Understanding Instruments in Macroeconomics - A Study of High-Frequency Identification

Pooyan Amir-Ahmadi   Christian Matthes
Mu-Chun Wang
*

July 8, 2022

Abstract

The effects of monetary policy shocks are regularly estimated using high-frequency surprises in asset prices around central bank meetings as an instrument. These studies assume a constant relationship between the instrument and the monetary policy shock. By allowing for time variation in this relationship, we show that only a few distinct periods are informative about monetary policy shocks. We thus build a narrative for instrument-based identification and sharpen results: for the instrument in Gertler & Karadi (2015), the effect of monetary policy shocks on the (log) price level is almost 50 percent larger than the standard specification would suggest. This result can be obtained using only 10 percent of all available observations of the instrument.

Keywords: High-Frequency Identification, Instruments, Monetary Policy
JEL Codes: C32, E31, E44

*Affiliations: University of Illinois and Amazon (Amir-Ahmadi), Indiana University (Matthes), and Deutsche Bundesbank (Wang). Previous versions of this paper circulated under the title "What Information Do Proxy VARs Use?". We thank Jonas Arias, Christiane Baumeister, Marco Del Negro, Raffaella Giacomini, Ed Herbst, Toru Kitagawa, Mikkel Plagborg-Møller, Giorgio Primiceri, Juan Rubio-Ramírez, Frank Schorfheide, and Jonathan Wright for very helpful comments. The paper has benefited from comments at the SBIES, the Bundesbank, the Applied Time Series Workshop at the St. Louis Fed, and the Barcelona Summer Forum. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Deutsche Bundesbank or the Eurosystem. This paper and its contents are not related to Amazon and do not reflect the position of the company and its subsidiaries.

1 Introduction

Identifying impulse responses via external instruments has become commonplace in empirical macroeconomics over the last decade (Stock & Watson 2012, Mertens & Ravn 2013, Gertler & Karadi 2015). These external instruments are interpreted as imperfect measurements of unobserved structural shocks. A key assumption of the underlying studies that use this approach is that there is a fixed, time-invariant relationship between the instrument and the shock of interest. However, in this paper we present evidence that for a common application of external instruments—the study of monetary policy shocks using high-frequency variation in asset prices around central bank announcements—there is actually substantial time variation in this relationship. To see this, Figure 1 plots the surprises in the three-month-ahead Fed Funds futures (FF4) in a 30-minute window around meetings of the Federal Open Market Committee (FOMC), a instrument popularized by Gertler & Karadi (2015) that we use as well. The figure shows there are periods where the dynamics of this instrument are substantially different from the rest of the sample, mainly the early 1990s, 2001, and during the Great Recession.

Building on this finding, we construct a vector autoregression (VAR) model that explicitly captures this time variation, using the Bayesian approach for VARs with instruments (commonly called proxy VARs) by Caldara & Herbst (2019). To motivate this approach further, we show that a standard VAR estimated via ordinary least squares (OLS) delivers forecast errors that are highly correlated with the instrument during periods where our approach estimates the instrument to be informative[1]

[1]These periods broadly line up with the aforementioned periods that stand out in

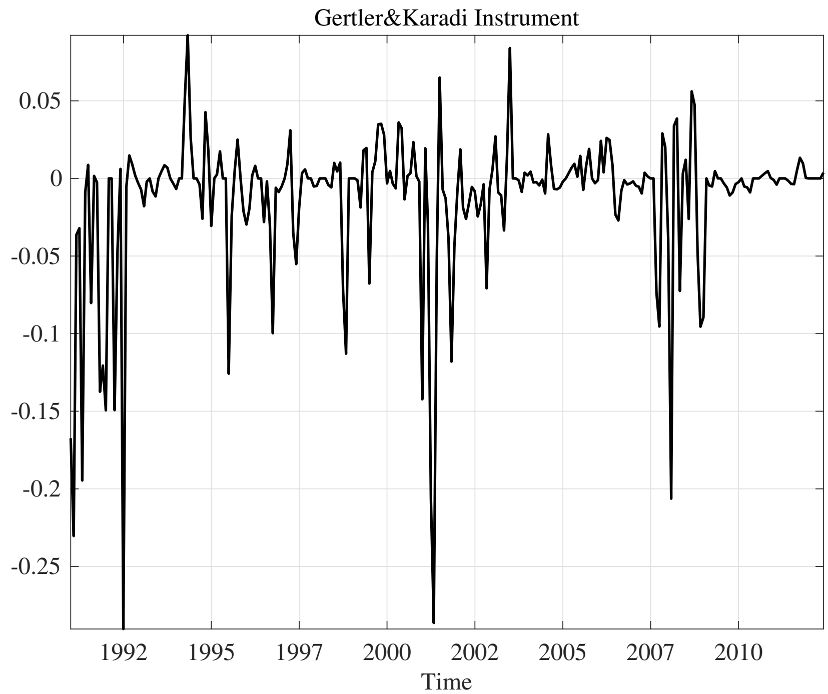


Figure 1: Surprise in 3-month-ahead Fed Funds futures (Gertler & Karadi 2015).

but not otherwise. Our model leaves intact the standard assumption of fixed coefficients in the structural VAR relationship itself (Baumeister & Hamilton 2015, Mertens & Ravn 2013, Caldara & Herbst 2019, Arias et al. 2021). This makes our approach and results directly comparable to the large majority of the structural VAR literature, particularly the part of the literature that uses external instruments.[2] Our approach comes at negligible additional computational cost relative to the previous literature. We interpret changes in the relationship between the instrument and the monetary policy shock as coming from changes in the parameter linking those two variables in a measurement equation. While changes in the volatility of monetary policy shocks or the noise part of the instrument could in theory rationalize Figure 1, we use Monte Carlo experi-

Figure 1.

[2]Mumtaz & Petrova (2021) estimate time-varying parameter VARs with external instruments, but in their application the relationship between the instrument and the shock of interest is actually time invariant.

ments to make the case that these specifications are not consistent with our other findings.

Our approach yields two important insights. First, we can infer periods where the instrument is most informative about monetary policy shocks, thus helping to answer the question as to where identification is coming from and allowing us to develop a narrative for identification. As such, our approach can be seen as complementing the narrative sign restrictions approach of Antolín-Díaz & Rubio-Ramírez (2018), who impose identification via sign restrictions (and related restrictions) for only certain periods. In fact, as shown by Plagborg-Møller & Wolf (2021) and highlighted by Giacomini et al. (2022), narrative sign restrictions can be recast as binary instruments. Our approach instead identifies informative periods for a given instrument. We find, for a standard US instrument, that high-frequency-based instruments for monetary policy shocks are only relevant for a small number of distinct periods. We show that even when we set 90 percent of the instrument observations for the standard Gertler & Karadi (2015) instrument to zero (while keeping those periods our approach estimates to be the most informative), we can recover the same impulse responses as when we use all available observations.

Second, because inference about the monetary policy shock is no longer contaminated by periods where the instrument is not actually informative (our algorithm discounts information contained in the instrument from these periods), we can gain a clearer picture of the effects of monetary policy shocks. In our application using the same instrument as Gertler & Karadi (2015), the effects on prices are almost 50 percent larger after four years, for example. Error bands for impulse responses are generally *not* wider than their fixed coefficient counterparts. Even in applications where our approach yields similar impulse responses to the benchmark fixed coefficient approach (which is something that is not known a priori), the sharpening of the identification narrative can be crucial for interpreting the results.

We next show via Monte Carlo simulations that in cases where both the traditional fixed coefficient approach and our approach are misspecified, either our approach does better (if misspecification is not too severe in a sense that we make precise) or both approaches yield very similar results (such as in the case of changes in the volatility of the true monetary policy shock or changes in the volatility of the noise term inherent in the instrument). More broadly, we make the case that allowing for volatility changes in the noise term in our specific applications is unappealing because it translates into a prior that puts very little faith in the instrument. The additional computational cost of our approach is small, so there seems little reason not to allow for a time-varying relationship between instruments and structural shocks in empirical applications.

The use of instruments in macroeconomics to identify the effects of monetary policy shocks was pioneered by Romer & Romer (2004), who estimate a sophisticated monetary policy rule using real-time data and obtain their instrument as the residual in that estimated monetary policy rule. More recently, the focus has shifted toward using instruments based on high-frequency variation in asset prices, first in event studies (Kuttner 2001, Gürkaynak & Wright 2013, Faust et al. 2007) and later as an instrument incorporated into time series models (Gertler & Karadi 2015, Jarociński & Karadi 2020, Caldara & Herbst 2019, Miranda-Agrippino & Ricco 2020), building on the work of Mertens & Ravn (2013), who introduced the proxy VAR framework.[3] Wolf (2020) highlights how an instrument-based approach mitigates issues that can arise when using standard sign restrictions to identify monetary policy shocks, giving us even more reason to dig deeper into the underpinnings of this approach. Miranda-Agrippino & Ricco (2020) develop an instrument that is also based on high-frequency-based asset price variation around FOMC meetings but further controls for information that the Federal Reserve had at the time of its meeting as well as possible autocorrelation in the instrument. We show in Section 3.4 that with this instrument, we also find relatively rare spikes in instrument relevance. The differences between the impulse responses using the standard approach and our method are substantially smaller with this instrument than with the Gertler & Karadi (2015) instrument. In fact, the impulse responses obtained using this instrument are similar to those obtained with our approach and the Gertler & Karadi (2015) instrument. However, our approach does not require central bank forecasts to clean the original instrument and thus provides a general purpose technology that can also be used for other instruments and shocks. Other papers directly use information from high-frequency variation in asset prices around monetary policy decisions as a right-hand side variable for regressions to estimate the effects of monetary policy shocks (Campbell et al. 2016, Nakamura & Steinsson 2018).

Section 2 lays out our framework and discusses various modeling choices. Section 3 discusses the results for the US, robustness with respect to the instrument used, and the modeling assumptions linking the instrument and the monetary policy shock. Section 4 examines Monte Carlo experiments using the Smets & Wouters (2007) model, and Section 5 concludes.

[3]The use of this type of identification is becoming more common. For example, Känzig (2021) uses a high-frequency-based identification to identify oil shocks.

2 A VAR Model to Study Changes in Instrument Relevance

We set out to study the response of an n dimensional vector of observables $\mathbf{y}_t$ to a monetary policy shock e_t^{MP}, which is one element of the n dimensional vector of structural shocks $\mathbf{e}_t$.[4] To estimate said response, we use a structural vector autoregression (SVAR) in equation (1):

$$\mathbf{y}_t = \mathbf{c} + \sum_{\ell=1}^{\mathcal{L}} \mathbf{A}_\ell \mathbf{y}_{t-\ell} + \mathbf{\Sigma} \mathbf{e}_t, \tag{1}$$

where $\mathbf{e}_t \sim_{iid} N(\mathbf{0}, \mathbf{I})$.

The well-known identification problem in Gaussian structural VARs (Canova 2011, Kilian & Luetkepohl 2018) implies that we need additional information to identify the column of the response matrix $\mathbf{\Sigma}$, which tells us how the elements of $\mathbf{y}_t$ respond to the monetary policy shock e_t^{MP}. The additional information that we exploit, following a substantial fraction of the recent literature in empirical macroeconomics, is an instrument m_t for the monetary policy shock e_t^{MP}. There are various frequentist (Mertens & Ravn 2013, Stock & Watson 2018) and Bayesian (Arias et al. 2021, Caldara & Herbst 2019, Drautzburg 2020) approaches to incorporating such information in an SVAR analysis.[5] Since our ultimate goal is to study possible changes in the relationship between the observable m_t and the unobserved monetary policy shock, we explicitly state the re-

[4] We use boldface for vectors and matrices.

