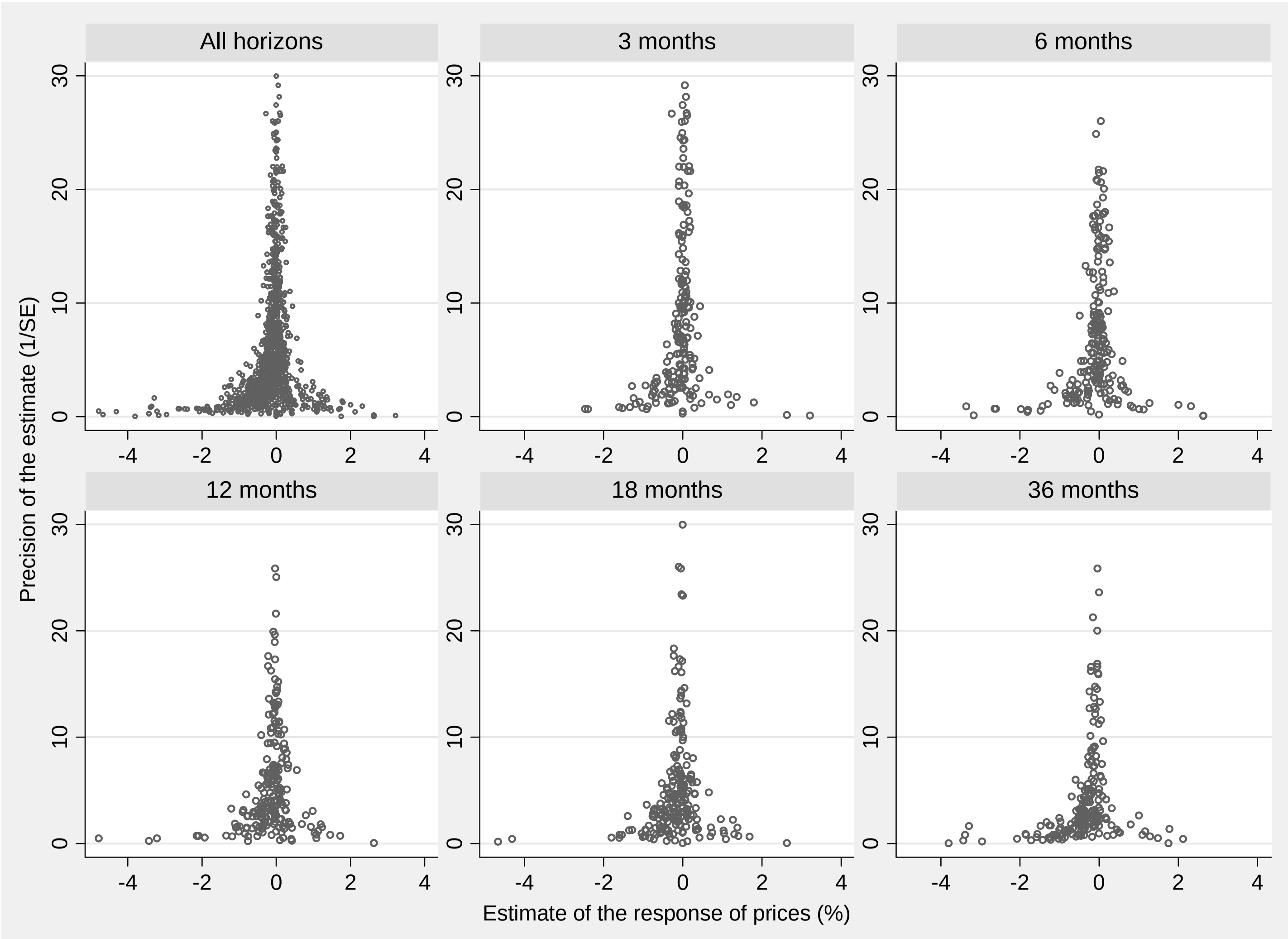
where β denotes the true underlying effect, $\hat{\beta}_j$ denotes the effect's j-th estimate, β_0 denotes the magnitude of publication bias, SE_j denotes the standard error of $\hat{\beta}_j$, and e_j denotes a disturbance term.

Specification (3) has become the cornerstone of modern meta-analysis in medicine and the social sciences, including economics. The question is whether the method is suitable for summarizing graphical results such as impulse responses. In order for this meta-analysis method to be valid, the distribution of empirical effects needs to be symmetrical *in the absence of publication bias* (usually it is assumed that the disturbance term in (3) is normally distributed). But impulse responses are nonlinear functions of the coefficients estimated in the VAR system; as discussed in Section 2, the confidence intervals around the individual estimates are often asymmetrical. If the pattern of asymmetry is not random across the individual estimates, the distribution of the impulse responses will not be symmetrical even in the absence of publication bias, and the test for publication bias will be invalid.

Systematic asymmetry of the distribution of impulse responses would manifest as a significant difference between the average distance from the point estimate of the impulse response to the lower and upper confidence bound. We select the 68% confidence bound, which for a symmetrical distribution would mean a distance of one standard error on both sides of the mean. The difference of the distances is significant at the 5% level for only one out of five horizons (the 12-month horizon), and even there the difference is small: the average lower confidence interval is 11.6% further from the mean. It is unlikely that such a small difference could explain the degree of asymmetry apparent from Figure 3. It cannot explain the asymmetry of the distribution of the collected point estimates of the impulse responses at the 12-month horizon, where the distance from the 16th to the 50th percentile is 53.1% larger than the distance from the 50th to the 84th percentile. For this reason, we employ the standard meta-analysis

[5] A few outlying estimates are trimmed from the funnels to ensure that the main pattern is clearly observable. Nevertheless, all estimates are included in the meta-regression analysis.

Figure 3: Funnel plots show publication bias



methodology—bearing in mind that the results concerning publication bias must be interpreted with some caution.

In practice, meta-analysts rarely estimate specification (3) directly since it suffers from heteroscedasticity by definition (the explanatory variable is a sample estimate of the standard deviation of the response variable). Instead, weighted least squares are used to gain efficiency, and they require that specification (3) be divided by SE_j, the measure of heteroscedasticity (Stanley, 2008):

$$\frac{\hat{\beta}_j}{SE_j} \equiv t_j = \beta_0 + \beta\left(\frac{1}{SE_j}\right) + \xi_j, \quad \xi_j|SE_j \sim N(0,\sigma^2), \tag{4}$$

where t_j denotes the approximated t-statistic of the estimate and the new disturbance term ξ_j has constant variance. Note that the intercept and the slope are now reversed: the slope measures the true effect and the intercept measures publication bias. In addition to removing heteroscedasticity, specification (4) gives more weight to more precise results, which represents a common approach in meta-analysis. Testing the significance of β_0 in this specification is analogous to testing the asymmetry of the funnel plot—it follows from rotating the funnel plot and dividing the values on the new vertical axis by SE_j. Testing the significance of β constitutes a test for the true underlying effect of monetary tightening on prices, corrected for publication selection.

Since we use all reported impulse responses we need to account for the potential dependence of estimates within one study (Disdier & Head, 2008); in such a case, (4) would be misspecified. As a remedy, researchers typically employ the mixed-effects multilevel model (Doucouliagos & Stanley, 2009):

$$t_{ij} = \beta_0 + \beta\left(\frac{1}{SE_{ij}}\right) + \alpha_j + \epsilon_{ij}, \quad \alpha_j|SE_{ij} \sim N(0,\psi), \quad \epsilon_{ij}|SE_{ij},\alpha_j \sim N(0,\theta), \tag{5}$$

where i and j denote estimate and study subscripts, respectively. The overall error term now consists of study-level random effects and estimate-level disturbances ($\xi_{ij} = \alpha_j + \epsilon_{ij}$), and its variance is additive since both components are assumed to be independent: $\text{Var}(\xi_{ij}) = \psi + \theta$, where ψ denotes between-study variance and θ within-study variance. If ψ approaches zero the benefit of using the mixed-effect estimator instead of ordinary least squares (OLS) dwindles. To put the magnitude of these variance terms into perspective the within-study correlation is useful: $\rho \equiv \text{Cor}(\xi_{ij}, \xi_{i'j}) = \psi/(\psi+\theta)$, which expresses the degree of dependence of estimates reported in the same study, or equivalently, the degree of between-study heterogeneity.

The mixed-effects multilevel model is analogous to the random-effects model commonly used in panel-data econometrics. We follow the terminology from multilevel data modeling, which calls the model "mixed effects" since it contains a fixed (β) as well as a random (α_j) part. For the purposes of meta-analysis the multilevel framework is more suitable because it takes into account the unbalancedness of the data (the restricted maximum likelihood estimator is used instead of generalized least squares), allows for nesting multiple random effects (study-, author-, or country-level), and is thus more flexible (Nelson & Kennedy, 2009).

Table 2: Test of true effect and publication bias

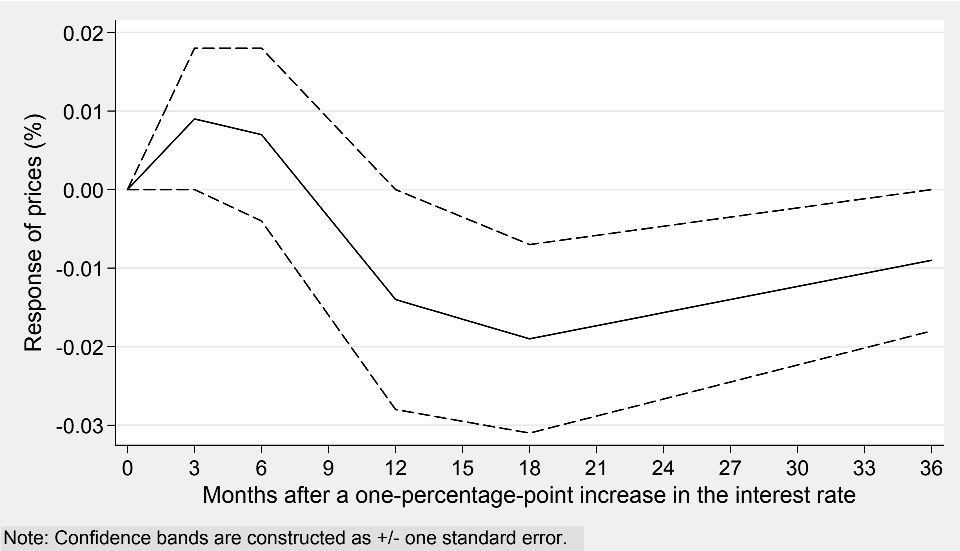
	Mixed-effects multilevel				
Horizon	3 months	6 months	12 months	18 months	36 months
Intercept (bias)	0.058	-0.088	-0.176	-0.325**	-0.806***
	(0.167)	(0.166)	(0.145)	(0.128)	(0.126)
$1/SE$ (effect)	0.009	0.007	-0.014	-0.019	-0.009
	(0.009)	(0.011)	(0.014)	(0.012)	(0.010)
Within-study correlation	0.43	0.56	0.46	0.41	0.14
Observations	208	215	215	217	205
Studies	69	70	70	70	63

Note: Standard errors in parentheses. Response variable: the approximated t-statistic of the estimate of the percentage response of prices to a one-percentage-point increase in the interest rate.
***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The outcomes of the mixed-effects estimator are presented in Table 2. OLS with standard errors clustered at the study level are reported in the Appendix: Table A1 gives even more significant results for publication bias. The within-study correlation is large, indicating that the mixed-effects estimator is more appropriate, which is confirmed by likelihood-ratio tests. We also experimented with several nested mixed-effects models (available in the online appendix), but they yield qualitatively similar outcomes. Compared with the simple average, the response of prices corrected for publication bias is more positive (that is, weaker), corroborating evidence for publication selection in favor of the stronger responses of prices to monetary policy contraction. Moreover, the magnitude of publication bias increases with the time horizon after the shock. This result is in line with Doucouliagos & Stanley (2008), who find stronger publication selection for research questions with weaker theory competition. For the short run, some disagreement occurs regarding the effects of monetary policy on prices because of the cost channel. On the other hand, a consensus emerges about the long-run effect: prices should eventually decrease after monetary policy tightening; estimates showing the opposite would be difficult to publish.

The impulse-response function corrected for publication bias is depicted in Figure 4: it exhibits the price puzzle. In the short run prices increase, but in the medium run they decrease and bottom out 18 months after the tightening. The maximum decrease in the price level, however, is negligible: only 0.02%. Compared to the average response reported in Figure 2, now the function shifts upwards—especially in the long run, because publication bias is filtered out. Figure 4 would be our best estimate of the underlying impulse response if all heterogeneity between studies was random; the estimate is unconditional on the characteristics of the countries examined and on the methodology used. In the next section we relax the assumption of random heterogeneity and explain the differences in the reported estimates.

Figure 4: Unconditional impulse response corrected for publication bias exhibits the price puzzle



5 What Explains Heterogeneity

As motivation for the empirical investigation of structural heterogeneity consider Figure 5, which depicts the differences in monetary transmission among selected countries. We use a simple random-effects meta-analysis to compute impulse-response functions. Simple meta-analysis weights each estimate by its precision and adds an estimate-specific random effect; it does not correct for publication bias. We use simple meta-analysis for estimation by countries since it requires fewer degrees of freedom than meta-regression. Figure 5 shows that the impulse responses for the United States, the United Kingdom, and Japan exhibit the price puzzle, but that monetary transmission in euro area countries seems to work intuitively and prices decline soon after a tightening. Nevertheless, a part of these differences may arise from the use of diverse methods since some countries are examined only in a few studies.

To account for heterogeneity we extend the meta-regression (5) to the following multivariate version:

$$t_{ij} = \beta_0 + \frac{\beta}{SE_{ij}} + \sum_{k=1}^{K} \frac{\gamma_k Z_{ijk}}{SE_{ij}} + \alpha_j + \epsilon_{ij}, \tag{6}$$

where Z denotes explanatory variables assumed to affect the reported estimates. The exogeneity assumptions become $\alpha_j | SE_{ij}, Z_{ijk} \sim N(0, \psi)$ and $\epsilon_{ij} | SE_{ij}, \alpha_j, Z_{ijk} \sim N(0, \theta)$.

Table 3 presents descriptions and summary statistics of all the explanatory variables we consider. In principle, they can be divided into five groups: variables capturing the fundamental characteristics of the economy (structural heterogeneity), data characteristics control for differences in the data used, specification characteristics control for differences in the basic design of the estimated models, estimation characteristics control for differences in econometric techniques, and publication characteristics control mainly for differences in quality not captured by other variables.

Figure 5: Aggregate impulse responses for selected countries suggest heterogeneity

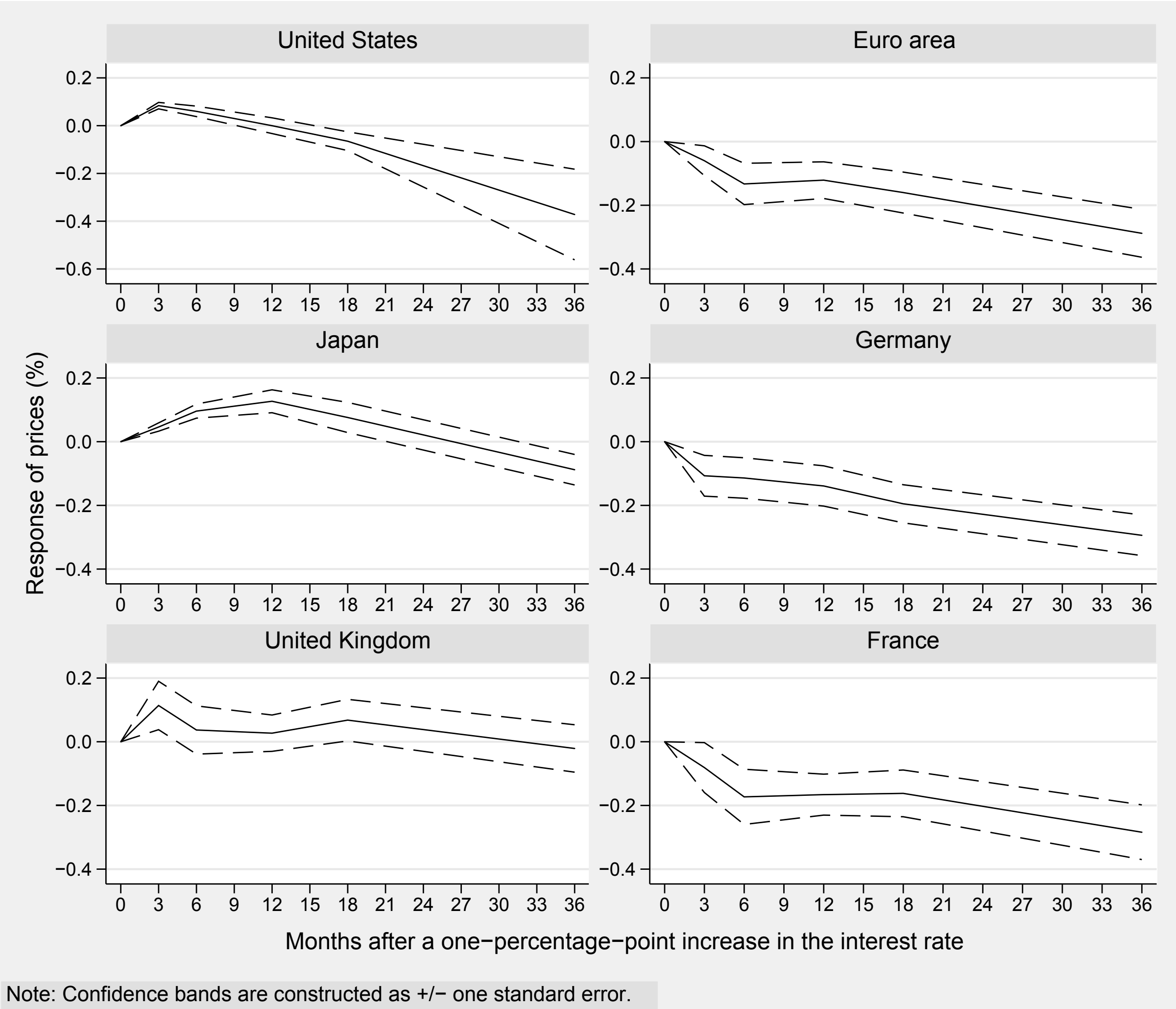


Table 3: Description and summary statistics of regression variables

Variable	Description	Mean	Std. dev.
Response (3M)	The percentage response of prices 3 months after a tightening.	-0.034	0.692
Response (6M)	The percentage response of prices 6 months after a tightening.	-0.067	0.883
Response (12M)	The percentage response of prices 12 months after a tightening.	-0.136	1.012
Response (18M)	The percentage response of prices 18 months after a tightening.	-0.216	1.327
Response (36M)	The percentage response of prices 36 months after a tightening.	-0.561	1.714
$1/SE$	The precision of the estimate of the response (all horizons).	6.805	7.821
Structural heterogeneity			
GDP per capita	The logarithm of the country's real GDP per capita.	9.881	0.414
GDP growth	The average growth rate of the country's real GDP.	2.668	1.035
Inflation	The average inflation of the country.	7.748	14.26
Inflation volatility	The standard deviation of the country's Hodrick-Prescott-filtered inflation.	6.234	33.43