[5] Frequentist inference using proxies in dynamic factor models was introduced by Stock & Watson (2012).

lationship we want to estimate in equation (2):[6]

$$m_t = \beta_t e_t^{MP} + \sigma_m u_t^M. \tag{2}$$

The key identification assumptions are twofold. First, we assume that u_t^M, which is distributed independently and identically over time as $N(0,1)$, is also independent of all other shocks in our model, both the vector of structural shocks e_t and any shocks determining the evolution of β_t, generally denoted as w_t. Second, the instrument is informative for the monetary policy shock, meaning that at least for some periods, $\beta_t \neq 0$. We can formalize these assumptions as

$$E\left[e_{j,t} m_t\right] = 0 \quad \text{for } j = 2, \ldots, n, \qquad \text{(exogeneity)} \tag{3}$$

$$E\left[u_t^M \mathrm{e}_t\right] = 0 \quad \text{and} \quad E\left[u_t^M w_t\right] = 0, \tag{4}$$

$$E\left[e_t^{MP} m_t\right] = \beta_t \neq 0 \quad \text{for some } t, \qquad \text{(relevance)} \tag{5}$$

where $e_{j,t}$ denotes the jth element of e_t and w_t is the time-t innovation to β_t.

Our Bayesian estimation approach will still be valid if $\beta_t = 0 \quad \forall t$. In that case, the impulse responses are not identified. In particular, our approach will automatically approximate the posterior distribution of β_t and the associated instrument reliability for each time period t. If those are always small (i.e., standard posterior bands include zero), we can infer that the instrument is weak.[7] We thus borrow the approach of directly

[6]In Section 3.4 we show that our results are robust when allowing for lagged macroeconomic variables to enter the right-hand side of equation (2), similar to specifications used in Arias et al. (2021) and Plagborg-Møller & Wolf (2021).

[7]Our approach also assumes invertibility of the monetary policy shock. For our monetary policy application, this seems to be a widely accepted assumption (Wolf 2020). For recent work on the link between inference using instruments and invertibility, see Miranda-Agrippino & Ricco (2022).

estimating the parameters of this measurement equation from Caldara & Herbst (2019). Unlike in that paper, we allow for changes in the parameter β_t governing the systematic relationship between these two variables.[8]

A useful summary statistic for assessing the strength of the instrument in different periods is a time-varying version of the common reliability (or relevance) statistic ρ_t:

$$\rho_t \equiv \frac{\beta_t^2}{\beta_t^2 + \sigma_m^2}. \tag{6}$$

This statistic represents a time-t approximation to the squared correlation between the instrument and the structural shock. The approximation is exact if parameters are constant.

We next estimate three specifications for β_t: (i) a constant parameter specification reminiscent of Caldara & Herbst (2019) as a benchmark (equation (7)), (ii) a random walk specification in the tradition of the literature on time-varying parameters in state space models and VARs (Cogley & Sargent 2002, Primiceri 2005, Stock & Watson 2007), and (iii) a Markov-switching specification that allows for infrequent changes in β_t (Hamilton 1989, Sims & Zha 2006). We assume that any random innovations to β_t are independent of u_t^M and e_t.

[8]Following Caldara & Herbst (2019), we normalize the relevant column of Σ so that the monetary policy shock increases interest rates on impact. Such a sign normalization is necessary for any structural VAR identification scheme. In our specific application, it furthermore allows us to center the prior for β_t at zero while still maintaining a standard interpretation of the estimated monetary policy shock.

$$\text{Constant: } \beta_t = \beta, \tag{7}$$

$$\text{Random Walk: } \beta_t = \beta_{t-1} + \sigma_\beta w_t, w_t \sim_{iid} N(0,1), \tag{8}$$

$$\text{Markov Switching: } \beta_t = \beta_{s_t}, Pr(s_t = i | s_{t-1} = j) = p_{ij}. \tag{9}$$

To approximate the posterior of our model, which consists of equations (1), (2), and one of the equations (7), (8), or (9), we modify the Metropolis-within-Gibbs sampling framework of Caldara & Herbst (2019) (the specification with equation (7) is exactly their specification). One important feature of our algorithm is that we do not require the same number of observations for the instrument m_t as for the macro variables collected in $\mathbf{y}_t$. Details about the algorithm can be found in Appendix A.

Two possible extensions of our model are immediate. First, instead of one instrument, we could use multiple instruments. In that case we would need to make a decision about possible correlation in the error terms of the measurement equations, for example. Since in our application the instruments yield similar results when used on their own, we do not pursue this extension here.

More importantly, we could allow for stochastic volatility in the law of motion for the parameter β_t.[9] If we had introduced stochastic volatility in equation (2) instead of time-varying parameters, this would change how our model interprets spikes in the instrument. Our current assumption means that spikes or outliers are interpreted as informative events;

[9]One could also, along the lines of Mumtaz & Petrova (2021), introduce stochastic volatility in equation (1). If the changes in volatility in our instrument (see Figure 1) were due to actual changes in the volatility of monetary shocks, we would expect to estimate a relatively smooth evolution of β_t, which is not what we find in our applications. We elaborate on this point in Section 4.

with stochastic volatility, they would be interpreted as noise. In most of our applications it turns out that the instruments are generally not very informative (low ρ_t) except for clearly delineated short periods of high reliability. Thus, using stochastic volatility would imply a prior that, in these specific applications, would put very little faith in the instruments.[10] This stands in contrast to standard priors in the proxy VAR literature (Arias et al. 2021, Caldara & Herbst 2019) that impose a prior that implies that the instruments are indeed useful/reliable.

Our assumption of time-varying parameters is thus best seen as a context-specific prior choice that implies that there is at least *some* instrument reliability, in line with the previous literature. We will show in the following sections that the periods our algorithm identifies as informative are generally those anecdotal evidence also identifies as periods where there was substantial uncertainty about the conduct of monetary policy and hence meaningful monetary policy shocks. We show in Section 4 via Monte Carlo experiments that our approach performs as well as the standard fixed coefficient approach when it is misspecified and there is either stochastic volatility in the measurement equation that is unrelated to monetary policy or there is stochastic volatility in the monetary policy shock. More importantly, those specifications for the data-generating process lead to estimated paths of β_t that are inconsistent with our findings based on US data, hence providing evidence for our modeling assumptions.

[10] Results with stochastic volatility will thus give results similar to one of the robustness checks in Section 3.3.

3 Effects of Monetary Policy Shocks Identified via High-Frequency Variation in Asset Prices

Section 3.1 studies the effects of allowing for time-varying reliability using the benchmark random-walk specification. Section 3.2 instead uses a Markov-switching approach to model time variation. In Section 3.3 we highlight that indeed only a few periods are informative for the effects of monetary policy shocks by estimating our model using an instrument that is set to zero except for the most informative periods. Finally, Section 3.4 estimates the effects of monetary policy shocks using the Miranda-Agrippino & Ricco (2020) instrument and an alternative version of the Gertler & Karadi (2015) instrument.

3.1 The Effects of Time-Varying Reliability

We first contrast the constant parameter specification with the random walk specification. Our application uses US data:[11] y_t consists of the log of the Consumer Price Index (CPI), the log of industrial production (IP), the interest rate on one-year government bonds i, and the excess bond premium (EBP) (Gilchrist & Zakrajsek 2012). As Caldara & Herbst (2019) highlight, including a measure of financial conditions like the EBP in our VAR is crucial in order to get the effects of monetary policy right. The sample for y_t runs from July 1979 to June 2012. We follow Gertler & Karadi (2015) in our choice of the instrument m_t and use the surprise in the three-month-ahead Fed Funds futures around FOMC meetings (the

[11] In Appendix E we apply our approach to UK data, using the high-frequency instrument of Cesa-Bianchi et al. (2020).

series depicted in Figure 1). The sample for m_t is January 1991 to June 2012. We use 12 lags in all VARs estimated on US data in this paper.

To motivate our assumption of variation in β_t, it is instructive to study the relationship between the instrument and various forecast errors implied by a VAR. We estimate a version of our benchmark VAR via OLS to make sure the estimation does not use any information on the instrument (as it would when we estimate our benchmark model using Bayesian methods). We then compute the correlation between the OLS-based one-step-ahead forecast errors for the variables in our VAR and the Gertler & Karadi (2015) instrument, for both periods where, according to our benchmark model, the instrument is informative and periods where it is not. Identification schemes for structural VARs generally posit a linear relationship between these forecast errors and the structural shocks of interest. Thus, these are key correlations that are exploited whenever researchers use an instrument for identification in a structural VAR (as discussed before, standard proxy VARs assume a time-invariant relationship, so this correlation should be constant across subsamples). The exact definition of the periods we use can be found in Section 3.3.

Table 1 shows the correlation for each forecast error in a different column and each period in a different row. In periods where our approach identifies the instrument as not being informative, the absolute value of the correlation between the instrument and the (one-step-ahead) forecast errors decreases by at least 70 percent and by as much as 82 percent, depending on the variable.[12] Furthermore, the absolute decrease in correlation is meaningful (a fall in correlation of 0.34 for inflation and EBP,

[12]The slightly awkward use of the absolute value of the correlation is necessary because the correlation of the instrument with CPI forecast errors becomes positive when the instrument is not informative.

for example). The sign of the correlation for IP might seem unusual, but we see below that after a small initial positive impulse response of IP to a monetary policy shock (confirming the sign of the correlation), the response quickly turns negative.

Table 1: Correlation between instrument and forecast errors in OLS version of our VAR.

	i	CPI	IP	EBP
Instrument is informative	0.48	-0.34	0.10	0.46
Instrument is not informative	0.14	0.06	0.03	0.12
Percent reduction in (absolute) correlation	0.71	0.82	0.70	0.74

For the rest of the analysis, we use a Bayesian approach. The priors we use throughout are standard in the literature and are described in detail in Appendix A. We make the priors as comparable as possible across the different specifications: the same parameters always have the same priors. Furthermore, the prior for β in the fixed coefficient variant is the same as the prior for β_0 in the random walk specification. Estimation results for models with time-varying coefficients can be somewhat sensitive to the choice of prior for the innovation standard deviation σ_β in the law of motion for the parameter. This parameter governs the amount of time variation. Sensitivity is less of an issue here because (i) we only have one time-varying parameter (in contrast to papers where all VAR parameter can vary, such as Cogley & Sargent 2002 and Primiceri 2005), and (ii) we do not have stochastic volatility in our model, which helps sharpen inference. Nonetheless, to make sure that this is not an issue, we follow some of our previous work (Amir-Ahmadi et al. 2020) and estimate the hyperparameters that enter the prior for σ_β. Details on the priors can be found in Appendix A.

Figure 2 shows the posterior path of β_t and ρ_t. In gray we plot the corresponding elements of the fixed coefficient version. We plot the posterior median as well as 68 percent equal-tail posterior bands.[13] It is striking that there are few short periods of high instrument relevance when allowing for time variation in β.[14] Three periods stand out, the first being the first half of the 1990s. The large value of β_t is not driven by our prior as our prior for the initial value of β_t is centered at zero. Instead, the first half of the 1990s was characterized by relatively high inflation at the beginning as well as a (mild) recession. Our model highlights the period coming out of the 1990s recession when annual CPI inflation was still high in 1991 (4.2 percent) as a period where the Federal Reserve was surprisingly accommodative (see Figure 1).

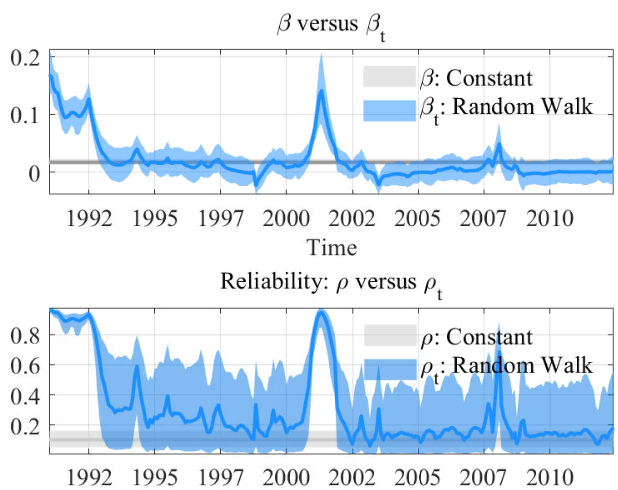


Figure 2: Posterior of β_t and ρ_t (median and 68 percent posterior bands).