Continued on the next page

Table 3: Description and summary statistics of regression variables (continued)

Variable	Description	Mean	Std. dev.
Financial development	The financial development of the country measured by (domestic credit to private sector)/GDP.	0.837	0.414
Openness	The trade openness of the country measured by (exports + imports)/GDP.	0.460	0.401
CB independence	A measure of central bank independence (Arnone *et al.*, 2009).	0.774	0.143
Data characteristics			
Monthly	=1 if monthly data are used.	0.630	0.483
Time span	The number of years of the data used in the estimation.	18.83	10.44
No. of observations	The logarithm of the number of observations used.	4.889	0.675
Average year	The average year of the data used (2000 as a base).	-8.926	7.881
Specification characteristics			
GDP deflator	=1 if the GDP deflator is used instead of the consumer price index as a measure of prices.	0.177	0.382
Single regime	=1 if the VAR is estimated over a period of a single monetary policy regime.	0.296	0.457
No. of lags	The number of lags in the model, normalized by frequency: lags/frequency	0.610	0.370
Commodity prices	=1 if a commodity price index is included.	0.607	0.489
Money	=1 if a monetary aggregate is included.	0.529	0.499
Foreign variables	=1 if at least one foreign variable is included.	0.441	0.497
Time trend	=1 if a time trend is included.	0.126	0.332
Seasonal	=1 if seasonal dummies are included.	0.146	0.354
No. of variables	The logarithm of the number of endogenous variables included in the VAR.	1.741	0.383
Industrial production	=1 if industrial production is used as a measure of economic activity.	0.430	0.495
Output gap	=1 if the output gap is used as a measure of economic activity.	0.028	0.165
Other measures	=1 if another measure of economic activity is used (employment, expenditures).	0.119	0.324
Estimation characteristics			
BVAR	=1 if a Bayesian VAR is estimated.	0.144	0.352
FAVAR	=1 if a factor-augmented VAR is estimated.	0.051	0.221
SVAR	=1 if non-recursive identification is employed.	0.295	0.456
Sign restrictions	=1 if sign restrictions are employed.	0.144	0.352
Publication characteristics			
Study citations	The logarithm of [(Google Scholar citations of the study)/(age of the study) + 1].	1.882	1.279
Impact	The recursive RePEc impact factor of the outlet.	0.888	2.274
Central banker	=1 if at least one co-author is affiliated with a central bank.	0.451	0.498
Policymaker	=1 if at least one co-author is affiliated with a Ministry of Finance, IMF, OECD, or BIS.	0.055	0.228
Native	=1 if at least one co-author is native to the investigated country.	0.446	0.497
Publication year	The year of publication (2000 as a base).	5.032	3.886

Structural heterogeneity When constructing the variables that capture structural heterogeneity, we use the average values which correspond with the sample employed in the estimation of the impulse response. For instance, in the case of inflation: When the impulse response comes from a VAR model estimated on the 1990:1–1999:12 Italian data, we use the average inflation rate in Italy for the period 1990–1999. This approach increases the variability in regressors and describes the estimates more precisely than using the same year of structural variables for all extracted impulse responses. The variable GDP per capita reflects the importance of the degree of economic development of the economy for monetary transmission. To investigate whether the strength of transmission depends on the phase of the economic cycle, we include the variable GDP growth in the meta-regression.The underlying reason is related to credit market imperfections, which could amplify the propagation of monetary policy shocks during bust periods (Bernanke & Gertler, 1989). Next, we examine the variables implied by the various channels of the transmission mechanism. We include the trade openness of the economy to capture the importance of foreign developments for domestic monetary policy as well as the exchange rate channel of monetary transmission. Furthermore, as pointed out by Bernanke & Gertler (1995) and Cecchetti (1999), differences in financial structure may explain important portions of heterogeneity in monetary transmission. We include a measure of financial development approximated by the ratio of private credit to GDP.

We add the average level and volatility of inflation, as these may influence price setting behavior as well as monetary transmission (Angeloni *et al.*, 2006). We expect that independent central banks are likely to have more credibility (Rogoff, 1985; Keefer & Stasavage, 2003; Perino, 2010). In consequence, economic subjects may respond more to monetary policy shocks. We test whether the degree of central bank independence affects the strength of monetary transmission.

Regarding the sources of the data, the trade openness, GDP growth, and GDP level per capita are obtained from Penn World Tables. The consumer price index, used to compute average inflation and inflation volatility, is obtained from the International Monetary Fund's International Financial Statistics. The ratio of domestic credit to GDP is obtained from the World Bank's World Development Indicators, and the index of central bank independence is extracted from Arnone *et al.* (2009).

Data characteristics We control for the frequency of the data used in the VAR model: 63% of specifications use monthly data, the rest rely on quarterly data. To account for possible changes in transmission not explained by the structural variables (for example, changes caused by globalization or financial innovations, see Boivin & Giannoni, 2006), we include the average year of the sample period used in the estimation. Finally, we add the total number of observations to assess whether smaller samples yield systematically different outcomes.

Specification characteristics To account for the different measures of the price level we include a dummy which equals one when the GDP deflator is used instead of the usual consumer price index (18% of specifications in primary studies). We add a dummy for the case where the data cover a period of a single monetary policy regime (30%). Next, we include the VAR's lag order normalized by the data frequency. We account for the cases where commodity prices, a money aggregate, foreign variables, a time trend, and seasonal dummies are included in the VAR. We also control for the number of endogenous variables in the model. Since the results might vary depending on the measure of economic activity, we introduce dummies for the cases where industrial production, the output gap, or another measure is used instead of GDP.

Estimation characteristics Most of the studies in our sample estimate VAR models using the standard methods (OLS or Maximum Likelihood); we control for studies using Bayesian methods to address the problem of overparameterization (14% of specifications in primary studies) and for studies using the FAVAR approach to address the problem of omitted variables (5%). As for identification strategies, most of the studies employ recursive identification; we include a dummy for non-recursive identification (30%) and a dummy for identification using sign restrictions (14%).

Publication characteristics To proxy study quality we use the recursive RePEc impact factor of the outlet (as the journal coverage of RePEc is much more comprehensive than in other databases) and the number of Google Scholar citations of the study normalized by the study's age. We add a dummy for authors affiliated with a central bank and a dummy for authors working at policy-oriented institutions such as a Ministry of Finance, the International Monetary Fund, or the Bank for International Settlements. We include a dummy for the case where at least one co-author is "native" to the examined country: such authors may be more familiar with the data at hand, which could contribute positively to the quality of the analysis; on the other hand, such authors may have a vested interest in the results. We consider authors native if they either were born in the country or obtained an academic degree there. Finally, we use the year of publication to account for possible improvements in methodology that are otherwise difficult to codify.

In the first step we estimate a general model containing all explanatory variables; the general model is not reported but is available in the online appendix. All variance inflation factors are lower than 10, indicating that the degree of multicollinearity is not too problematic. In the second step, we drop the variables which are, for each horizon, jointly insignificant at the 10% level.

For example, GDP per capita, the number of lags used, and most publication characteristics belong to the dropped variables. The insignificance of publication characteristics suggests that the quality of a given study is to a large extent captured by the methods used.

The resulting, more parsimonious, model is presented in Table 4. The specifications reported in this section are based on the mixed-effects multilevel estimator, but the inference would be similar from an OLS with standard errors clustered at the study level; these robustness checks are available in Appendix A. The similarity between the outcomes of these two estimators indicates that the exogeneity assumptions made in the mixed-effects estimation are not seriously violated; in meta-analysis it is difficult to test exogeneity formally because the extreme unbalancedness of the data (some studies report only one impulse response) does not permit the construction of a reasonable fixed-effects model. We prefer mixed effects over OLS because likelihood-ratio tests reject the hypothesis of zero within-study variance, suggesting that the OLS is misspecified.

Concerning structural heterogeneity, the results reported in Table 4 suggest that GDP growth, the openness of the economy, the level and volatility of inflation, and the degree of central bank independence systematically affect the estimated impulse response of prices to monetary tightening in the medium to long run. The importance of monetary policy shocks weakens in periods of higher GDP growth. This result is consistent with Bernanke & Gertler (1989), who argue that asymmetric information and other credit market frictions could amplify the effects of monetary policy through the so-called financial accelerator. In periods of lower GDP growth and especially during recessions, firms' dependence on external financing increases, and changes in the interest rate become more important.

Table 4: Explaining the differences in impulse responses

	Mixed-effects multilevel				
Horizon	3 months	6 months	12 months	18 months	36 months
Intercept (publication bias)	-0.112	-0.134	-0.219*	-0.208*	-0.604***
	(0.131)	(0.133)	(0.132)	(0.124)	(0.150)
$1/SE$	-0.075	-0.125	-0.287	-0.252	-0.154
	(0.117)	(0.147)	(0.181)	(0.169)	(0.202)
Structural heterogeneity					
GDP growth	-0.006	0.009	0.023**	0.023**	0.040***
	(0.008)	(0.010)	(0.011)	(0.011)	(0.012)
Inflation	0.001	-0.001	0.003	0.004	0.009***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Inflation volatility	-0.0004	0.0004	-0.0011	-0.0019	-0.0044***
	(0.0011)	(0.0014)	(0.0014)	(0.0012)	(0.0013)
Financial development	0.101***	0.080*	0.144**	0.072	-0.024
	(0.036)	(0.048)	(0.064)	(0.062)	(0.070)
Openness	-0.028	-0.048	-0.068	-0.090*	-0.283***
	(0.039)	(0.049)	(0.056)	(0.048)	(0.042)
CB independence	0.088	-0.015	-0.040	-0.167*	-0.290***
	(0.070)	(0.089)	(0.097)	(0.085)	(0.079)
Data characteristics					
No. of observations	0.011	0.027	0.049*	0.080***	0.148***
	(0.017)	(0.023)	(0.028)	(0.028)	(0.032)
Average year	0.002	-0.001	0.002	0.005*	0.013***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Specification characteristics					
GDP deflator	0.011	0.039	0.126***	0.157***	0.148
	(0.023)	(0.030)	(0.043)	(0.051)	(0.092)
Single regime	0.028	0.033	0.031	0.026	0.095**
	(0.020)	(0.025)	(0.033)	(0.035)	(0.037)

Continued on the next page

Table 4: Explaining the differences in impulse responses (continued)

	Mixed-effects multilevel				
Horizon	3 months	6 months	12 months	18 months	36 months
Commodity prices	-0.045***	-0.066***	-0.127***	-0.151***	-0.226***
	(0.016)	(0.021)	(0.030)	(0.031)	(0.033)
Foreign variables	0.011	0.032	0.062**	0.065*	0.130***
	(0.017)	(0.023)	(0.031)	(0.034)	(0.045)
No. of variables	-0.018	-0.024	-0.034	-0.056**	-0.183***
	(0.014)	(0.015)	(0.022)	(0.025)	(0.049)
Industrial production	0.030	0.060**	0.061*	0.064*	-0.011
	(0.023)	(0.027)	(0.035)	(0.038)	(0.039)
Output gap	-0.249	-0.303**	-0.219***	-0.131*	0.015
	(0.162)	(0.136)	(0.084)	(0.070)	(0.036)
Other measures	-0.072**	-0.036	-0.059	-0.041	-0.026
	(0.029)	(0.037)	(0.054)	(0.063)	(0.093)
Estimation characteristics					
BVAR	0.113***	0.085**	0.112**	0.160**	0.153
	(0.033)	(0.036)	(0.055)	(0.070)	(0.132)
FAVAR	-0.135***	-0.182***	-0.105	0.035	0.299**
	(0.036)	(0.059)	(0.082)	(0.085)	(0.122)
SVAR	-0.068***	-0.109***	-0.123***	-0.139***	-0.070***
	(0.016)	(0.018)	(0.023)	(0.022)	(0.026)
Sign restrictions	-0.294***	-0.280***	-0.334***	-0.369***	-0.271*
	(0.036)	(0.051)	(0.069)	(0.083)	(0.141)
Publication characteristics					
Central banker	0.034	0.052*	0.074**	0.076**	0.133***
	(0.022)	(0.027)	(0.033)	(0.035)	(0.038)
Policymaker	-0.057*	-0.029	0.051	0.092**	0.174***
	(0.034)	(0.043)	(0.040)	(0.038)	(0.045)
Within-study correlation	0.32	0.37	0.32	0.37	0.43
Observations	208	215	215	217	205
Studies	69	70	70	70	63

Note: Standard errors in parentheses. Response variable: the approximated t-statistic of the estimate of the percentage response of prices to a one-percentage-point increase in the interest rate. All explanatory variables are divided by the approximated standard error of the estimate at the corresponding horizon. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The expectation channel of monetary transmission can explain why the impact of monetary policy diminishes in periods of higher inflation: high inflation impedes the credibility of the central bank and restricts its ability to control the price level. Furthermore, our results indicate that a higher volatility of inflation strengthens the effect on prices in the long run. This is likely to be a consequence of monetary policy shocks having more lasting effects in more volatile environments (Mohanty & Turner, 2008). Next, monetary policy is more effective in open economies, where its impact can be amplified through the exchange rate channel. Following a contractionary monetary policy shock, the real exchange rate appreciates through the uncovered interest parity condition. As a result, imported goods become less expensive, amplifying the drop in the aggregate price level caused by monetary tightening (Dennis *et al.*, 2007). As expected, monetary policy is more powerful if the central bank enjoys more independence, which corresponds with the findings of Rogoff (1985) and Perino (2010).

In contrast, the structural variables (that is, those related to fundamentals) are not so significant for the short-run response, with the exception of the financial development indicator. Our results suggest that higher development of the financial system weakens the short-run impact of monetary policy. This finding complies with Cecchetti (1999), who reports that the effects of monetary policy are more important in countries with many small banks, less healthy banking systems, and underdeveloped capital markets.

Concerning data characteristics, the results presented in Table 4 indicate that the number of observations systematically influences the estimated long-run effect: more data make the reported response of prices weaker. In line with Boivin & Giannoni (2006), who put forward that globalization coupled with financial innovations may dampen the effects of monetary policy shocks on the economy, the reported long-run response weakens when newer data are used. Specification characteristics are found to be important as well. The GDP deflator reacts less to monetary tightening than does the consumer price index. The inclusion of commodity prices is important for all horizons and amplifies the estimated decrease in prices. When industrial production is used instead of GDP as a measure of economic activity, the reported response is typically weaker; on the other hand, the reported response strengthens when the output gap is used.

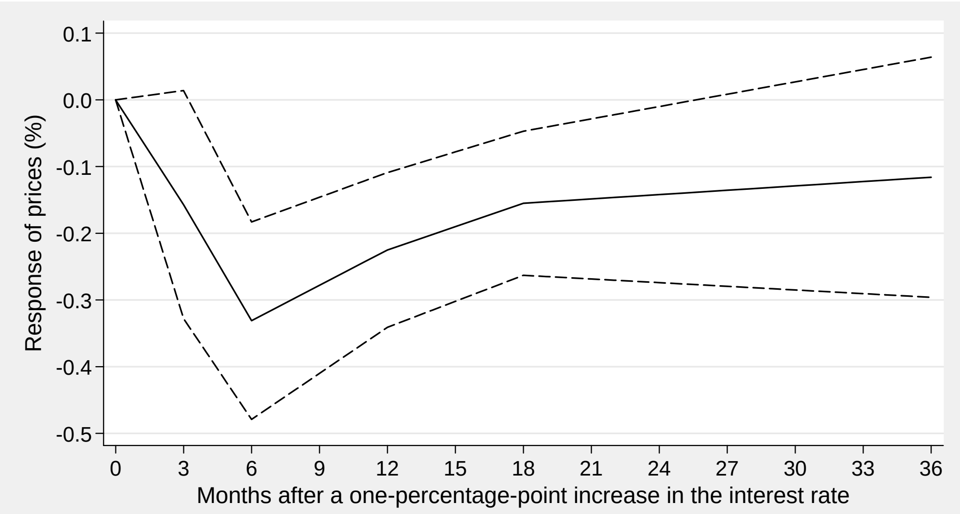
Estimation methods are important especially for the short-run response. For the 3- and 6-month horizons, Bayesian estimation produces a smaller decrease in prices compared with a simple VAR. The use of FAVAR, non-recursive identification, and sign restrictions contributes to reporting more potent monetary policy. It is worth noting that all methodological explanations of the price puzzle that were discussed in Section 3 indeed contribute to alleviating the puzzle and therefore to estimating intuitive impulse responses (with the exception of the effect of a single regime of monetary policy, which has the opposite sign, but is statistically insignificant).

Our results suggest that authors affiliated with central banks report less powerful monetary policy (that is, are more likely to report the price puzzle). This seems counterintuitive since it might be expected that central bankers have a vested interest in presenting a well-functioning monetary transmission mechanism. On the other hand, central bank employees may engage less in publication selection—they produce papers needed by their employers and often submit them to academic journals only as a by-product.