The second period our model highlights is around September 2001. Unsurprisingly, the exact actions of the Federal Reserve were hard to

[13]All posterior bands in this paper are 68 percent posterior bands.

[14]The posterior median reliability of our time-varying parameter specification is almost always larger than its fixed coefficient counterpart because the estimated variance of the noise part is substantially larger in the fixed coefficient version—in the fixed coefficient case part of the time variation is soaked up in the noise term. This is also evident from the posterior of β_t.

predict around that time even though the direction was clear. The third period of high instrument relevance is the Great Recession around 2008. Our framework thus helps us understand what information is contained in the instruments.

We next examine if this time variation in instrument relevance matters for impulse responses.

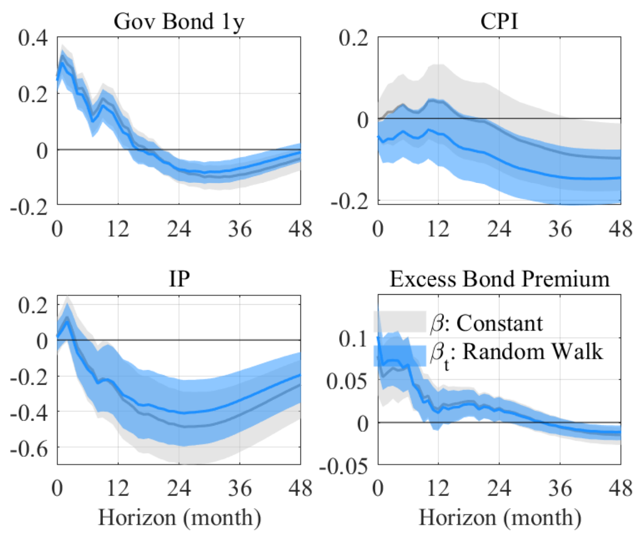


Figure 3: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock.

Figure 3 shows the impulse to a one standard deviation monetary policy shock under the fixed coefficient (gray) and random walk (blue) specifications. We plot the posterior median as well as the 68 percent error bands. For bond yields, IP, and the EBP, the impulse responses are similar. For log CPI, the differences are instead *substantial*. With fixed coefficients, we see a price puzzle appearing, whereas this is not the case at all for the posterior median of the impulse responses when we allow for time variation in instrument relevance. Furthermore, the response of log CPI is larger in magnitude—after four years, the posterior median of the response is almost 50 percent larger when we allow for time variation in instrument relevance. Our approach discounts periods where the instrument is not informative and hence can lead to substantially different impulse responses. As mentioned before, this does not come at a cost in terms of the width of the error bands in our example.

Looking back at Figure 2, one possible criticism is that the estimated path for β_t might be better characterized by a Markov-switching model. We think of the random walk as our benchmark exactly because it is flexible enough to approximate many patterns of time variation, including sudden changes as observed in Figure 2. Nonetheless, in the next section we estimate a two-state Markov-switching specification and show that it yields very similar results.

3.2 A Markov-Switching Alternative

We now change the law of motion for β_t. The only difference between the Markov-switching specification and the random walk benchmark is the law of motion for β_t as detailed in equations (8) and (9), respectively. Figure 4 shows the impulse response of log CPI to a one standard deviation monetary policy shock in the two-state Markov-switching model for β_t. We focus here on the response of CPI because that is where the major differences between fixed coefficient and time-varying parameter results occurred in the previous section. That impulse response is very similar to the random walk specification.

Figure 5 plots instrument relevance for our benchmark random walk specification in blue and the two-state Markov-switching model in red. We can see that both specifications identify largely the same periods of high instrument relevance. The random walk specification is somewhat conservative in that it has fewer spikes, but this does not lead to any

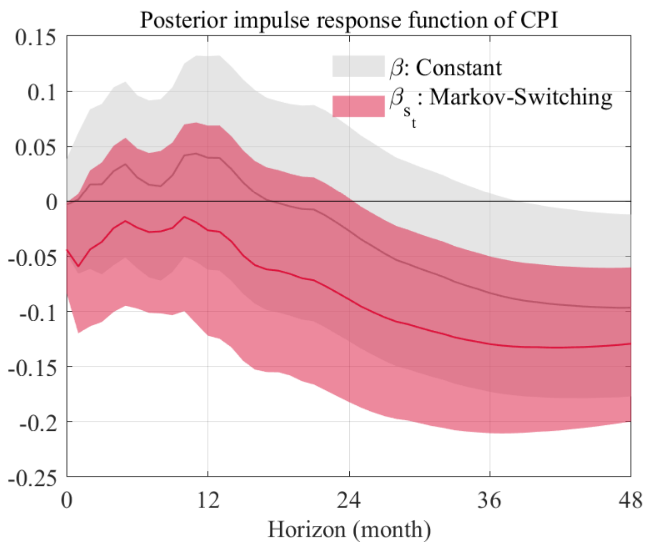


Figure 4: Impulse response of log CPI for Markov-switching specification (median and 68 percent posterior bands).

meaningful difference in impulse responses, as discussed above. The choice for a specific law of motion for β ultimately comes down to the application in mind as well as preferences. We recommend the random walk as a default choice because of its flexibility.

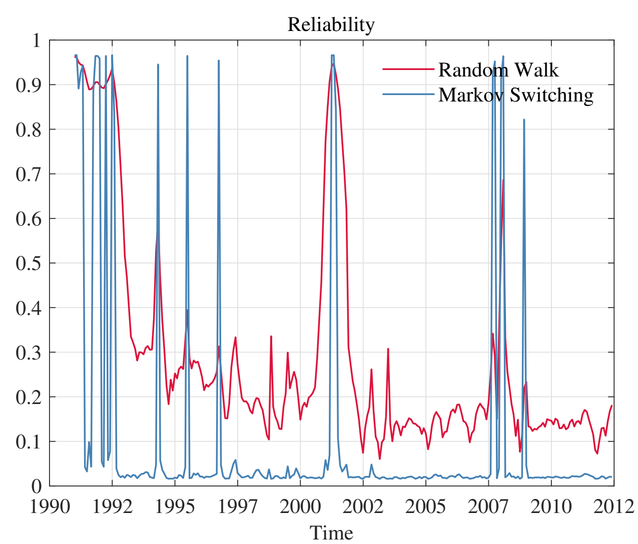


Figure 5: Posterior reliability for Markov-switching and random walk specifications (posterior median).

3.3 Shutting Down Periods Where the Instrument Is Informative/Not Informative

To get a better sense of the role that periods with high instrument relevance play in shaping the posterior distribution of the impulse responses, we now carry out two diametrically opposite thought experiments. First, we compute the posterior probability that $\beta_t = 0$ for each time period t, using the approach in Koop et al. (2010) and our original instrument m_t. We then create two instruments, $\tilde{m}_t$ and $\overline{m}_t$, from our instrument according to the following two rules:

1. $\tilde{m}_t = m_t$ if $Pr(\beta_t = 0) < 0.5$, $\tilde{m}_t = 0$ else

2. $\overline{m}_t = m_t$ if $Pr(\beta_t = 0) \geq 0.5$, $\overline{m}_t = 0$ else

$\tilde{m}_t$ only keeps the original realizations of the instrument that our model deems informative, whereas $\overline{m}_t$ only keeps relatively uninformative realizations. The threshold probability of 0.5 only selects the early 1990s and 2001 as informative periods.

Figure 6 shows the results when we use $\tilde{m}_t$ as our instrument. For the sake of comparison, the fixed coefficient VAR in this figure uses our original instrument m_t. We see that our approach still estimates the same periods to have high instrument relevance.[15] The impulse responses (we highlight CPI in this figure but show all responses in the Appendix) are very similar to those in our original setting, making clear that it is indeed *only* those high instrument relevance periods that inform the impulse responses. Naturally, this result depends on the specific application. Had the instrument relevance been reasonably high outside of the

[15]To economize on notation, we also call this parameter β_t, but it is a different object from β_t when we use the instrument m_t.

spikes in the instrument relevance we document, the procedure in this section would have led to a meaningful loss of information.

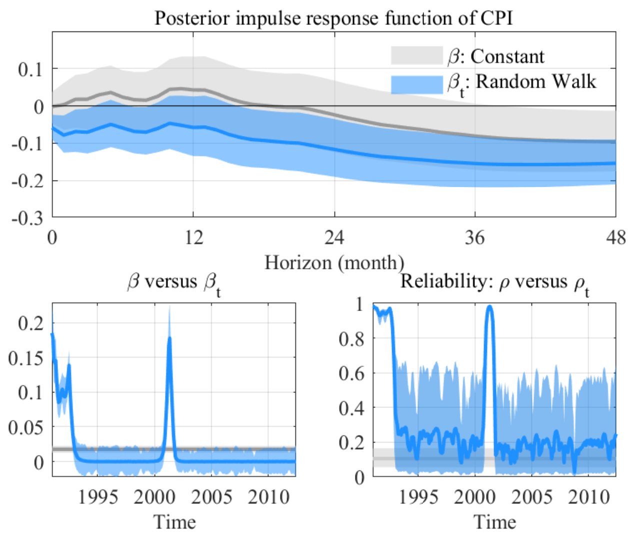


Figure 6: Results for $\tilde{m}_t$ (median and 68 percent posterior bands). The fixed coefficient VAR is based on the original m_t instrument.

Figure 7 plots the corresponding results when we only keep the original instrument if its relevance is low. Zero is now included in the 68 percent posterior bands for all horizons. Posterior instrument relevance is low for all periods. Even though we keep 90 percent of the observations from the original sample, there is little information contained in those observations. In the Appendix (Figures A-8 and A-9) we show impulse responses obtained with a fixed coefficient VAR and these modified instruments—the resulting impulse responses are basically indistinguishable from the responses obtained with a fixed coefficient VAR and our original instrument. In order to effectively exploit the instrument when it is informative, we need allow for time variation so that the instrument is not used for identification when it is not informative.

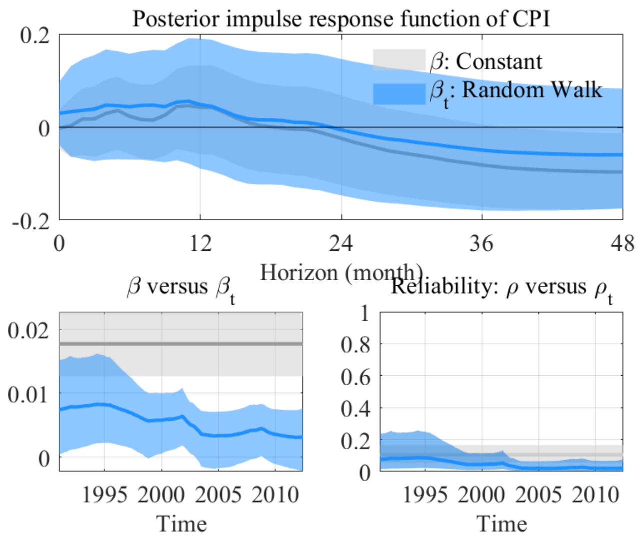


Figure 7: Results for $\overline{m}_t$ (median and 68 percent posterior bands). The fixed coefficient VAR is based on the original m_t instrument.

3.4 Alternative Instruments

Our various observation equations linking m_t and the unobserved monetary shock all imply that m_t is iid, borrowing from Caldara & Herbst (2019). Other papers in the literature have posited more flexible relationships where the instrument can be contaminated by past macro variables and/or lags of the instrument. To assess whether this is an issue in our application, we progress in two steps. First, we regress our instrument on two lags of itself and the variables y_t in the VAR. The key results are summarized in Figure 8 and are very similar to our benchmark. The only difference is that the spike in ρ_t and β_t surrounding the Great Recession is less pronounced. The impulse response of CPI is basically unchanged (other impulse responses can be found in Appendix D).