The multivariate meta-regression corroborates the evidence for publication selection reported in Section 4. The intercept, a measure of publication bias, is statistically significant for the 12-, 18-, and 36-month horizons. The estimate of the true effect in the multivariate model, however, is not simply represented by the regression coefficient for $1/SE$, but is conditional on the variables capturing heterogeneity. In order to estimate the true effect we need to choose the preferred values of the explanatory variables, thus defining some sort of best practice; in this way we create a synthetic study with ideal parameters. A suitably defined best-practice estimation can filter out misspecification bias from the literature, although the approach is subjective since different researchers may have different opinions on what constitutes best practice.

We define best practice by selecting methodology characteristics based on the discussion in Section 3: we prefer the output gap over GDP as a measure of economic activity, non-recursive identification over the simple VAR, data covering a single monetary policy regime over mixing more regimes, and the inclusion of commodity prices and foreign variables instead of omitting them. In addition, we prefer Bayesian estimation since overparameterization can be a problem even for systems of modest size (Banbura *et al.*, 2010). We insert sample maximums for the number of observations, the year of the data, and the number of endogenous variables. Country-specific variables and dummy variables for central bankers and policymakers are set to their sample means.

Figure 6: Impulse response implied by best practice: fast transmission



Note: Confidence bands are constructed as +/- one standard error.

The estimated impulse response implied by best practice is depicted in the bottom part of Figure 7: after controlling for both publication and misspecification biases, the price puzzle is not present and prices bottom out six months after a one-percentage-point increase in the interest rate. The maximum decrease in the price level reaches 0.33% and is statistically significant at the 5% level. The transmission of monetary policy shocks is quick, which contrasts with the view held at many central banks that there are long lags in the effects of monetary policy on prices (for instance, Bank of England, 1999; European Central Bank, 2010). The absence of the price puzzle is robust both individually and cumulatively to other possible definitions of best practice: selecting the FAVAR approach instead of the Bayesian approach, selecting the specification using sign restrictions instead of non-recursive identification, or selecting the sample mean of the number of endogenous variables in the VAR system instead of the sample maximum. The price puzzle does not occur even if we set the level of financial development to the sample maximum.

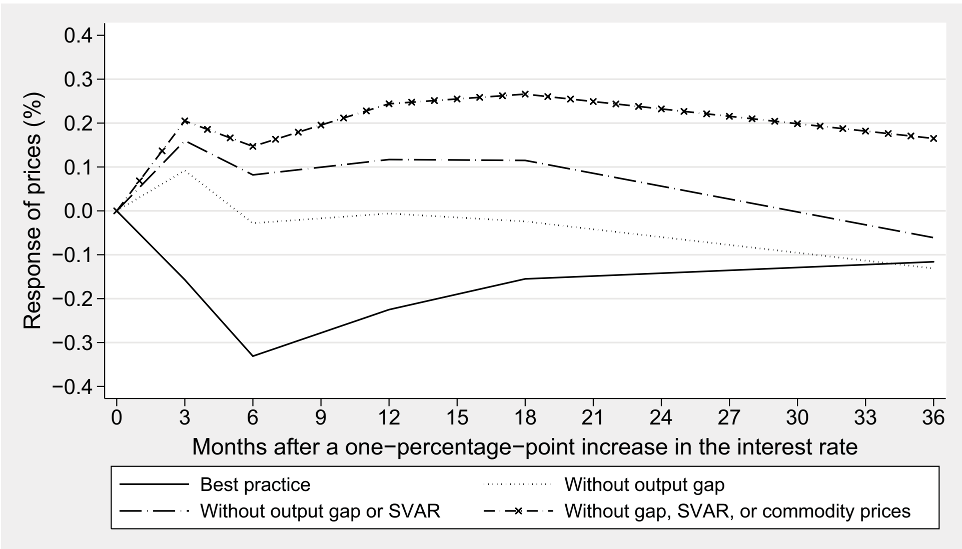
Table 5: Consequences of misspecifications

Horizon	Linear combination of regressors' values				
	3 months	6 months	12 months	18 months	36 months
Best practice	-0.157	-0.331**	-0.225*	-0.155	-0.116
Without output gap	0.092	-0.028	-0.006	-0.024	-0.131
Without gap and SVAR	0.160**	0.082	0.117	0.115	-0.061
Without gap, SVAR, and commodity prices	0.205**	0.147**	0.244**	0.266**	0.165

Note: The values represent the percentage response of prices to a one-percentage-point increase in the interest rate. Without output gap = Best practice omitting output gap. Without gap and SVAR = Best practice omitting output gap and using recursive identification. Without gap, SVAR, and commodity prices = Best practice omitting output gap, using recursive identification, and omitting commodity prices.
***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

To illustrate the consequences of misspecifications for the reported impulse responses, Table 5 and Figure 7 investigate the cases where some aspects of methodology deviate from best practice. When the model does not control for the potential output of the economy, the price puzzle occurs, but prices decline in the medium and long run. When the model combines the omission of the output gap with the use of recursive identification, the puzzle gets stronger, becomes statistically significant, and prices decline below the initial level only after 18 months. When the model additionally omits a measure of commodity prices, the price level is reported never to decline below the initial level during the 3-year horizon after monetary tightening. In sum, our analysis of the VAR studies on monetary transmission indicates that the price puzzle arises largely from misspecifications of the estimated models.

Figure 7: Deviations from best practice: Misspecifications cause the price puzzle



6 Conclusion

We examine the impact of monetary policy shocks on the price level by quantitatively reviewing the impulse-response functions from previously published VAR studies on monetary transmission. We collect impulse response estimates for 31 countries produced by 103 researchers and regress the estimates on variables capturing study design and author and country characteristics. To account for within-study dependence in the estimates, we employ mixed-effects multilevel meta-regression. Recently developed meta-analysis methods allow us to estimate the underlying effect of monetary policy implied by the entire literature net of the bias caused by publication selection and the misspecifications of some VAR models in primary studies.

Our results indicate that the estimates reporting a more powerful monetary policy (that is, a greater decrease in the price level following monetary tightening) tend to be preferentially selected for publication. The longer the horizon after a tightening, the stronger the selection. In the short run, some theory competition exists for the direction of the response since the counterintuitive increase in prices can be explained by the cost channel. In contrast, no widely accepted theory can explain why prices should stay above the initial level in the long run. This relation between publication selection and theory competition corroborates the findings of Doucouliagos & Stanley (2008), who report a similar phenomenon for many other areas of empirical research. The VAR literature on monetary transmission, on average, seems to exaggerate the long-run response of prices to monetary policy shocks.

The responses are systematically affected by study design and country-specific structural characteristics. Study design is important in particular for the short run. The reported short-run increase in prices after a tightening is well explained by the effects of commonly questioned aspects of methodology, such as the omission of commodity prices, the omission of potential output, or the use of recursive identification. When these are filtered out, the impulse-response function inferred from the entire literature becomes hump-shaped with no evidence of the price puzzle. The maximum decrease in the price level following a one-percentage-point increase in the interest rate reaches 0.33% and already occurs half a year after the tightening.

The long-run response depends on the characteristics of the examined country; on average, the decrease in prices is relatively persistent. The effect of monetary policy is weaker in countries with higher average inflation, possibly because high inflation hampers the credibility of the central bank. The effect is stronger in more open economies, in countries with a more independent central bank, and during recessions.

The robustness of our results could be further examined by adding all the unpublished studies to the data sample; this would require collecting information from hundreds of additional manuscripts, but would enable the researcher, for instance, to focus exclusively on one selected country. Researchers may conduct a meta-analysis of the effect of monetary policy on the rate of inflation (in this paper we include only studies using the price level for compatibility). In addition, the presented method of quantitative synthesis for graphical results can be applied to any other field that uses VARs as a research tool.

References

Abreu, M., H. L. F. de Groot, & R. J. G. M. Florax (2005): "A Meta-Analysis of β-Convergence: the Legendary 2%." *Journal of Economic Surveys* **19(3)**: pp. 389–420.

Angeloni, I., L. Aucremanne, M. Ehrmann, J. Galí, A. Levin, & F. Smets (2006): "New Evidence on Inflation Persistence and Price Stickiness in the Euro Area: Implications for Macro Modeling." *Journal of the European Economic Association* **4(2-3)**: pp. 562–574.

Arnone, M., B. J. Laurens, J.-F. Segalotto, & M. Sommer (2009): "Central Bank Autonomy: Lessons from Global Trends." *IMF Staff Papers* **56(2)**: pp. 263–296.

Ashenfelter, O. & M. Greenstone (2004): "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review* **94(2)**: pp. 454–460.

Banbura, M., D. Giannone, & L. Reichlin (2010): "Large Bayesian vector auto regressions." *Journal of Applied Econometrics* **25(1)**: pp. 71–92.

Bank of England (1999): "The Transmission Mechanism of Monetary Policy. A Paper by the Monetary Policy Committee." *Bank of England*, April 1999.

Barth, M. J. & V. A. Ramey (2002): "The Cost Channel of Monetary Transmission." In "NBER Macroeconomics Annual 2001, Volume 16," NBER Chapters, pp. 199–256. National Bureau of Economic Research.

Benati, L. (2008): "The 'Great Moderation' in the United Kingdom." *Journal of Money, Credit, and Banking* **40(1)**: pp. 121–147.

Bernanke, B., J. Boivin, & P. S. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-augmented Vector Autoregressive (FAVAR) Approach." *The Quarterly Journal of Economics* **120(1)**: pp. 387–422.

Bernanke, B. & M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review* **79(1)**: pp. 14–31.

Bernanke, B. S. (1986): "Alternative explanations of the money-income correlation." *Carnegie-Rochester Conference Series on Public Policy* **25(1)**: pp. 49–99.

Bernanke, B. S. & A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review* **82(4)**: pp. 901–21.

Bernanke, B. S. & M. Gertler (1995): "Inside the Black Box: The Credit Channel of Monetary Policy Transmission." *Journal of Economic Perspectives* **9(4)**: pp. 27–48.

Bjornland, H. C. & K. Leitemo (2009): "Identifying the Interdependence between US Monetary Policy and the Stock Market." *Journal of Monetary Economics* **56(2)**: pp. 275 – 282.

Blanchard, O. J. & D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* **79(4)**: pp. 655–73.

Blanchard, O. J. & M. W. Watson (1986): "Are Business Cycles All Alike?" In "The American Business Cycle: Continuity and Change," NBER Chapters, pp. 123–180. National Bureau of Economic Research.

Boivin, J. & M. P. Giannoni (2006): "Has Monetary Policy Become More Effective?" *Review of Economics and Statistics* **88(3)**: pp. 445–462.

Borys, M., R. Horvath, & M. Franta (2009): "The effects of monetary policy in the Czech Republic: an empirical study." *Empirica* **36(4)**: pp. 419–443.

Canova, F. & G. D. Nicolo (2002): "Monetary disturbances matter for business fluctuations in the G-7." *Journal of Monetary Economics* **49(6)**: pp. 1131–1159.

Card, D., J. Kluve, & A. Weber (2010): "Active Labor Market Policy Evaluations: A Meta-analysis." *The Economic Journal* **120(548)**: pp. F452–F477.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-analysis." *American Economic Review* **85(2)**: pp. 238–43.

Carlstrom, C. T., T. S. Fuerst, & M. Paustian (2009): "Monetary Policy Shocks, Choleski Identification, and DNK Models." *Journal of Monetary Economics* **56(7)**: pp. 1014 – 1021.

Castelnuovo, E. & P. Surico (2010): "Monetary Policy, Inflation Expectations and the Price Puzzle." *The Economic Journal* **120(549)**: pp. 1262–1283.

Cecchetti, S. G. (1999): "Legal structure, financial structure, and the monetary policy transmission mechanism." *Federal Reserve Bank of New York Economic Policy Review* **5(2)**: pp. 9–28.

Christiano, L. J., M. Eichenbaum, & C. L. Evans (1999): "Monetary policy shocks: What have we learned and to what end?" In J. B. Taylor & M. Woodford (editors), "Handbook of Macroeconomics," volume 1 of *Handbook of Macroeconomics*, chapter 2, pp. 65–148. Elsevier.

Christiano, L. J., M. Eichenbaum, & C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* **113(1)**: pp. 1–45.

Clarida, R., J. Galí, & M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *The Quarterly Journal of Economics* **115(1)**: pp. 147–180.

Clarida, R. & J. Gali (1994): "Sources of real exchange-rate fluctuations: How important are nominal shocks?" *Carnegie-Rochester Conference Series on Public Policy* **41(1)**: pp. 1–56.

DeLong, J. B. & K. Lang (1992): "Are all economic hypotheses false?" *Journal of Political Economy* **100(6)**: pp. 1257–72.

Dennis, R., K. Leitemo, & U. Söderström (2007): "Monetary policy in a small open economy with a preference for robustness." *Working Paper Series 2007-04*, Federal Reserve Bank of San Francisco.

Disdier, A.-C. & K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics* **90(1)**: pp. 37–48.

Doucouliagos, H. & T. D. Stanley (2008): "Theory Competition and Selectivity: Are All Economic Facts Greatly Exaggerated?" *Economics Series 06*, Deakin University.

Doucouliagos, H. & T. D. Stanley (2009): "Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis." *British Journal of Industrial Relations* **47(2)**: pp. 406–428.

Egert, B. & R. MacDonald (2009): "Monetary Transmission Mechanism in Central and Eastern Europe: Surveying the Surveyable." *Journal of Economic Surveys* **23(2)**: pp. 277–327.

Elbourne, A. & J. de Haan (2006): "Financial structure and monetary policy transmission in transition countries." *Journal of Comparative Economics* **34(1)**: pp. 1–23.

European Central Bank (2010): "Monthly Bulletin." *European Central Bank*, May 2010.

Faust, J. (1998): "The robustness of identified VAR conclusions about money." *Carnegie-Rochester Conference Series on Public Policy* **49(1)**: pp. 207–244.

Fry, R. & A. Pagan (2011): "Sign Restrictions in Structural Vector Autoregressions: A Critical Review." *Journal of Economic Literature* **(forthcoming)**.

Giordani, P. (2004): "An alternative explanation of the price puzzle." *Journal of Monetary Economics* **51(6)**: pp. 1271–1296.

de Grauwe, P. & C. C. Storti (2004): "The Effects of Monetary Policy: A Meta-Analysis." *CESifo Working Paper Series 1224*, CESifo Group Munich.

Hanson, M. S. (2004): "The price puzzle reconsidered." *Journal of Monetary Economics* **51(7)**: pp. 1385–1413.

Henzel, S., O. Hülsewig, E. Mayer, & T. Wollmershäuser (2009): "The price puzzle revisited: Can the cost channel explain a rise in inflation after a monetary policy shock?" *Journal of Macroeconomics* **31(2)**: pp. 268–289.

Keefer, P. & D. Stasavage (2003): "The limits to delegation: Veto players, central bank independence, and the credibility of monetary policy." *American Political Science Review* **97(3)**: p. 407–423.

Kim, S. (1999): "Do Monetary Policy Shocks Matter in the G-7 Countries? Using Common Identifying Assumptions about Monetary Policy across Countries." *Journal of International Economics* **48(2)**: pp. 387–412.

Kim, S. & N. Roubini (2000): "Exchange Rate Anomalies in the Industrial Countries: A Solution with a Structural VAR Approach." *Journal of Monetary Economics* **45(3)**: pp. 561–586.

Mohanty, M. S. & P. Turner (2008): "Monetary policy transmission in emerging market economies: what is new?" In "Transmission mechanisms for monetary policy in emerging market economies," volume 35 of *BIS Papers chapters*, pp. 1–59. Bank for International Settlements.

Nelson, J. & P. Kennedy (2009): "The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment." *Environmental & Resource Economics* **42(3)**: pp. 345–377.

Perino, G. (2010): "How delegation improves commitment." *Economics Letters* **106(2)**: pp. 137–139.

Rabanal, P. (2007): "Does inflation increase after a monetary policy tightening? Answers based on an estimated DSGE model." *Journal of Economic Dynamics and Control* **31(3)**: pp. 906–937.

Rogoff, K. (1985): "The Optimal Degree of Commitment to an Intermediate Monetary Target." *The Quarterly Journal of Economics* **100(4)**: pp. 1169–89.

Romer, C. D. & D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* **94(4)**: pp. 1055–1084.

Rose, A. K. & T. D. Stanley (2005): "A Meta-Analysis of the Effect of Common Currencies on International Trade." *Journal of Economic Surveys* **19(3)**: pp. 347–365.

Sims, C. A. (1980): "Macroeconomics and Reality." *Econometrica* **48(1)**: pp. 1–48.

Sims, C. A. (1992): "Interpreting the macroeconomic time series facts: The effects of monetary policy." *European Economic Review* **36(5)**: pp. 975–1000.

Sims, C. A. & T. Zha (1998): "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review* **39(4)**: pp. 949–68.

Sims, C. A. & T. Zha (1999): "Error Bands for Impulse Responses." *Econometrica* **67(5)**: pp. 1113–1156.

Smith, V. K. & J.-C. Huang (1995): "Can Markets Value Air Quality? A Meta-analysis of Hedonic Property Value Models." *Journal of Political Economy* **103(1)**: pp. 209–27.

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

Stanley, T. D. (2008): "Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

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

Stock, J. H. & M. W. Watson (2001): "Vector Autoregressions." *Journal of Economic Perspectives* **15(4)**: pp. 101–115.

Uhlig, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics* **52(2)**: pp. 381–419.

Uhlig, H. (2010): "Economics and Reality." *NBER Working Papers 16416*, National Bureau of Economic Research, Inc.