We then use the instrument introduced by Miranda-Agrippino & Ricco (2020). The authors start off with a high-frequency-based instrument like our benchmark choice but then remove any autocorrelation and information available to the FOMC at the time of their meetings (as en-

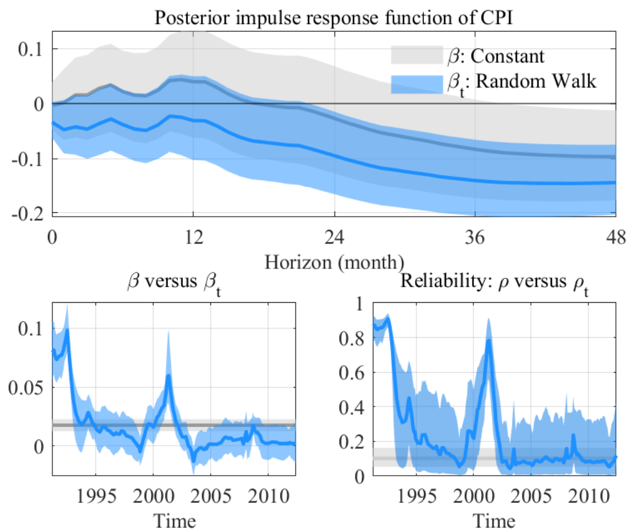


Figure 8: Results for the case of the modified instrument (median and 68 percent posterior bands).

coded in the Green Book).[16] The sample for this instrument is January 1991 to December 2009; it is shorter due to the need for Green Book data, which is published with a lag. Figure 9 shows that for this instrument, the largest spike in β_t by far now appears around 2001. We still see a clear tightening of the error bands for instrument reliability in the early 1990s and around the Great Recession as well, but these movements are less pronounced than in our benchmark.

Interestingly, the posterior median path of instrument reliability is substantially higher than in our benchmark or in Figure 9. Since the reliability does not fall as much between spikes as in our benchmark, it is not surprising that the difference between the fixed coefficient version of the impulse response of log CPI and its random walk counterpart are very similar—and both are similar to the random walk–based results from our benchmark instrument. While using our approach with the Miranda-

[16]As highlighted recently by Bauer & Swanson (2022), in this context the information content of professional forecasts is very similar to those forecasts contained in the Green Book.

Agrippino & Ricco (2020) instrument does not change the conclusions in terms of impulse responses, it adds substantial interpretability. For example, the most informative period (largest β_t value and largest instrument reliability) is around 2001.

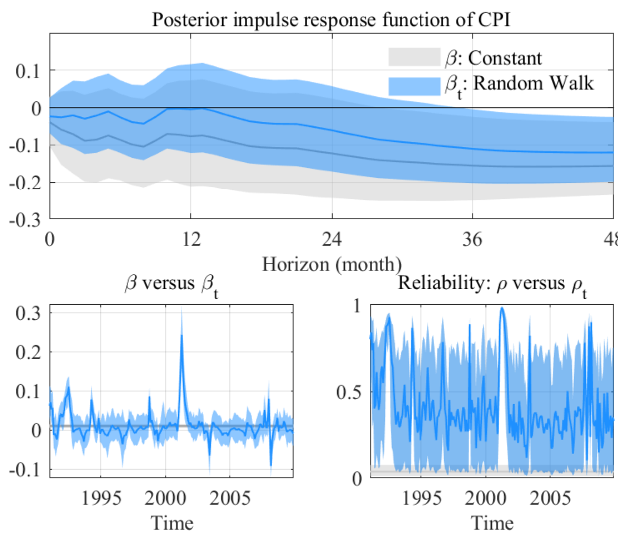


Figure 9: Results for the Miranda-Agrippino & Ricco (2020) instrument (median and 68 percent posterior bands).

4 Monte Carlo Studies

To assess the properties of our approach, we now turn to a series of Monte Carlo experiments (details on the exact calibration of the data-generating processes can be found in Appendix B). We follow Wolf (2020) and use the Smets & Wouters (2007) model as a laboratory.[17] We use three observables in our VAR: output, inflation, and nominal interest rates. Since the Smets-Wouters model is a quarterly model, we set the lag length in our

[17] Wolf (2020) studies an instrumental variables approach in VARs, but in addition to his assumption of fixed coefficients, there are two substantial differences relative to our setup: we use a standard sample size in our simulations, whereas Wolf (2020) studies population properties, and we introduce measurement error in our instrument, which we calibrate to have 25 percent of the variance of the actual monetary shock in our first experiment.

VAR to four. We simulate the instrument using equation (2) for various specifications of β_t, the volatility of the monetary shock, and the volatility of the measurement error σ_m.[18]

The first Monte Carlo experiment uses time variation in the measurement equation of the instrument that is along the lines of our estimated models (even though the random walk specification that we use here is still misspecified). The data-generating process features parameter changes in β_t. We choose an extreme scenario where β_t can only take on the values zero and one. The path of β_t in the data-generating process is fixed across all Monte Carlo samples for this first experiment.
We simulate 100 samples of length 250 using the posterior mode as in Wolf (2020). For approximately 10 percent of those periods,[19] we set $\beta_t = 1$ in the data-generating process; otherwise it is zero (hence the instrument is just noise). Figure 10 shows the true impulse response to a one standard deviation monetary shock in black as well as the Monte Carlo average of the 68 percent posterior bands for our approach and the fixed coefficient version. Our results confirm those of the population analysis in Wolf (2020): even without perfect invertibility, the true responses are well approximated by our VAR. The fixed coefficient version generally has wider error bands, which leads the average posterior bands for output to include 0 for all horizons and a more pronounced probability of a price puzzle for inflation.

Figure 11a shows the estimated posterior median of instrument reliability in blue and the true reliability path in black (periods with high values of β, and hence higher instrument volatility, are denoted by gray

[18] One difference between the equation used to simulate the instrument and equation (2) in our estimated model is that the simulated monetary policy shock does not have unit variance, a point we come back to below.

[19] From periods 1 to 4, 141 to 149, and 231 to 250.

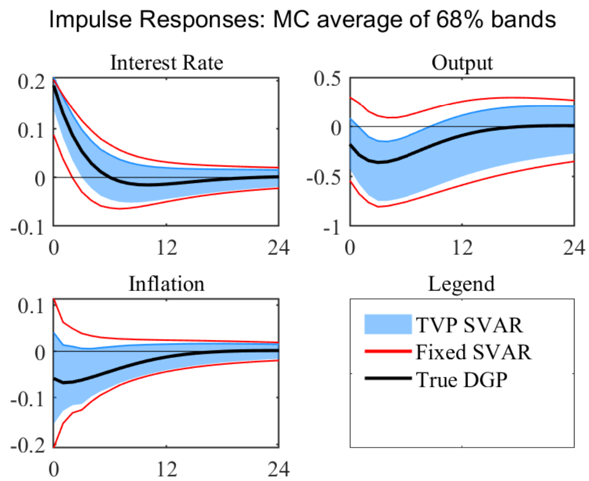
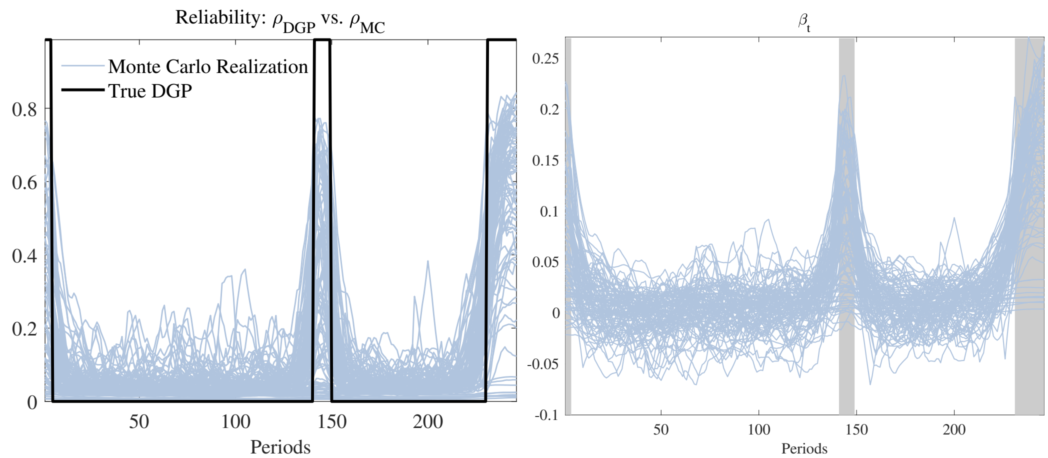


Figure 10: Impulse responses for the data-generating process with time-varying β_t and the Monte Carlo replications, first Monte Carlo experiment.

bars). Our approach captures the changes in instrument reliability well.

It is instructive to directly study the posterior median paths of β_t. As we discuss in more detail in Appendix B, we must rescale the true β_t values described there by the standard deviation of the monetary shock to make them comparable to our estimation results since our estimated model assumes monetary policy shocks with unit variance; all scale effects in our estimated model are captured by Σ. In this Monte Carlo exercise, the properly rescaled true β_t values (which can be directly compared to the estimated values) are 0 and 0.23 (the original values multiplying the monetary policy shock with non-unit variance were 0 and 1).



(a) Reliability (posterior medians).

(b) Posterior medians of β_t.

Figure 11: Reliability and posterior paths of β_t, first Monte Carlo experiment.

The posterior paths plotted in Figure 11b closely resemble the true paths and, more importantly, the patterns we find in US data: close to 0 for most periods, with distinct increases when the true value is non-zero.

Next we study two more Monte Carlo experiments where our specification of the measurement equation for the instrument is more severely misspecified. We focus here on the posterior impulse responses and the posterior median paths for β_t. First, we simulate the data so that there is stochastic volatility in the measurement error u_t^M. We choose parameter values to keep the overall volatility of the instrument at each point in time to be the same as in the benchmark case discussed immediately above (see Appendix B for details). Figure 12 shows that our approach performs as well as the fixed coefficient version that is standard in the literature even though both are misspecified in different ways. It is not surprising that both specifications do reasonably well in this specification since β_t is always non-zero in the data-generating process, so the instrument conveys some information about the true monetary policy shock each period, in contrast to the previous experiment.

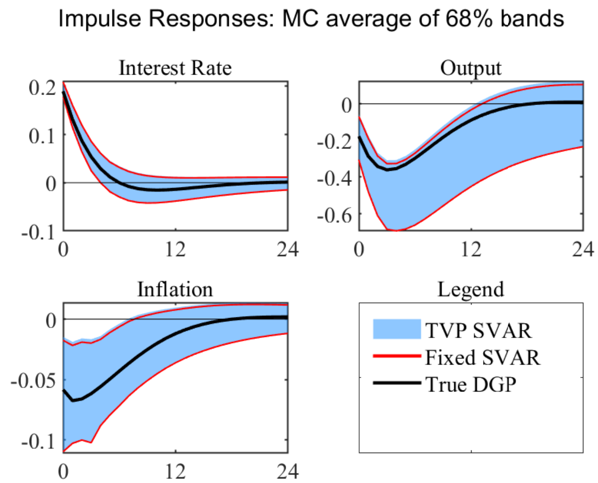


Figure 12: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in measurement error.

Turning to the posterior paths of β_t, the properly rescaled true β value in this experiment (which can be directly compared to the estimated values) is $0.5 * 0.23 \approx 0.12$. The posterior paths plotted in Figure 13 are remarkably stable and do generally not hover around zero, a stark contrast to the posterior paths obtained using US data.

Finally, we ask how our approach fares when confronted with data where there is stochastic volatility in the true monetary policy shock.[20] We again keep the paths of the instrument's volatility the same as in our benchmark specification. Figure 14 shows that our approach again performs as well as the fixed coefficient version.[21] Intuitively, since the coefficients of the estimated VAR itself are constant over time, the estimated one-step-ahead forecast errors will, in this situation, feature changes in volatility even if the estimated model is misspecified because the esti-

[20]We solve the model linearly and then change the volatility of the monetary policy shock in some periods, along the lines of Justiniano & Primiceri (2008).

[21]Since the volatility of the monetary policy shock changes over time in the data-generating process, we scale the impulse response for the data-generating process so that the variance of the monetary shock is the simple average of the two possible realizations of the variance we consider.

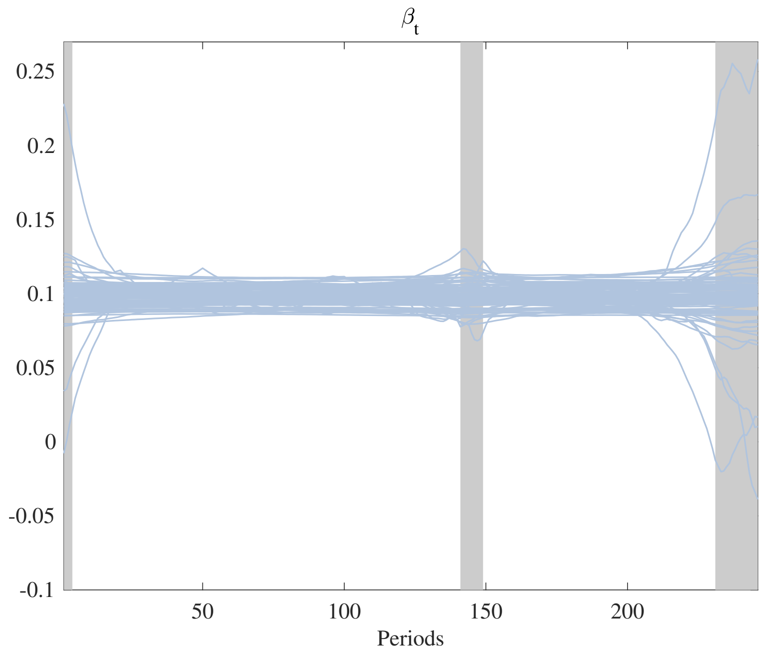


Figure 13: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of v_t.

mated forecast errors will be a function of the true forecast errors (although they are not equal in finite samples).

Our measurement equation for the instruments then links instruments that feature changes in volatility by construction with estimated structural shocks, which inherit the changes in volatility from the estimated one-step-ahead forecast errors. Both the left- and right-hand side variables in the measurement equation are thus influenced by the same changes in volatility, and we find estimated paths for β_t that do not change much over time and do not decrease toward zero, in contrast to those obtained using US data or in the first experiment (see Figure 15). Due to the relative stability, it is then also not surprising that our approach does as well as the fixed coefficient approach in this case. Comparing across experiments, the reason that both algorithms have a harder time identifying significant effects in this experiment is due to our calibration implying that the monetary policy shock is less volatile in most periods compared to the second experiment. This final Monte Carlo ex-

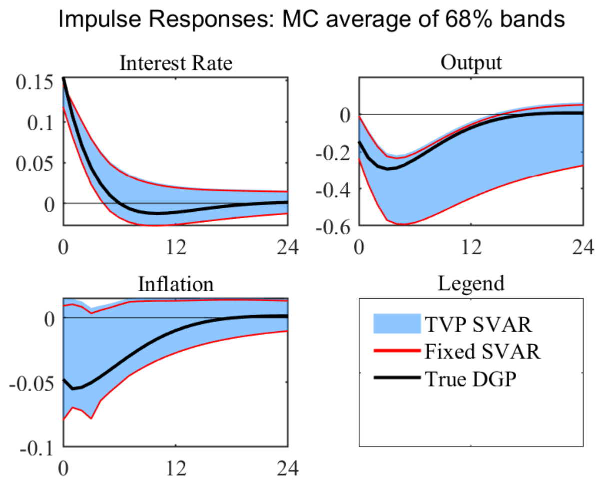


Figure 14: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in the monetary policy shock.

periment involves one subtlety that was not present in the earlier experiments: Having knowledge of the true data-generating process, we could rescale the true β_t value (which is constant) by taking into account the changing volatility of the true monetary shock to give β_t values of 0.12 and 0.26 that multiply rescaled versions of monetary policy shock that have unit variance throughout the sample.[22] Such a transformation is always possible in models with stochastic volatility. This does, however, not mean that this last experiment is similar to the first experiment: In the first experiment the VAR in equation (1) is correctly specified, whereas in this last experiment the VAR is misspecified because it assumes constant volatility of the shocks. As discussed before, our estimated monetary policy shock (which in a correctly specified world would have unit

[22]In terms of equations, the true data generating process is $m_t = e_t^{MP} + \sigma_m u_t^M = \beta_t \bar{e}_t + \sigma_m u_t^M$, where $\bar{e}_t$ is the standardized monetary policy shock that has unit variance, while e_t^{MP} features changes in volatility.

variance) inherits changes in volatility via the VAR that is estimated to have fixed coefficients and fixed forecast error variance. Hence the estimated movements in β_t are much more muted than what we would get if the true unit variance (i.e., rescaled) monetary policy shock from the data-generating process were observable and we would directly estimate the measurement equation for the instrument. Nonetheless, the average level is broadly in line with the rescaled values discussed above.[23]

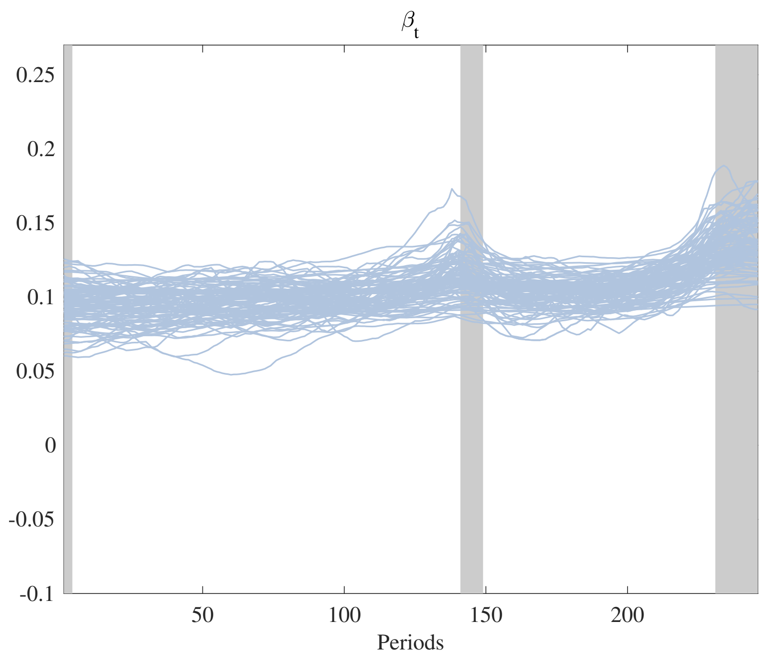


Figure 15: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of the monetary policy shock.

Summing up, only the first Monte Carlo experiment can qualitatively capture the patterns of the posterior β_t paths that we obtained in our empirical applications. We view this as substantial evidence in favor of our modeling assumptions.

[23]To convince readers that the issue of rescaling the monetary policy shock has no impact on our results, we show in Appendix B.3.1 that the estimated impulse responses remain unchanged in a hypothetical scenario where the instrument is directly linked to the period-by-period rescaled (and thus unit variance) monetary policy shock.

5 Conclusion

In this paper we study how instrument relevance changes over time in a common application of instrument-based identification in structural VARs. We find substantial time variation in instrument relevance, thus allowing us to isolate periods where instruments are informative, helping to build a narrative for a given instrument. Furthermore, our approach can substantially alter conclusions by discounting periods where the instrument is not informative, as in the case of the Gertler & Karadi (2015) instrument. As a practical recommendation, we show in our application that "cleaning" an instrument and removing periods that are not informative will generally not help unless a researcher is willing to model time variation in instrument relevance.

While we focus in our application on monetary policy shocks, the estimation approach we develop is general and can be used for any application of external instruments in VARs, such as the effects of government spending shocks, tax shocks, or financial shocks.

References

Amir-Ahmadi, P., Matthes, C. & Wang, M.-C. (2020), 'Choosing Prior Hyperparameters: With Applications to Time-Varying Parameter Models', *Journal of Business & Economic Statistics* **38**(1), 124–136.
URL: *https://ideas.repec.org/a/taf/jnlbes/v38y2020i1p124-136.html*

Antolín-Díaz, J. & Rubio-Ramírez, J. F. (2018), 'Narrative sign restrictions for svars', *American Economic Review* **108**(10), 2802–29.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20161852*

Arias, J. E., Rubio-Ramírez, J. F. & Waggoner, D. F. (2021), 'Inference in Bayesian Proxy-SVARs', *Journal of Econometrics* **225**(1), 88–106.
URL: *https://ideas.repec.org/a/eee/econom/v225y2021i1p88-106.html*

Bauer, M. D. & Swanson, E. T. (2022), 'A reassessment of monetary policy surprises and high-frequency identification', *NBER Macro Annual* .

Baumeister, C. & Hamilton, J. D. (2015), 'Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information', *Econometrica* **83**(5), 1963–1999.
URL: *https://ideas.repec.org/a/wly/emetrp/v83y2015i5p1963-1999.html*

Caldara, D. & Herbst, E. (2019), 'Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs', *American Economic Journal: Macroeconomics* **11**(1), 157–192.
URL: *https://ideas.repec.org/a/aea/aejmac/v11y2019i1p157-92.html*

Campbell, J. R., Fisher, J. D. M., Justiniano, A. & Melosi, L. (2016), Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis, *in* 'NBER Macroeconomics Annual 2016, Volume 31', NBER Chapters, National Bureau of Economic Research, Inc, pp. 283–357.
URL: *https://ideas.repec.org/h/nbr/nberch/13764.html*

Canova, F. (2011), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton, NJ.

Carter, C. K. & Kohn, R. (1994), 'On Gibbs sampling for state space models', *Biometrika* **81**(3), 541–553.

Cesa-Bianchi, A., Thwaites, G. & Vicondoa, A. (2020), 'Monetary policy transmission in the United Kingdom: A high frequency identification approach', *European Economic Review* **123**(C).
URL: *https://ideas.repec.org/a/eee/eecrev/v123y2020ics0014292120300076.html*

Cogley, T. & Sargent, T. J. (2002), Evolving Post-World War II US Inflation Dynamics, *in* 'NBER Macroeconomics Annual 2001, Volume 16', NBER Chapters, National Bureau of Economic Research, Inc, pp. 331–388.
URL: *https://ideas.repec.org/h/nbr/nberch/11068.html*

Drautzburg, T. (2020), 'A narrative approach to a fiscal DSGE model', *Quantitative Economics* **11**(2), 801–837.
URL: *https://ideas.repec.org/a/wly/quante/v11y2020i2p801-837.html*

Faust, J., Rogers, J. H., Wang, S.-Y. B. & Wright, J. H. (2007), 'The high-frequency response of exchange rates and interest rates to macroeconomic announcements', *Journal of Monetary Economics*

54(4), 1051–1068.
URL: *https://www.sciencedirect.com/science/article/pii/S0304393206001565*

Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, Springer New York.
URL: *https://books.google.com/books?id=f8KiI7eRjYoC*

Gertler, M. & Karadi, P. (2015), 'Monetary Policy Surprises, Credit Costs, and Economic Activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.
URL: *https://ideas.repec.org/a/aea/aejmac/v7y2015i1p44-76.html*

Giacomini, R., Kitagawa, T. & Read, M. (2022), Narrative restrictions and proxies, Technical report, University College London.

Giannone, D., Lenza, M. & Primiceri, G. E. (2015), 'Prior Selection for Vector Autoregressions', *The Review of Economics and Statistics* **97**(2), 436–451.
URL: *https://ideas.repec.org/a/tpr/restat/v97y2015i2p436-451.html*

Gilchrist, S. & Zakrajsek, E. (2012), 'Credit Spreads and Business Cycle Fluctuations', *American Economic Review* **102**(4), 1692–1720.
URL: *https://ideas.repec.org/a/aea/aecrev/v102y2012i4p1692-1720.html*

Gürkaynak, R. S. & Wright, J. H. (2013), 'Identification and Inference Using Event Studies', *Manchester School* **81**, 48–65.
URL: *https://ideas.repec.org/a/bla/manchs/v81y2013ip48-65.html*

Hamilton, J. D. (1989), 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle', *Econometrica*

57(2), 357–384.

URL: *https://ideas.repec.org/a/ecm/emetrp/v57y1989i2p357-84.html*

Jarociński, M. & Karadi, P. (2020), 'Deconstructing monetary policy surprises—the role of information shocks', *American Economic Journal: Macroeconomics* **12**(2), 1–43.

URL: *https://www.aeaweb.org/articles?id=10.1257/mac.20180090*

Justiniano, A. & Primiceri, G. E. (2008), 'The Time-Varying Volatility of Macroeconomic Fluctuations', *American Economic Review* **98**(3), 604–641.

URL: *https://ideas.repec.org/a/aea/aecrev/v98y2008i3p604-41.html*

Kilian, L. & Luetkepohl, H. (2018), *Structural Vector Autoregressive Analysis*, number 9781107196575 *in* 'Cambridge Books', Cambridge University Press.

URL: *https://ideas.repec.org/b/cup/cbooks/9781107196575.html*

Koop, G., Leon-Gonzalez, R. & Strachan, R. W. (2010), 'Dynamic probabilities of restrictions in state space models: An application to the phillips curve', *Journal of Business & Economic Statistics* **28**(3), 370–379.

URL: *https://doi.org/10.1198/jbes.2009.07335*

Kuttner, K. N. (2001), 'Monetary policy surprises and interest rates: Evidence from the Fed funds futures market', *Journal of Monetary Economics* **47**(3), 523–544.

URL: *https://ideas.repec.org/a/eee/moneco/v47y2001i3p523-544.html*

Känzig, D. R. (2021), 'The macroeconomic effects of oil supply news: Evidence from opec announcements', *American Economic Review* **111**(4), 1092–1125.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20190964*

Mertens, K. & Ravn, M. O. (2013), 'The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States', *American Economic Review* **103**(4), 1212–1247.
URL: *https://ideas.repec.org/a/aea/aecrev/v103y2013i4p1212-47.html*

Miranda-Agrippino, S. & Ricco, G. (2020), 'The Transmission of Monetary Policy Shocks', *American Economic Journal: Macroeconomics (forthcoming)* .

Miranda-Agrippino, S. & Ricco, G. (2022), 'Identification with External Instruments in Structural VARs under Partial Invertibility', *Journal of Monetary Economics (forthcoming)* .

Mumtaz, H. & Petrova, K. (2021), Algorithms for proxy svar models with time-varying parameters, Technical report.

Nakamura, E. & Steinsson, J. (2018), 'High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*', *The Quarterly Journal of Economics* **133**(3), 1283–1330.
URL: *https://doi.org/10.1093/qje/qjy004*

Plagborg-Møller, M. & Wolf, C. K. (2021), 'Local Projections and VARs Estimate the Same Impulse Responses', *Econometrica* **89**(2), 955–980.

URL: *https://ideas.repec.org/a/wly/emetrp/v89y2021i2p955-980.html*

Primiceri, G. E. (2005), 'Time Varying Structural Vector Autoregressions and Monetary Policy', *Review of Economic Studies* **72**(3), 821–852.
URL: *https://ideas.repec.org/a/oup/restud/v72y2005i3p821-852.html*

Romer, C. D. & Romer, D. H. (2004), 'A New Measure of Monetary Shocks: Derivation and Implications', *American Economic Review* **94**(4), 1055–1084.
URL: *https://ideas.repec.org/a/aea/aecrev/v94y2004i4p1055-1084.html*

Rubio-Ramírez, J. F., Waggoner, D. F. & Zha, T. (2010), 'Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference', *Review of Economic Studies* **77**(2), 665–696.
URL: *https://ideas.repec.org/a/oup/restud/v77y2010i2p665-696.html*

Sims, C. A. & Zha, T. (2006), 'Were There Regime Switches in U.S. Monetary Policy?', *American Economic Review* **96**(1), 54–81.
URL: *https://ideas.repec.org/a/aea/aecrev/v96y2006i1p54-81.html*

Smets, F. & Wouters, R. (2007), 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach', *American Economic Review* **97**(3), 586–606.
URL: *https://ideas.repec.org/a/aea/aecrev/v97y2007i3p586-606.html*

Stock, J. H. & Watson, M. W. (2007), 'Why has u.s. inflation become

harder to forecast?', *Journal of Money, Credit and Banking* **39**(s1), 3–33.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4616.2007.00014.x*

Stock, J. H. & Watson, M. W. (2012), 'Disentangling the Channels of the 2007-09 Recession', *Brookings Papers on Economic Activity* **43**(1 (Spring), 81–156.
URL: *https://ideas.repec.org/a/bin/bpeajo/v43y2012i2012-01p81-156.html*

Stock, J. H. & Watson, M. W. (2018), 'Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments', *Economic Journal* **128**(610), 917–948.
URL: *https://ideas.repec.org/a/wly/econjl/v128y2018i610p917-948.html*

Wolf, C. K. (2020), 'Svar (mis)identification and the real effects of monetary policy shocks', *American Economic Journal: Macroeconomics* **12**(4), 1–32.
URL: *https://www.aeaweb.org/articles?id=10.1257/mac.20180328*

Online Appendix to "What Information Do Proxy-VARs Use? A Study of High Frequency Identification in Macroeconomics"

A Algorithms and Priors

A.1 Random Walk Specification

The first three steps of the algorithm follows exactly Algorithm 1 of Caldara & Herbst (2019), whose notation we largely borrow. The law of motion of β_t is given by

$$\beta_t = \beta_{t-1} + w_t, w_t \overset{iid}{\sim} N\left(0, \sigma_w^2\right).$$

In addition, we assume following priors:

$$\begin{aligned} p\left(\sigma_w^2\right) &\sim IG(\tau/2, \tau q/2). \\ p(\beta_0) &\sim N(b_0, V_0). \end{aligned}$$

The scale parameter q of the IG prior is crucial for controlling the time variation. We follow the procedure outlined in Amir-Ahmadi et al. (2020) to estimate this parameter.

Our VAR can be stated in companion form as

$$\mathbf{Y}_t = \Phi \mathbf{X}_t + \mathbf{U}_t \tag{A-1}$$

where $\mathbf{Y}_t$ stack current and lagged values of our vector of observables $\mathbf{y}_t$, $\mathbf{X}_t$ contains lags of $\mathbf{Y}_t$ as well as a vector of ones to capture the intercept, and $\mathbf{U}_t \sim_{iid} N(\mathbf{0}, \check{\Sigma})$.

Algorithm 1. For $i = 1, \ldots, N$. At iteration i

1. Draw $\check{\Sigma}, \Phi \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. For $\check{\Sigma}$ we use a mixture proposal distribution (suppressing dependence on parameters for notational convenience):

$$q\left(\check{\Sigma} \mid \check{\Sigma}^{i-1}\right) = \gamma p\left(\check{\Sigma} \mid \mathbf{Y}_{1:T}\right) + (1-\gamma)\mathcal{IW}\left(\check{\Sigma}; \check{\Sigma}^{i-1}, d\right)$$

where $p\left(\check{\Sigma} \mid \mathbf{Y}_{1:T}\right)$ is the known posterior distribution of $\check{\boldsymbol{\Sigma}}$ under $\mathbf{Y}_{1:T}$ and $\mathcal{IW}\left(\cdot; \check{\Sigma}^{i-1}, d\right)$ is an Inverse Wishart distribution with scaling matrix $\check{\Sigma}^{i-1}$ and d degrees of freedom. For Φ we use the known distribution $p\left(\Phi \mid \mathbf{Y}_{1:T}, \check{\Sigma}\right)$ as a proposal in an independence MH step:

   - Draw $\check{\Sigma}^*$ according to $q\left(\check{\Sigma} \mid \check{\Sigma}^{i-1}\right)$.
   - Draw Φ^* according to $p\left(\Phi \mid \mathbf{Y}_{1:T}, \check{\Sigma}^*\right)$.
   - With probability α, set $\Phi^i = \Phi^*$ and $\check{\Sigma}^i = \check{\Sigma}^*$, otherwise set $\Phi^i = \Phi^{i-1}$ and $\check{\Sigma}^i = \check{\Sigma}^{i-1}$, defined as

$$\alpha = \min\left\{\frac{p\left(\mathbf{M}_{1:T}, \mathbf{Y}_{1:T}, \Phi^*, \check{\Sigma}^*, \Omega^{i-1}, \beta^{i-1}, \sigma_\nu^{i-1}\right) p\left(\check{\Sigma}^*\right)}{p\left(\mathbf{M}_{1:T}, \mathbf{Y}_{1:T}, \Phi^{i-1}, \check{\Sigma}^{i-1}, \Omega^{i-1}, \beta^{i-1}, \sigma_\nu^{i-1}\right) p\left(\check{\Sigma}^{i-1}\right)} \frac{q\left(\check{\Sigma}^{i-1} \mid \check{\Sigma}^*\right)}{q\left(\check{\Sigma}^* \mid \check{\Sigma}^{i-1}\right)}, 1\right\}$$

2. Draw $\Omega \mid \mathbf{Y}_{1:T}, \mathbf{M}_t, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. Use an Independence Metropolis-Hastings sampler step using the Haar measure on the space of orthogonal matrices:

   - Draw Ω^* using Theorem 9 in Rubio-Ramírez et al. (2010).
   - With probability α, set $\Omega^i = \Omega^*$, otherwise $\Omega^i = \Omega^{i-1}$ is defined

as

$$\alpha = \min\left\{\frac{p\left(\mathbf{M}_{1:T} \mid \mathbf{Y}_{1:T}, \Phi^i, \check{\Sigma}^i, \mathbf{\Omega}^*, \beta^{i-1}, \sigma_\nu^{i-1}\right)}{p\left(\mathbf{M}_{1:T} \mid \mathbf{Y}_{1:T}, \Phi^i, \check{\Sigma}^i, \mathbf{\Omega}^{i-1}, \beta^{i-1}, \sigma_\nu^{i-1}\right)}, 1\right\}$$

3. Draw $\sigma_m^2 \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. Sample σ_m^2 from $IG\left(\bar{s}_1/2, \bar{s}_2/2\right)$, the known conditional posterior distribution associated with σ_m^2.

4. Draw $\beta_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. Conditional on all other parameters, the law of motion forms a linear Gaussian state space system. This step can be carried out using the simulation smoother detailed in Carter & Kohn (1994) or Primiceri (2005).

5. Draw $\sigma_w^2 \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. Sample σ_w^2 from $IG\left(\bar{w}_1/2, \bar{w}_2/2\right)$, the known conditional posterior distribution associated with σ_w^2.

6. Draw $q \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, \sigma_w^{i-1}, q^{i-1}$. The scale parameter is sampled with a MH step outlined in Amir-Ahmadi et al. (2020).

A.2 Markov switching

In the case of Markov switching in β_t, we assume that β_t follows a two state Markov process with

$$\begin{aligned} \beta_t &= \beta_{s_t} \\ \Pr(s_t &= i|s_{t-1} = j) = p_{ij} \\ i,j &\in \{1,2\}. \end{aligned}$$

We assume following priors

$$
\begin{aligned}
p(\beta_{s_t=1}) &\sim N(b_1, V_1). \\
p(\beta_{s_t=2}) &\sim N(b_2, V_2). \\
p_{11} &\sim beta\left(a_{11}, b_{11}\right). \\
p_{22} &\sim beta\left(a_{22}, b_{22}\right).
\end{aligned}
$$

Algorithm 2. For $i = 1, \ldots, N$. At iteration i. The first 3 steps of the algorithm are the same as Algorithm 1.

4. Draw $\beta_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. Sample β_t from $N(\bar{b}_1, \bar{V}_1)$ if $s^{i-1} = 1$ and from $N(\bar{b}_2, \bar{V}_2)$ if $s^{i-1} = 2$. Both are known conditional normal distributions.

5. Draw $p_{11}, p_{22} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. Sample p_{11} from $beta(\bar{a}_{11}, \bar{b}_{11})$ and p_{22} from $beta(\bar{a}_{22}, \bar{b}_{22})$. Both are known conditional beta distributions (see Frühwirth-Schnatter (2006), page 330).

6. Draw $s_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{\Sigma}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_m^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. Sample $s_{1:T}$ using the Multi-Move sampler outlined in Frühwirth-Schnatter (2006), algorithm 11.5.

A.3 More on Priors

We use the benchmark Minnesota prior setting for the VAR with a very loose overall tightness parameter equal to 10.[1] The diagonal elements of the location matrix of the inverse Wishart prior are fixed to estimates based on pre-sample data (lag length). We use fairly uninformative prior

[1] For the exact definition of is parameter see Giannone et al. (2015).

for the residual variance $\sigma_m^2 \sim IG(s_1/2, (s_1 s_2^2/2)$. For the estimation of the time variation q, we adopt the half-Cauchy prior with scale parameter θ. The prior hyperparameters are summarized in the following table:

Table A-1: TVP Benchmark Prior Hyperparameters

s_1	s_2	b_0	V_0	τ	θ
2	0.2	0	1	2	0.01

In case of Markov Switching, the Minnesota prior specification remains the same, the other prior hyperparameters are summarized in the following table:

Table A-2: Markov Switching Benchmark Prior Hyperparameters

s_1	s_2	b_1	V_1	b_2	V_2	a_{11}	b_{11}	a_{22}	b_{22}
2	0.2	0	1	0	1	6	1	6	1

All posterior results except for the Monte Carlo experiments are based on $500,000$ draws from the MCMC sampler.

B Details on Monte Carlo Exercises

All of our Monte Carlo setups consist of two regimes. Our goal is to match the variance of the instrument for a given regime across specifications. We assume that in the benchmark the monetary policy shocks are $N(0, \sigma_e^2)$ and $\beta = 1$ in one regime and equal 0 in the other. Furthermore, we will assume that in the benchmark the variance of the measurement error v_t is a fixed fraction κ of the variance of the monetary policy shock.[2] Note that in contrast to our estimated model (where e_t^{MP} is assumed to

[2]Compared to the main text, we use non-unit variance shocks in the data-generating process, whereas the shocks entering the estimated model in the main text (monetary shock e_t^{MP} and measurement error u_t^M) are unit variance shocks.

have unit variance and all scaling is captured in the impact matrix Σ), the true monetary policy shock does not have unit variance. This affects the scale of the estimated β_t coefficients and needs to be taken into account when comparing to the true values stated here (we give more details when discussing the estimated paths of β_t).

In our Monte Carlo exercise, we simulate 100 samples of length $T = 250$ each. The variables we use in Monte Carlo exercise are the nominal interest rate, output, inflation, and the monetary policy shock from an estimated Smets-Wouters model. The VAR contains simulated nominal interest rate, output and inflation and the lag length is set to 4. In each of the Monte Carlo repetitions (in total 100), posterior results are based on 50,000 MCMC draws. The prior specification is exactly the same as in the empirical estimation.

B.1 Benchmark

The measurement equation and the variance in the two regimes are:

$$m_t = e_t + v_t, Var(m_t) = (1+\kappa)\sigma_e^2 \tag{A-2}$$

$$m_t = v_t, Var(m_t) = \kappa\sigma_e^2 \tag{A-3}$$

We set $\sigma_e^2 = 0.2290^2$ equal to the DGP value and $\kappa = 0.25$. For $T = 5, ..., 140$ and $T = 150, ..., 230$, $\beta = 0$. Otherwise, $\beta = 1$. These values are chosen to be comparable to the Gertler-Karadi instrument.

B.2 Changing Volatility in the measurement error

We now assume that the measurement error v_t has a variance that switches between regimes with values $\sigma_{v,1}^2$ and $\sigma_{v,2}^2$. The measurement equations are given by:

$$m_t = \overline{\beta} e_t + v_t, Var(m_t) = \overline{\beta}^2 \sigma_e^2 + \sigma_{v,1}^2 \tag{A-4}$$

$$m_t = \overline{\beta} e_t + v_t, Var(m_t) = \overline{\beta}^2 \sigma_e^2 + \sigma_{v,2}^2 \tag{A-5}$$

We now need to solve the following two equations:

$$\overline{\beta}^2 \sigma_e^2 + \sigma_{v,1}^2 = (1+\kappa)\sigma_e^2 \tag{A-6}$$

$$\overline{\beta}^2 \sigma_e^2 + \sigma_{v,2} = \kappa\sigma_e^2 \tag{A-7}$$

We actually have three unknowns and two equations here. Since all variances have to be positive, we have additional constraints though. We set $\overline{\beta} = \sqrt{\kappa}$ and $\sigma_{v,2}^2 = 0$. This implies $\sigma_{v,1}^2 = \sigma_e^2$.

We set $\sigma_e^2 = 0.2290^2$ (equal to the DGP value) and $\kappa = 0.25$. For $T = 5, ..., 140$ and $T = 150, ..., 230$, $\sigma_{v,2}^2 = 0$. Otherwise, $\sigma_{v,1}^2 = \sigma_e^2$.

B.3 Changing Volatility in e_t

We now assume that the variance in the monetary policy shocks changes, with variances $\sigma_{e,1}^2$ and $\sigma_{e,2}^2$. We also allow the measurement error variance $\tilde{\sigma}_v^2$ and the coefficient $\tilde{\beta}$ to be different than in the other specifications (they are fixed across regimes though). The equations in this MC are given by

$$m_t = \tilde{\beta} e_t + v_t, Var(m_t) = \tilde{\beta}^2 \sigma_{e,1}^2 + \tilde{\sigma}_v^2 \tag{A-8}$$

$$m_t = \tilde{\beta} e_t + v_t, Var(m_t) = \tilde{\beta}^2 \sigma_{e,2}^2 + \tilde{\sigma}_v^2 \tag{A-9}$$

The equations we need to solve are:

$$\tilde{\beta}^2\sigma_{e,1}^2 + \tilde{\sigma}_v^2 = (1+\kappa)\sigma_e^2 \quad \text{(A-10)}$$

$$\tilde{\beta}^2\sigma_{e,2}^2 + \tilde{\sigma}_v^2 = \kappa\sigma_e^2 \quad \text{(A-11)}$$

We impose $\tilde{\beta} = 1$ and $\tilde{\sigma}_v^2 = 0$, which implies $\sigma_{e,2}^2 = \kappa\sigma_e^2$ and $\sigma_{e,2}^2 = (1+\kappa)\sigma_e^2$.

We set $\sigma_e^2 = 0.2290^2$ and $\kappa = 0.25$. For $T = 5, ..., 140$ and $T = 150, ..., 230$, $\sigma_{e,2}^2 = \kappa\sigma_e^2$. Otherwise, $\sigma_{e,2}^2 = (1+\kappa)\sigma_e^2$.

B.3.1 Alternative Measurement Equation

Since the previous exercise is somewhat cumbersome to interpret, we also carried out an alternative where the data-generating process for all variables except the instrument is the same as before. For the instrument, we now assume that

$$m_t = \overline{e}_t + v_t, \quad \text{(A-12)}$$

where $\overline{e}_t$ is the normalized monetary policy shock that has unit variance each period. We set $v_t \sim N(0, 0.25)$. This exercise has the disadvantage that the path of the instrument's volatility is not the same as in the previous exercise. The advantage is that the instrument equation is independent of changes in the monetary policy shock's volatility. Furthermore, this exercise is certainly not as realistic as the others because the instrument is linked to the normalized true shock. Figure A-1 plots the impulse responses under this alternative specification - results are basically indistinguishable from the original exercise in the main text, as can be seen when comparing Figure A-1 with Figure 14. The posterior median paths of β_t are now flat (Figure A-2).

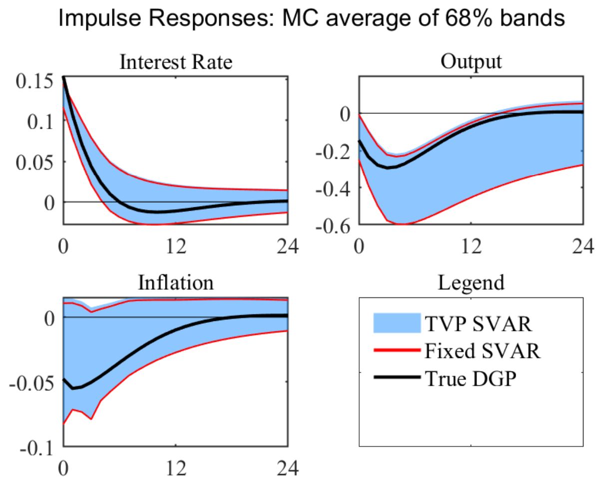


Figure A-1: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in the monetary policy shock, alternative instrument.

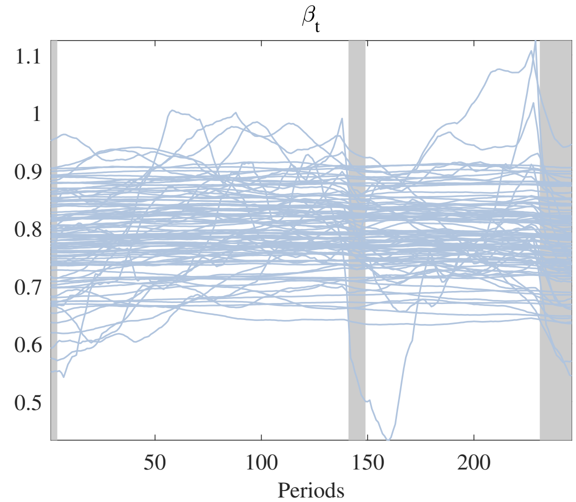


Figure A-2: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of the monetary policy shock, alternative instrument.

C Data Sources

For the US economy, we follow Gertler & Karadi (2015) and obtained industrial production (INDPRO), consumer price index (CPIAUCSL) and 1-year treasury rate (GS1) from FRED (`https://fred.stlouisfed.org/`). The data for the excess bond premium is obtained from Board of Governors (`https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/files/ebp_csv.csv`). The instrument of Gertler & Karadi (2015) is obtained from the replication file of the paper (`https://www.openicpsr.org/openicpsr/project/114082/version/V1/view`). The instrument of Miranda-Agrippino & Ricco (2022) is obtained from the personal website of Silvia Miranda-Agrippino (`http://silviamirandaagrippino.com/s/Instruments_web-x8wr.xlsx`). For the UK economy, we use the replication data and instrument of Cesa-Bianchi et al. (2020) from `https://github.com/ambropo/MP_HighFrequencyUK/`.

D Additional Figures

Here we show the full set of impulse responses for various specifications in the main text.

D.1 Markov Switching

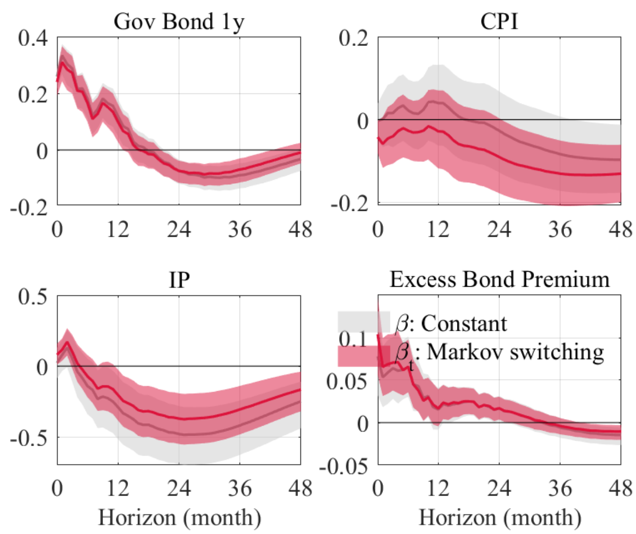


Figure A-3: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - Markov-switching specification.

D.2 Shutting Down Periods Where the Instrument is Informative/Not Informative

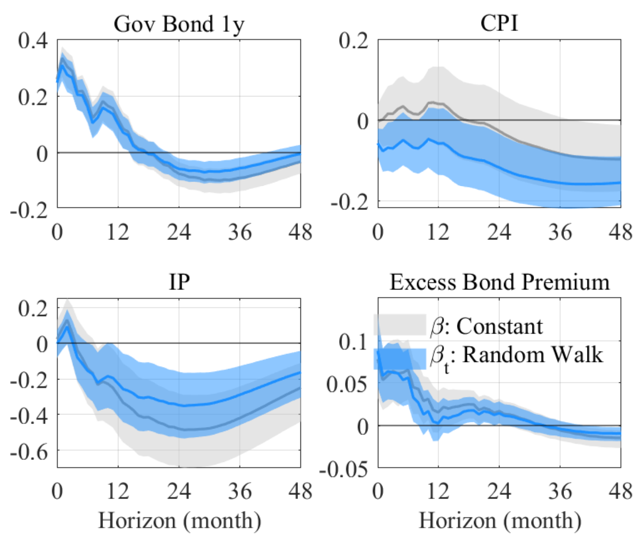


Figure A-4: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - $\tilde{m}_t$.

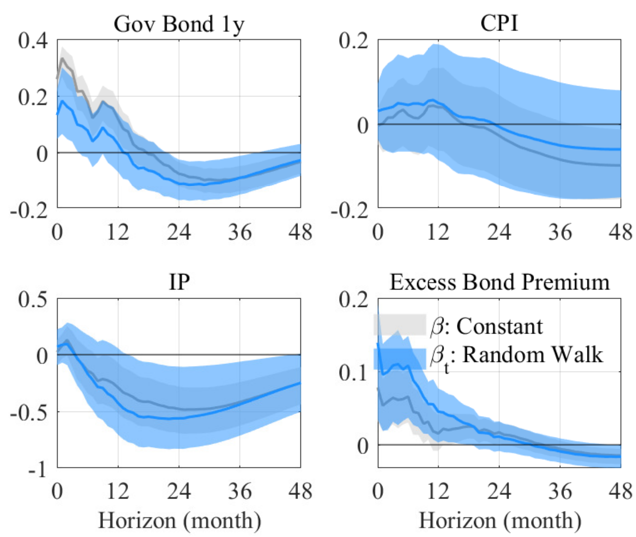


Figure A-5: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - $\overline{m}_t$.

D.3 Alternative Instruments

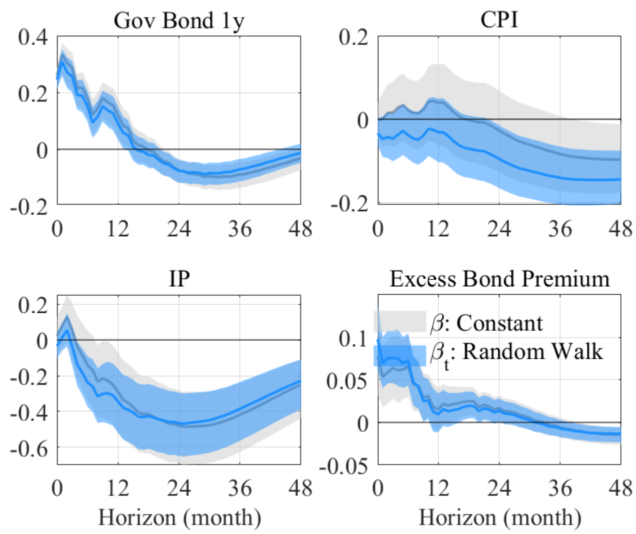


Figure A-6: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - modified instrument.

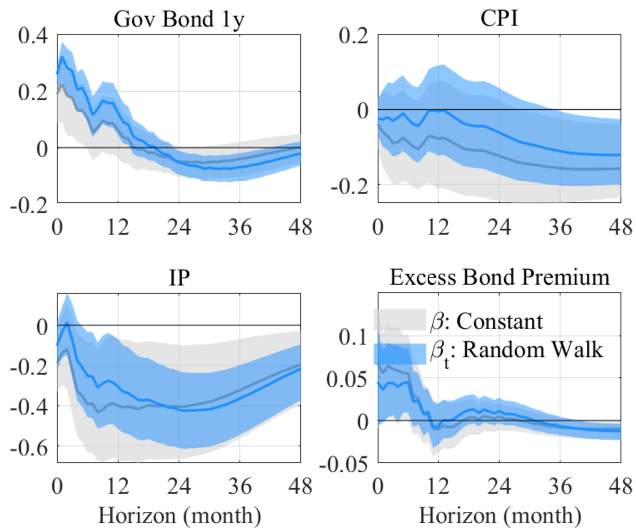


Figure A-7: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - Miranda-Agrippino & Ricco (2020) instrument.

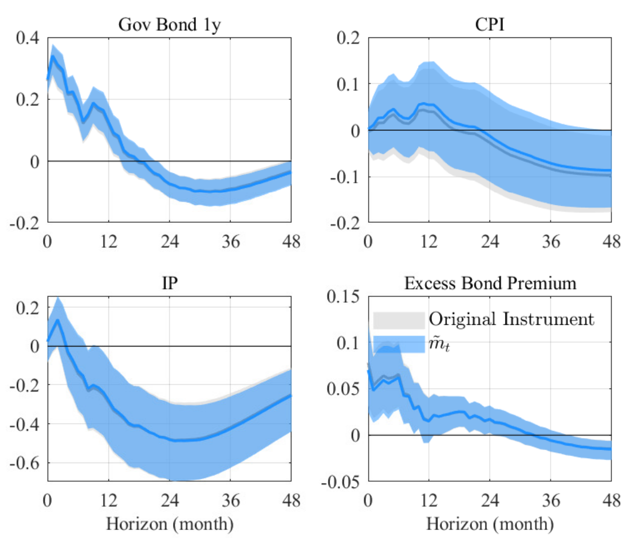


Figure A-8: Impulse responses (median and 68 percent posterior bands) based on fixed coefficient VAR for benchnark instrument and $\tilde{m}_t$

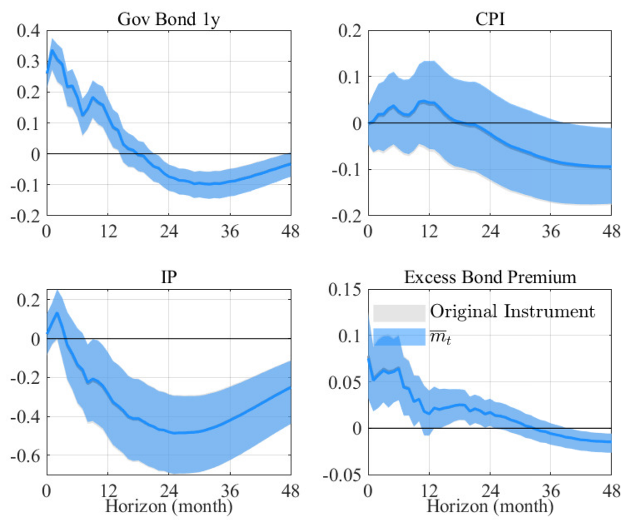


Figure A-9: Impulse responses (median and 68 percent posterior bands) based on fixed coefficient VAR for benchnark instrument and $\overline{m}_t$

E Evidence from the UK

Finally, we present evidence for high-frequency based identification of monetary policy shocks in the United Kingdom. We use both the instrument and the VAR specification (i.e. the choice of variables entering $\mathbf{y}_t$) of Cesa-Bianchi et al. (2020).

Figure A-10 shows impulse responses for all UK variables in the VAR. In contrast to the US, we find little difference between fixed coefficient-based responses and random walk-based responses. A potential reason can be seen in Figure A-11: the sample for the UK (both for the VAR variables and the instrument) is much shorter, and within that shorter time-span there are more periods where the instrument is informative, making the fixed coefficient estimation generally more informative.

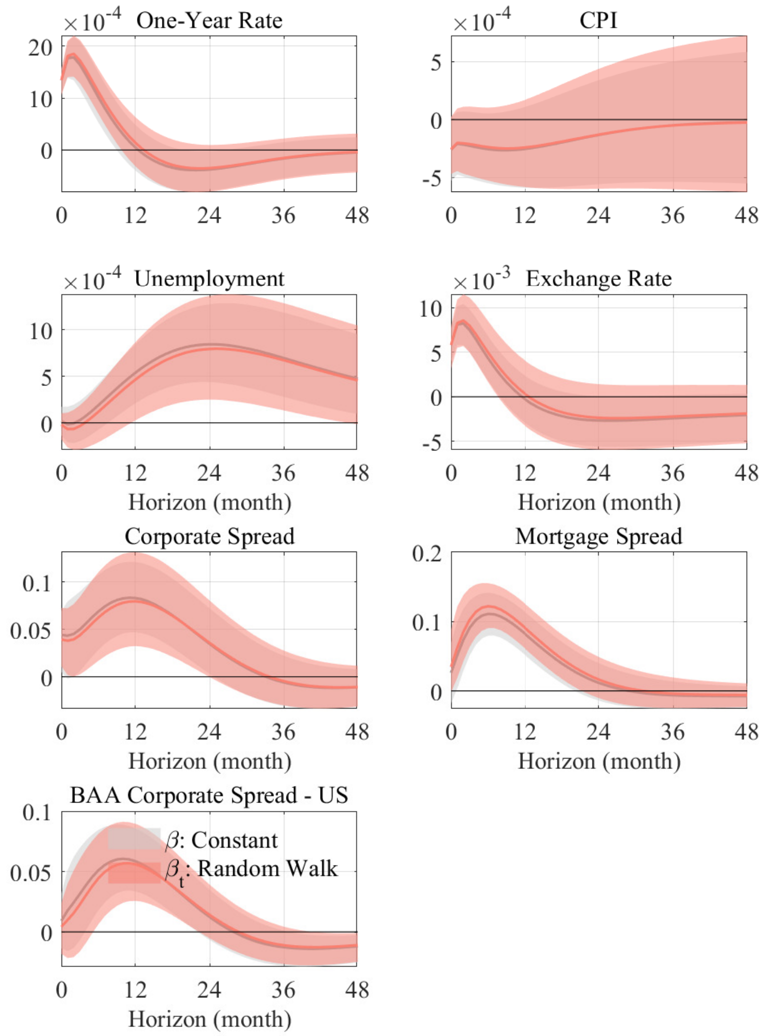


Figure A-10: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock, UK.

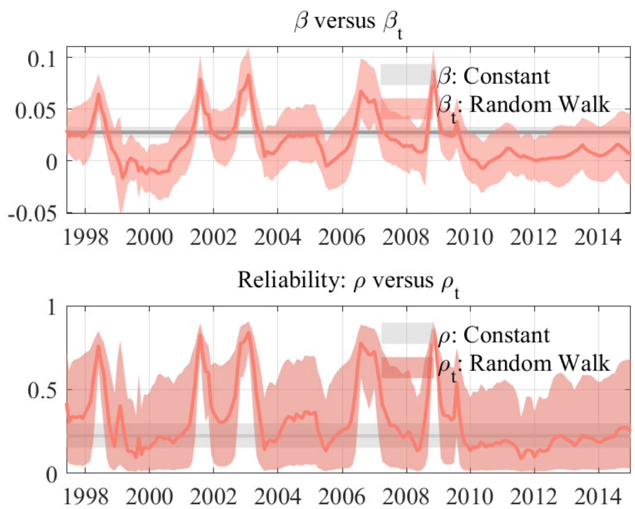


Figure A-11: Posterior of β_t and ρ_t, UK (median and 68 percent posterior bands).