Appendix A: Robustness Checks

Table A1: Test of publication bias and true effect, OLS

	OLS with clustered standard errors				
Horizon	3 months	6 months	12 months	18 months	36 months
Intercept (bias)	-0.277	-0.407**	-0.341**	-0.393***	-0.784***
	(0.176)	(0.186)	(0.156)	(0.147)	(0.122)
$1/SE$ (effect)	0.032**	0.033	-0.007	-0.025*	-0.018**
	(0.014)	(0.021)	(0.016)	(0.014)	(0.008)
R^2	0.05	0.03	0.00	0.02	0.01
Observations	208	215	215	217	205
Studies	69	70	70	70	63

Note: Standard errors, clustered at the study level, in parentheses. Response variable: the approximated t-statistic of the estimate of the percentage response of prices to a one-percentage-point increase in the interest rate.
***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A2: Explaining the differences in impulse responses, OLS

	OLS with clustered standard errors				
Horizon	3 months	6 months	12 months	18 months	36 months
Intercept (bias)	-0.131	-0.127	-0.240*	-0.221*	-0.538***
	(0.151)	(0.133)	(0.128)	(0.120)	(0.130)
$1/SE$	-0.058	-0.106	-0.237	-0.168	-0.028
	(0.068)	(0.115)	(0.178)	(0.174)	(0.212)
Structural heterogeneity GDP growth	-0.008	0.010	0.024*	0.027*	0.037
	(0.008)	(0.010)	(0.013)	(0.014)	(0.024)
Inflation	-0.000	-0.003	0.003	0.005**	0.008***
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Inflation volatility	-0.000	0.001	-0.001	-0.002**	-0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Financial development	0.093***	0.079	0.174**	0.110	-0.054
	(0.030)	(0.054)	(0.076)	(0.073)	(0.067)
Openness	-0.026	-0.052	-0.089*	-0.130***	-0.258**
	(0.031)	(0.048)	(0.048)	(0.048)	(0.117)
CB independence	0.038	-0.141	-0.135	-0.258**	-0.338***
	(0.068)	(0.106)	(0.133)	(0.123)	(0.061)
Data characteristics					
No. of observations	0.020*	0.043**	0.053**	0.074***	0.127***
	(0.011)	(0.019)	(0.023)	(0.025)	(0.047)
Average year	0.001	-0.001	0.004	0.006**	0.012***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Specification characteristics					
GDP deflator	-0.004	0.023	0.119***	0.141***	0.119*
	(0.013)	(0.021)	(0.039)	(0.046)	(0.060)
Single regime	0.038**	0.034	0.024	0.021	0.109**

Continued on the next page

Table A2: Explaining the differences in impulse responses, OLS (continued)

	OLS with clustered standard errors				
Horizon	3 months	6 months	12 months	18 months	36 months
	(0.015)	(0.022)	(0.028)	(0.032)	(0.053)
Commodity prices	-0.047***	-0.070***	-0.139***	-0.158***	-0.212***
	(0.008)	(0.018)	(0.023)	(0.027)	(0.059)
Foreign variables	0.009	0.041***	0.068**	0.071*	0.082
	(0.015)	(0.013)	(0.030)	(0.038)	(0.055)
No. of variables	-0.022*	-0.024**	-0.039**	-0.059***	-0.153***
	(0.012)	(0.011)	(0.016)	(0.022)	(0.038)
Industrial production	0.024	0.062***	0.065**	0.069*	-0.026
	(0.016)	(0.018)	(0.032)	(0.040)	(0.041)
Output gap	-0.259***	-0.330***	-0.235***	-0.140***	0.012
	(0.090)	(0.102)	(0.060)	(0.039)	(0.031)
Other measure	-0.094***	-0.066**	-0.065	-0.044	0.018
	(0.022)	(0.030)	(0.058)	(0.077)	(0.079)
Estimation characteristics					
BVAR	0.136***	0.099***	0.105*	0.146	0.131
	(0.026)	(0.027)	(0.055)	(0.089)	(0.164)
FAVAR	-0.084***	-0.118***	-0.073	0.029	0.270***
	(0.025)	(0.037)	(0.054)	(0.063)	(0.068)
SVAR	-0.089***	-0.142***	-0.139***	-0.147***	-0.050
	(0.018)	(0.026)	(0.031)	(0.030)	(0.033)
Sign restrictions	-0.300***	-0.299***	-0.347***	-0.396***	-0.250
	(0.031)	(0.042)	(0.061)	(0.096)	(0.172)
Publication characteristics					
Central banker	0.024*	0.058**	0.089**	0.102**	0.125***
	(0.014)	(0.023)	(0.035)	(0.040)	(0.036)
Policymaker	-0.051**	-0.006	0.070**	0.089***	0.119***
	(0.023)	(0.022)	(0.033)	(0.032)	(0.033)
R^2	0.59	0.58	0.48	0.47	0.45
Observations	208	215	215	217	205
Studies	69	70	70	70	63

Note: Standard errors, clustered at the study level, in parentheses. Response variable: the approximated t-statistic of the estimate of the percentage response of prices to a one-percentage-point increase in the interest rate. All explanatory variables are divided by the approximated standard error of the estimate at the corresponding horizon.

***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Appendix B: Studies Used in the Meta-Analysis

Andries, M. A. (2008): "Monetary Policy Transmission Mechanism in Romania—A VAR Approach." *Theoretical and Applied Economics* **11(528)(11)**: pp. 250–260.

Anzuini, A. & A. Levy (2007): "Monetary Policy Shocks in the New EU Members: A VAR Approach." *Applied Economics* **39(7-9)**: pp. 1147–1161.

Arin, K. P. & S. P. Jolly (2005): "Trans-Tasman Transmission of Monetary Shocks: Evidence from a VAR Approach." *Atlantic Economic Journal* **33(3)**: pp. 267–283.

Bagliano, F. C. & C. A. Favero (1998): "Measuring Monetary Policy with VAR Models: An Evaluation." *European Economic Review* **42(6)**: pp. 1069–1112.

Bagliano, F. C. & C. A. Favero (1999): "Information from Financial Markets and VAR Measures of Monetary Policy." *European Economic Review* **43(4-6)**: pp. 825 – 837.

Banbura, M., D. Giannone, & L. Reichlin (2010): "Large Bayesian vector auto regressions." *Journal of Applied Econometrics* **25(1)**: pp. 71–92.

Belviso, F. & F. Milani (2006): "Structural Factor-Augmented VARs (SFAVARs) and the Effects of Monetary Policy." *B.E. Journal of Macroeconomics: Topics in Macroeconomics* **6(3)**: pp. 1–44.

Bernanke, B., J. Boivin, & P. S. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-augmented Vector Autoregressive (FAVAR) Approach." *The Quarterly Journal of Economics* **120(1)**: pp. 387–422.

Bernanke, B. S., M. Gertler, & M. Watson (1997): "Systematic Monetary Policy and the Effects of Oil Price Shocks." *Brookings Papers on Economic Activity* **1(1)**: pp. 91–142.

Berument, H. (2007): "Measuring monetary policy for a small open economy: Turkey." *Journal of Macroeconomics* **29(2)**: pp. 411–430.

Boivin, J. & M. P. Giannoni (2007): "Global Forces and Monetary Policy Effectiveness." In "International Dimensions of Monetary Policy," NBER Chapters, pp. 429–478. National Bureau of Economic Research, Inc.

Borys, M., R. Horvath, & M. Franta (2009): "The effects of monetary policy in the Czech Republic: an empirical study." *Empirica* **36(4)**: pp. 419–443.

Bredin, D. & G. O'Reilly (2004): "An Analysis of the Transmission Mechanism of Monetary Policy in Ireland." *Applied Economics* **36(1)**: pp. 49–58.

Brissimis, S. N. & N. S. Magginas (2006): "Forward-Looking Information in VAR Models and the Price Puzzle." *Journal of Monetary Economics* **53(6)**: pp. 1225–1234.

Brunner, A. D. (2000): "On the Derivation of Monetary Policy Shocks: Should We Throw the VAR out with the Bath Water?" *Journal of Money, Credit, and Banking* **32(2)**: pp. 254–279.

Buckle, R. A., K. Kim, H. Kirkham, N. McLellan, & J. Sharma (2007): "A Structural VAR Business Cycle Model for a Volatile Small Open Economy." *Economic Modelling* **24(6)**: pp. 990–1017.

Cespedes, B., E. Lima, & A. Maka (2008): "Monetary Policy, Inflation and the Level of Economic Activity in Brazil after the Real Plan: Stylized Facts from SVAR Models." *Revista Brasileira de Economia* **62(2)**: pp. 123–160.

Christiano, L. J., M. Eichenbaum, & C. Evans (1996): "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds." *The Review of Economics and Statistics* **78(1)**: pp. 16–34.

Christiano, L. J., M. Eichenbaum, & C. L. Evans (1999): "Monetary policy shocks: What have we learned and to what end?" In J. B. Taylor & M. Woodford (editors), "Handbook of Macroeconomics," volume 1 of *Handbook of Macroeconomics*, chapter 2, pp. 65–148. Elsevier.

Croushore, D. & C. L. Evans (2006): "Data Revisions and the Identification of Monetary Policy Shocks."

Journal of Monetary Economics **53(6)**: pp. 1135–1160.

Cushman, D. O. & T. Zha (1997): "Identifying monetary policy in a small open economy under flexible exchange rates." *Journal of Monetary Economics* **39(3)**: pp. 433–448.

De Arcangelis, G. & G. Di Giorgio (2001): "Measuring Monetary Policy Shocks in a Small Open Economy." *Economic Notes* **30(1)**: pp. 81–107.

Dedola, L. & F. Lippi (2005): "The monetary transmission mechanism: Evidence from the industries of five OECD countries." *European Economic Review* **49(6)**: pp. 1543–1569.

EFN (2004): "Monetary transmission in acceding countries." In "European Forecasting Network Annex 53," p. 97–142. European University Institute.

Eichenbaum, M. (1992): "Comment on 'Interpreting the macroeconomic time series facts: The effects of monetary policy'." *European Economic Review* **36(5)**: pp. 1001–1011.

Eickmeier, S., B. Hofmann, & A. Worms (2009): "Macroeconomic Fluctuations and Bank Lending: Evidence for Germany and the Euro Area." *German Economic Review* **10**: pp. 193–223.

Elbourne, A. (2008): "The UK Housing Market and the Monetary Policy Transmission Mechanism: An SVAR Approach." *Journal of Housing Economics* **17(1)**: pp. 65–87.

Elbourne, A. & J. de Haan (2006): "Financial structure and monetary policy transmission in transition countries." *Journal of Comparative Economics* **34(1)**: pp. 1–23.

Elbourne, A. & J. de Haan (2009): "Modeling Monetary Policy Transmission in Acceding Countries: Vector Autoregression versus Structural Vector Autoregression." *Emerging Markets Finance and Trade* **45(2)**: pp. 4–20.

Forni, M. & L. Gambetti (2010): "The dynamic effects of monetary policy: A structural factor model approach." *Journal of Monetary Economics* **57(2)**: pp. 203–216.

Fujiwara, I. (2004): "Output Composition of the Monetary Policy Transmission Mechanism in Japan." *Topics in Macroeconomics* **4(1)**: pp. 1–21.

Gan, W. B. & L. Y. Soon (2003): "Characterizing the Monetary Transmission Mechanism in a Small Open Economy: The Case of Malaysia." *Singapore Economic Review* **48(2)**: pp. 113–134.

Gavin, W. T. & D. M. Kemme (2009): "Using extraneous information to analyze monetary policy in transition economies." *Journal of International Money and Finance* **28(5)**: pp. 868–879.

Hanson, M. S. (2004): "The price puzzle reconsidered." *Journal of Monetary Economics* **51(7)**: pp. 1385–1413.

Horvath, R. & M. Rusnak (2009): "How Important Are Foreign Shocks in a Small Open Economy? The Case of Slovakia." *Global Economy Journal* **9(1)**.

Hulsewig, O., E. Mayer, & T. Wollmershauser (2006): "Bank Loan Supply and Monetary Policy Transmission in Germany: An Assessment Based on Matching Impulse Responses." *Journal of Banking and Finance* **30(10)**: pp. 2893–2910.

Jang, K. & M. Ogaki (2004): "The Effects of Monetary Policy Shocks on Exchange Rates: A Structural Vector Error Correction Model Approach." *Journal of the Japanese and International Economies* **18(1)**: pp. 99–114.

Jarocinski, M. (2009): "Responses to monetary policy shocks in the east and the west of Europe: a comparison." *Journal of Applied Econometrics* **25(5)**: pp. 833–868.

Jarocinski, M. & F. R. Smets (2008): "House Prices and the Stance of Monetary Policy." *Federal Reserve Bank of St. Louis Review* **90(4)**: pp. 339–365.

Kim, S. (2001): "International Transmission of U.S. Monetary Policy Shocks: Evidence from VARs." *Journal of Monetary Economics* **48(2)**: pp. 339–372.

Kim, S. (2002): "Exchange rate stabilization in the ERM: identifying European monetary policy reactions." *Journal of International Money and Finance* **21(3)**: pp. 413–434.

KRUSEC, D. (2010): "The price puzzle in the monetary transmission VARs with long-run restrictions." *Economics Letters* **106(3)**: pp. 147–150.

KUBO, A. (2008): "Macroeconomic Impact of Monetary Policy Shocks: Evidence from Recent Experience in Thailand." *Journal of Asian Economics* **19(1)**: pp. 83–91.

LAGANA, G. & A. MOUNTFORD (2005): "Measuring Monetary Policy in the UK: A Factor-Augmented Vector Autoregression Model Approach." *Manchester School* **73(s1)**: pp. 77–98.

LANGE, R. H. (2010): "Regime-switching monetary policy in Canada." *Journal of Macroeconomics* **32(3)**: pp. 782–796.

LEEPER, E. M., C. A. SIMS, & T. ZHA (1996): "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity* **27(2)**: pp. 1–78.

LI, Y. D., T. B. ISCAN, & K. XU (2010): "The impact of monetary policy shocks on stock prices: Evidence from Canada and the United States." *Journal of International Money and Finance* **29(5)**: pp. 876–896.

MCMILLIN, W. D. (2001): "The Effects of Monetary Policy Shocks: Comparing Contemporaneous versus Long-Run Identifying Restrictions." *Southern Economic Journal* **67(3)**: pp. 618–636.

MERTENS, K. (2008): "Deposit rate ceilings and monetary transmission in the US." *Journal of Monetary Economics* **55(7)**: pp. 1290–1302.

MINELLA, A. (2003): "Monetary Policy and Inflation in Brazil (1975-2000): A VAR Estimation." *Revista Brasileira de Economia* **57(3)**: pp. 605–635.

MOJON, B. (2008): "When Did Unsystematic Monetary Policy Have an Effect on Inflation?" *European Economic Review* **52(3)**: pp. 487–497.

MOJON, B. & G. PEERSMAN (2003): "A VAR description of the effects of monetary policy in the individual countries of the Euro area." In A. K. I ANGELONI & B. MOJON (editors), "Monetary Policy Transmission in the Euro Area," chapter 1, pp. 56–74. Cambridge University Press.

MOUNTFORD, A. (2005): "Leaning into the Wind: A Structural VAR Investigation of UK Monetary Policy." *Oxford Bulletin of Economics and Statistics* **67(5)**: pp. 597–621.

NAKASHIMA, K. (2006): "The Bank of Japan's Operating Procedures and the Identification of Monetary Policy Shocks: A Reexamination Using the Bernanke-Mihov Approach." *Journal of the Japanese and International Economies* **20(3)**: pp. 406–433.

NORMANDIN, M. & L. PHANEUF (2004): "Monetary Policy Shocks: Testing Identification Conditions under Time-Varying Conditional Volatility." *Journal of Monetary Economics* **51(6)**: pp. 1217–1243.

OROS, C. & C. ROMOCEA-TURCU (2009): "The Monetary Transmission Mechanisms In The CEECs: A Structural VAR Approach." *Applied Econometrics and International Development* **9(2)**: pp. 73–86.

PEERSMAN, G. (2004): "The Transmission of Monetary Policy in the Euro Area: Are the Effects Different across Countries?" *Oxford Bulletin of Economics and Statistics* **66(3)**: pp. 285–308.

PEERSMAN, G. (2005): "What Caused the Early Millennium Slowdown? Evidence Based on Vector Autoregressions." *Journal of Applied Econometrics* **20(2)**: pp. 185–207.

PEERSMAN, G. & F. SMETS (2003): "The Monetary Transmission Mechanism in the Euro Area: More Evidence from VAR Analysis." In A. K. I.ANGELONI & B.MOJON (editors), "Monetary Policy Transmission on the Euro Area," chapter 2, pp. 36–55. Cambridge University Press.

PEERSMAN, G. & R. STRAUB (2009): "Technology Shocks and Robust Sign Restrictions in a Euro Area SVAR." *International Economic Review* **50(3)**: pp. 727–750.

POBRE, M. L. (2003): "Sources of Shocks and Monetary Policy in the 1997 Asian Crisis: The Case of Korea and Thailand." *Osaka Economic Papers* **53(3)**: pp. 362–373.

RAFIQ, M. S. & S. K. MALLICK (2008): "The Effect of Monetary Policy on Output in EMU3: A Sign Restriction

Approach." *Journal of Macroeconomics* **30(4)**: pp. 1756–1791.

Romer, C. D. & D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* **94(4)**: pp. 1055–1084.

Shioji, E. (2000): "Identifying Monetary Policy Shocks in Japan." *Journal of the Japanese and International Economies* **14(1)**: pp. 22–42.

Sims, C. A. & T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* **96(1)**: pp. 54–81.

Smets, F. (1997): "Measuring Monetary Policy Shocks in France, Germany and Italy: The Role of the Exchange Rate." *Swiss Journal of Economics and Statistics* **133(3)**: pp. 597–616.

Sousa, J. & A. Zaghini (2008): "Monetary Policy Shocks in the Euro Area and Global Liquidity Spillovers." *International Journal of Finance and Economics* **13(3)**: pp. 205–218.

Vargas-Silva, C. (2008): "Monetary Policy and the US Housing Market: A VAR Analysis Imposing Sign Restrictions." *Journal of Macroeconomics* **30(3)**: pp. 977–990.

Voss, G. & L. Willard (2009): "Monetary policy and the exchange rate: Evidence from a two-country model." *Journal of Macroeconomics* **31(4)**: pp. 708–720.

Wu, T. (2003): "Stylized Facts on Nominal Term Structure and Business Cycles: An Empirical VAR Study." *Applied Economics* **35(8)**: pp. 901–906.

Working Paper Series
ISSN 1211-3298
Registration No. (Ministry of Culture): E 19443

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the European Structural Fund (within the Operational Programme Prague Adaptability), the budget of the City of Prague, the Czech Republic's state budget and the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2011);
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2011);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: http://www.cerge-ei.cz

The paper is available online at http://www.cerge-ei.cz/publications/working_papers/.

ISBN 978-80-7343-249-2 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-241-5 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz