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

Pooyan Amir-Ahmadi*
Amazon
and
Christian Matthes
Department of Economics, Indiana University
and
Mu-Chun Wang
Deutsche Bundesbank

March 27, 2025

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, insofar as they explicitly model the relationship between instrument and structural shock, 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. Therefore, we build a narrative for instrument-based identification. For the instrument in Gertler & Karadi (2015), the effect on the (log) price level is almost 50 percent larger than the standard specification would suggest.

Keywords: High-Frequency Identification, Instruments, Monetary Policy

*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, Domenico Giannone, Ed Herbst, Toru Kitagawa, Karel Mertens, Franck Portier, Mikkel Plagborg-Møller, Giorgio Primiceri, Josefine Quast, Juan Rubio-Ramírez, Frank Schorfheide, Eric Swanson, Andrea Tambalotti, 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, the Barcelona Summer Forum, the NBER Summer Institute, and the 2022 ESOBE conference. 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 past decade (Stock & Watson 2012, Mertens & Ravn 2013, Gertler & Karadi 2015). These external instruments are interpreted as imperfect measurements of unobserved structural shocks. An instrument-based approach mitigates the issues that can arise when using standard sign restrictions to identify monetary policy shocks, as highlighted by Wolf (2020).

A key assumption of studies that use this approach while conducting Bayesian inference 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), an instrument popularized by Gertler & Karadi (2015) that we also use. The figure shows that there are periods where the dynamics and volatility 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 vector autoregressions (VARs) that explicitly capture this time variation, using the Bayesian approach for VARs with instruments (commonly called proxy VARs).[1] We show that the pattern observed in Figure 1 can be explained parsimoniously by moving to a nonlinear measurement equation linking the instrument and the structural shock of interest, while maintaining a linear and Gaussian structure for the VAR itself to be comparable to the bulk of the literature.[2] We introduce this nonlinearity

[1]Time-varying identification strength is also a feature of the non-parametric framework in Rambachan & Shephard (2021).

[2]In theory, time-varying volatility of monetary policy shocks could also lead to the pattern described

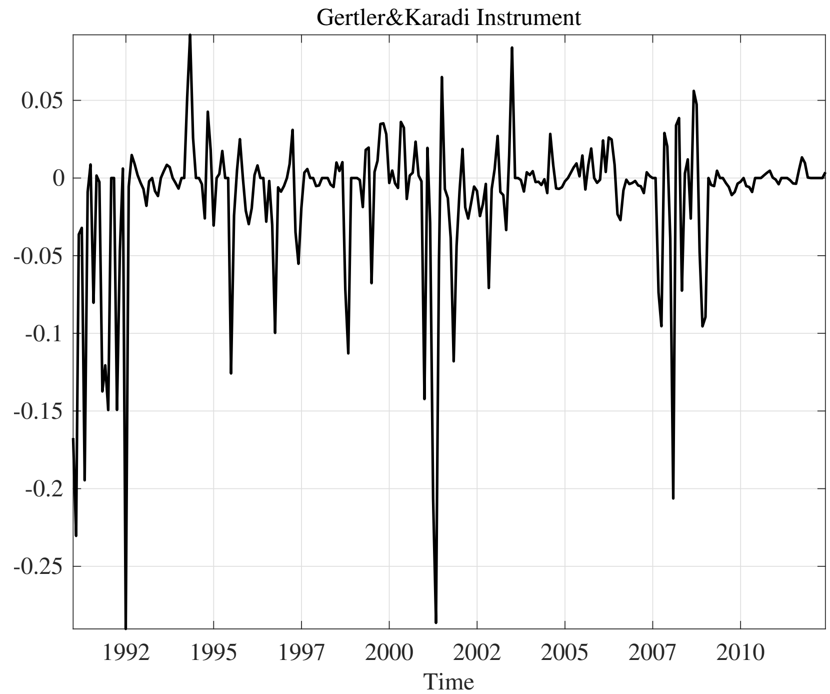


Figure 1: Surprise in 3-month-ahead Fed Funds futures (Gertler & Karadi 2015).

by allowing either changes in the volatility of the noise term or changes in the parameter multiplying the unobserved shock of interest.[3] These two assumptions translate changes in the volatility of the instrument into time variation in identification strength in opposite ways - volatile realizations of the instrument are deemed informative when we use the time-varying parameter approach, but are considered uninformative (i.e. a source of weak identification) when we estimate a model with stochastic volatility in the noise term. We show that for our specific application, assuming changes in the volatility of the noise term leads to results that most economists will find questionable. First, a standard VAR estimated via ordinary least squares (OLS) delivers forecast errors that are highly correlated with the instrument during periods where volatility is high, but not otherwise. Second, the implied impulse responses in the case of stochastic volatility in the noise term are estimated with a large amount of uncertainty and have the wrong sign for the response of

here. We show in a Monte Carlo exercise in Appendix B that such a data-generating process would lead to estimates of time-varying parameters that are qualitatively very different to those obtained using U.S. data.

[3]Our approach comes at negligible additional computational cost relative to the previous literature.

prices. Therefore, we interpret changes in the volatility of the instrument through the lens of a model with time variation in the parameter linking the instrument and the monetary policy shock instead.

We discuss what the implications of various types of misspecification[4] are on estimates obtained using our preferred specification. These implications are at odds with our estimates using US data, which we confirm using Monte Carlo experiments.

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 comes 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 certain periods only. 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 monetary policy shocks 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. Using the same instrument as Gertler & Karadi (2015) in our application yields, for example, effects on prices that are almost 50 percent larger after four

[4]One source of misspecification we study is stochastic volatility in the monetary policy shock itself.

years.

Error bands for impulse responses are generally *not* wider than their fixed-coefficient counterparts. Even in applications where our approach yields impulse responses similar 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.

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 that are 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 Stock (2008) and Mertens & Ravn (2013), who introduced the proxy VAR framework.[5] 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).

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 4.4 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

[5]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.

and the Gertler & Karadi (2015) instrument.

Our work is related to (Abadie et al. 2023), who show that exploiting heterogeneity in instrument strength in the first stage of a standard instrumental variable setting can substantially improve the mean squared error of instrumental variables estimators. We present evidence that exploiting heterogeneity across time can lead to improved estimators of impulse responses in applied macro settings.

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 $\mathbf{e}_t$.[6] To estimate said response, we use a structural vector autoregression (SVAR) in equation (1):

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

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

The well-known identification problem in Gaussian structural VARs (Canova 2011, Baumeister & Hamilton 2015, Kilian & Luetkepohl 2018) implies that we need additional information to identify the column of the response matrix $\mathbf{B}$, 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

[6] We use boldface for vectors and matrices.

in an SVAR analysis.[7] Since our ultimate goal is to study possible changes in the relationship between the observable instrument m_t[8] and the unobserved monetary policy shock, we explicitly model the relationship between the instrument and the structural shock of interest. Within the standard homoskedastic VAR framework in the literature (Mertens & Ravn 2013, Caldara & Herbst 2019), there are two diametrically opposite assumptions that can generate the patterns observed in Figure 1.[9] We introduce these cases in Equations (2) and (3):

$$m_t = \beta_t e_t^{MP} + \sigma_v v_t. \tag{2}$$

$$m_t = \beta e_t^{MP} + \sigma_{v,t} v_t. \tag{3}$$

v_t is distributed independently and identically over time as $N(0,1)$ in either equation. Equation (2) accounts for the observed patterns in the instrument by allowing the coefficient on the monetary shock to change, while Equation (3) allows for changes in the noise variance. These alternatives lead to very different interpretations of the observed data, as we show below. Measurement equations of the kind we use are standard in the literature dating back to Mertens & Ravn (2013), but the common assumption is that $\beta_t = \beta \forall t$ and $\sigma_{v,t} = \sigma_v \forall t$, especially when Bayesian inference is conducted.[10] The measurement equations can be generalized to include lags for observed variables on the right-hand side (Arias et al. 2021), and we show in Appendix G that such an extension does not change the conclusions in our empirical application.

The key identification assumptions are twofold. First, we assume that v_t is independent

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

[8]Instead of one scalar instrument, we could use multiple instruments. In that case we would, for example, need to make a decision about possible correlation in the error terms of the measurement equations.

[9]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 time-invariant.

[10]One exception is Mertens & Ravn (2013), where the authors allow for censoring of the instrument (so that the entire right-hand side is multiplied by an indicator function), which is conceptually distinct from the type of time variation we study.

of all other shocks in our model, both the vector of structural shocks $\mathbf{e}_t$ and any shocks determining the evolution of β_t or $\sigma_{v,t}$, generally denoted as w_t.[11] Second, the instrument is informative for the monetary policy shock, meaning that at least for some periods, $\beta_t \neq 0$. Our assumptions then imply

$$E\left[e_{j,t} m_t\right] = 0 \quad \text{for } j = 2, \dots, n, \text{(exogeneity)} \tag{4}$$

$$E\left[v_t \mathbf{e}_t\right] = 0 \quad \text{and} \quad E\left[v_t w_t\right] = 0, \tag{5}$$

$$\beta_t \neq 0 \quad \text{for some } t \text{ in Equation (2) or } \beta \neq 0 \quad \text{in Equation (3), (relevance)} \tag{6}$$

where $e_{j,t}$ denotes the jth element of $\mathbf{e}_t$.

Our Bayesian estimation approach is still valid even if $\beta_t = 0 \quad \forall t$ in Equation (2) or $\beta = 0$ in Equation (3) in the sense that our approach will asymptotically identify that the instrument is not relevant ($\beta_t = 0 \forall t$), although, naturally, in that case the instrument will not aid identification of the shock of interest. Our approach automatically approximates the posterior distribution of all time-varying parameters 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.[12] We borrow the approach of directly estimating the parameters of this measurement equation from Caldara & Herbst (2019). Unlike in that paper, we allow for changes in parameters that govern the systematic relationship between instruments and shocks.[13]

A useful summary statistic for assessing the strength of the instrument in different

[11] As is common in the literature on time-varying VAR (Cogley & Sargent 2002, Primiceri 2005), we also assume that the structural shocks $\mathbf{e}_t$ are independent of innovations to parameters w_t.

[12] 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).

[13] Following Caldara & Herbst (2019), we normalize the relevant column of $\mathbf{B}$ 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 also allows us to center the prior for β_t or β at zero while still maintaining a standard interpretation of the estimated monetary policy shock.

periods is a time-varying version of the common reliability (or relevance) statistic ρ_t:

$$\rho_t \equiv \begin{cases} \frac{\beta_t^2}{\beta_t^2+\sigma_v^2} & \text{if Equation (2) holds} \\ \frac{\beta^2}{\beta^2+\sigma_{v,t}^2} & \text{if Equation (3) holds} \end{cases} \tag{7}$$

This statistic represents the squared correlation between the instrument and the structural shock at time t and measures time-varying identification strength.

Equations (2) and (3) provide different theories for changes in volatility of the instrument m_t. These different theories have opposite effects on ρ_t: If an increase in volatility of m_t is driven by an increase in β_t this will lead to an increase in ρ_t, whereas an increase in $\sigma_{v,t}$ will lead to a decrease in ρ_t. The first specification therefore interprets periods where the instrument is volatile as more informative for the identification of the effects of monetary policy shocks. The stochastic volatility specification instead discounts these periods because it attributes these fluctuations in the instrument to noise and instead identifies the effects of monetary policy shocks using periods with low instrument volatility.

In Section 4.1, we first compare three different specifications for the time variation in parameters:[14]

$$\text{Constant: } \beta_t = \beta, \sigma_{v,t} = \sigma_v \tag{8}$$

$$\text{Time variation in Equation (2): } \beta_t = \beta_{t-1} + \sigma_\beta w_t \tag{9}$$

$$\text{Time variation in Equation (3): } \log(\sigma_{v,t}^2) = \log(\sigma_{v,t-1}^2) + \sigma_u w_t \tag{10}$$

In Equations (9) and (10), we assume that $w_t \sim_{iid} N(0,1)$. The first specification is a constant parameter specification reminiscent of Caldara & Herbst (2019) as a benchmark

[14]We focus on these diametrically opposite cases here. One could entertain stochastic volatility $\sigma_{v,t}$ and time-varying β_t jointly, but in that case the prior in the relative variability on β_t and $\log(\sigma_{v,t})$ would be crucial due to weak identification (for a discussion in the context of time-varying VAR models, see Amir-Ahmadi et al. (2020). More importantly, our choice of diametrically opposite cases helps us interpret our findings.

(equation (8)), the second specification is a Gaussian random walk specification for β_t 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 the third specification is a standard specification for stochastic volatility, where log volatility follows a random walk (Kim et al. 1998, Cogley & Sargent 2002, Primiceri 2005). We choose these specifications not only because they are common in the literature, but, more importantly, because estimates obtained using these specifications can capture many patterns of time variation even if the random walk specifications are misspecified (see, for example, Amir-Ahmadi et al. (2020)).

In finite samples, a misspecified fixed coefficient approach will use information from all time periods, even those where the instrument is not informative.

Standard frequentist approach will *asymptotically* be valid regardless as long as the instrument is relevant (i.e. there are enough periods where the instrument is informative).However, in finite samples, using irrelevant information (i.e. assuming Equations (8) when the true instrument relevance is time-varying) will lead to inefficient estimates. If there are only relatively few periods where the instrument is relevant, then our approach naturally takes into account that there is not much information about the structural shock in the instrument, thus automatically taking into account possibly weak identification. Note that this is different from having an inefficient estimator - weak identification, if present, is a feature of the data-generating process if the model is correctly specified. In Appendix F, we study the performance of our approach as we vary the number of informative periods in a Monte Carlo setting to highlight that even with very few informative periods, our approach outperforms the fixed coefficient alternative.

To approximate the posterior of our model, which consists of equations (1), (2), and one of the equations (8), (9), or (10), we modify the Metropolis-within-Gibbs sampling framework of Caldara & Herbst (2019) (the specification with equation (8) is exactly their specification). An 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. In Section 3, we show how our approach is related to, but distinct from, identification based on heteroskedasticity (Rigobon 2003).

3 Identification With Time-Varying Instrument Relevance

In this section, we use environments where the parameters follow a two-state discrete Markov chain to derive analytical results that can help us understand where our identification comes from and when it breaks down. In order to analyze the relationship to identification via changes in volatility, we first stack our original VAR and the measurement equation for the instrument m_t (which we assume to be scalar). We first assume that only β_t can vary over time before turning to a more general environment. $\mathbf{A}(L)$ is a polynomial in the lag operator. $\mathbf{e}_{2,t}$ is a vector that collects all structural shocks except for the monetary shock e_t^{MP}. The associated impact effects on $\mathbf{y}_t$ are collected in $\mathbf{B}^{MP}$ and $\mathbf{B}^2$.

$$z_t = \begin{bmatrix} m_t \\ \mathbf{y}_t \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{c} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{A}(L) \end{bmatrix} \begin{bmatrix} m_{t-1} \\ \mathbf{y}_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \sigma_v & \beta_t & 0 \\ 0 & \mathbf{B}^{MP} & \mathbf{B}^2 \end{bmatrix}}_{\mathbf{B}_t} \underbrace{\begin{bmatrix} v_t \\ e_t^{MP} \\ \mathbf{e}_{2,t} \end{bmatrix}}_{\mathbf{u}_t}$$

Note that in contrast to standard identification via heteroskedasticity (Rigobon 2003), the volatility of the shocks $E(\mathbf{u}_t\mathbf{u}_t')$ is time-invariant, but the impact matrix $\mathbf{B}_t$ varies because β_t can change over time.[15] We show now that if β_t follows a two regime Markov-

[15] We assume throughout, as before, that $E(\mathbf{u}_t\mathbf{u}_t') = I$.

switching process, we can identify the effects of $\varepsilon_{1,t}$. The Markov-switching structure is only assumed for simplicity. More general laws of motion for β_t would yield the same insights. We call the two possible values for B_t B_0 and B_1. These matrices only differ in their values for β, which is equal to either β_0 or β_1.

To obtain our result, we start by writing out $\mathbf{B}_1\mathbf{B}_1'$:

$$\begin{bmatrix} \sigma_v & \beta_1 & 0 \\ 0 & \mathbf{B}^{MP} & \mathbf{B}^2 \end{bmatrix} \times \begin{bmatrix} \sigma_v & 0 \\ \beta_1 & \mathbf{B}^{MP\prime} \\ 0 & \mathbf{B}^{2\prime} \end{bmatrix} = \begin{bmatrix} \sigma_v^2 + \beta_1^2 & \beta_1 \mathbf{B}^{MP\prime} \\ \beta_1 \mathbf{B}^{MP} & \mathbf{B}^{MP}\mathbf{B}^{MP\prime} + \mathbf{B}^2\mathbf{B}^{2\prime} \end{bmatrix}$$

We also get a similar expression for $\mathbf{B}_0\mathbf{B}_0'$, with β_1 replaced by β_0. The first row of those two quadratic forms allows us to identify $\beta_1^2 - \beta_0^2$ (by taking the difference of the first elements in the first row) and $\frac{\beta_1}{\beta_0}$ (by taking the ratio of any element of $\beta_1\mathbf{B}^{MP\prime}$ and the corresponding element of $\beta_0\mathbf{B}^{MP\prime}$). With those two pieces of information, we can uniquely pin down β_1 and β_0 up to one sign normalization. The other elements of the first row of $\mathbf{B}_1\mathbf{B}_1'$ or $\mathbf{B}_0\mathbf{B}_0'$ except for the first element pin down $\mathbf{B}^{MP}$, which identifies the effects of e_t^{MP}.

Because we restrict only β_t (a scalar random variable) to vary across regimes, each regime gives us $(n+1)$ new pieces of information from the regime-specific variance of the forecast errors.[16] Furthermore, our model implies that the lower nn block of the forecast error covariance matrix is regime-independent. Thus, our model of time-varying β_t coefficients is overidentified.[17]

[16] The overall forecast error variance $\mathbf{B}_t\mathbf{B}_t'$ has $\frac{(n+1)(n+2)}{2}$ unique elements, but $\frac{(n+1)n}{2}$ elements of that matrix do not depend on β_t.

[17] In contrast, when we consider general time variation in $\mathbf{B}_t$, an issue we come back to in the next section, we have n^2+1 free parameters in $\mathbf{B}_t$ in each regime and still only $\frac{(n+1)(n+2)}{2}$ free elements in the forecast error covariance amtrix, meaning that such a model would not be identified for $n > 3$.

3.1 Identification Breakdown When General Time Variation is Present

We will now consider an environment with two regimes where all relevant parameters are allowed to vary, including both the volatility of the measurement error and the volatility of the monetary policy shock. We will study two macroeconomic outcomes stacked in the vector y_t. For simplicity, assume that all data are iid (none of the arguments above relied on the data being persistent). Furthermore, we assume that $\mathbf{y}_t$ is two-dimensional. Again, nothing hinges on this assumption, but it hopefully makes the exposition clearer. We can then again stack the model for y_t with the measurement equation for the instrument m_t to get:

$$\begin{bmatrix} m_t \\ y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \sigma_v & \beta_t & 0 \\ 0 & B^{MP}_{y1,t} & B_{y1,t} \\ 0 & B^{MP}_{y2,t} & B_{y2,t} \end{bmatrix} \begin{bmatrix} \tilde{\sigma}_{v,t} & 0 & 0 \\ 0 & \tilde{\sigma}^{MP}_t & 0 \\ 0 & 0 & \tilde{\sigma}_t \end{bmatrix} \begin{bmatrix} v_t \\ e^{MP}_t \\ e_{2,t} \end{bmatrix} \tag{11}$$

We can identify the variance of the one-step ahead forecast error in each of the two regimes, which in total gives us 12 independent pieces of information that can be used to help with identification. Let us further assume that the time varying volatilities $\tilde{\sigma}$ are normalized to be equal to 1 in one regime. This leaves us with 14 parameters to be estimated, so there is no chance that identification can be achieved. To be clear, we do not recommend to estimate such a general system without further identification restrictions.[18] However, counting the number of restrictions only gives us a *necessary* condition for identification, and even if we incorporate additional restrictions the situation is less than clear: the mapping from parameters to elements of the covariance matrix of the one-step ahead forecast error is not linear. In fact, the mapping is represented by multivariate higher order polyno-

[18] In a factor-augmented VAR context, Liao et al. (2023) explore the role of time-varying instrument relevance and allow all VAR parameters to vary over time. As our example shows, priors play a crucial role in such a model even asymptotically.

mial equations,[19] which generally have multiple solutions that would need to characterized numerically on a case-by-case basis.

In Appendix E, we show how our approach compares empirically to an approach that assumes stochastic volatility in the monetary policy shock itself, but shuts down all time variation in the relationship between the instrument and the structural shock.

4 Effects of Monetary Policy Shocks Identified Through High-Frequency Variation in Asset Prices

Section 4.1 studies the effects of allowing for time-varying reliability using the specifications outlined above. Crucially, we will make our case for modeling time variation in β_t *for this specific application* and analyze further variations of that specification. Section 4.2 uses a Markov-switching approach to model time variation in β_t for reasons that will become obvious once we study the posterior of β_t in our benchmark random-walk specification. In Section 4.3, we highlight that indeed only a few periods are informative for the effects of monetary policy shocks by estimating our model with an instrument that is set to zero except for the most informative periods. Finally, Section 4.4 estimates the effects of monetary policy shocks with the Miranda-Agrippino & Ricco (2020) instrument, the Bauer & Swanson (2022) instrument, and an alternative version of the Gertler & Karadi (2015) instrument.

[19] In general the equations will feature at least some third order terms even if we treat the squared values of the $\tilde{\sigma}$ volatilities directly as objects of interest, thus reducing the degree of the polynomial that needs to be solved.

4.1 The Effects of Time-Varying Reliability

We first contrast the constant parameter specification with the random walk specification for β_t and the stochastic volatility in noise specification. Our application uses US data:[20] $\mathbf{y}_t$ 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 such as the EBP in our VAR is crucial in order to get the effects of monetary policy right. The sample for $\mathbf{y}_t$ runs from July 1979 to June 2012 to match the sample used by Gertler & Karadi (2015). We also follow Gertler & Karadi (2015) in our choice of instrument m_t and use the surprise in the three-month-ahead Fed Funds futures around FOMC meetings (the 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 preference for the 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 that 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 models, the instrument is volatile and periods where it is not.[21] 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

[20] In Appendix H, we apply our preferred specification to UK data, using the high-frequency instrument of Cesa-Bianchi et al. (2020).

[21] More specifically, we use the periods that our model with time variation in β_t identifies as informative as described in Section 4.3. Using the stochastic volatility model to identify these volatile periods leads to very similar results.

time-invariant relationship, so this correlation should be constant across subsamples).

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 volatile, 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.[22] Furthermore, the absolute decrease in correlation is meaningful (a decrease in correlation of 0.34 for inflation and EBP, for example). Not only is the correlation between the instruments and forecast errors stronger when the instrument is volatile, in periods where the instrument is volatile, the *signs* of the correlation between the instrument and the forecast errors in the VAR are more in line with correlations implied by standard New Keynesian theories: a contractionary shock raises interest rates, but lowers prices. The sign of the correlation for IP might seem unusual, but we will see a small initial positive impulse response of IP to a monetary policy shock (confirming the sign of the correlation) that quickly turns negative in our preferred specification below.

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

	i	CPI	IP	EBP
High volatility in m_t	0.48	-0.34	0.10	0.46
Low volatility in m_t	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 that 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 often be somewhat sensitive to the choice

[22]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.

of prior for the innovation standard deviations σ_β and σ_{vol} 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 only have *either* time-varying parameters or stochastic volatility in our models, which helps sharpen inference. Nevertheless, to ensure 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 priors for σ_β and σ_{vol}. Details on the priors can be found in Appendix A.

We first analyze the case of time variation in β. Figure 2 shows the posterior path of β_t and ρ_t. We plot the corresponding elements of the fixed coefficient version in gray. We show the posterior median as well as 68 percent equal-tail posterior bands.[23] It is striking that there are few short periods of high instrument relevance when allowing for time variation in β.[24] Three periods stand out, which we now discuss in turn. The first period is the first half of the 1990s. It is useful to point out that the large posterior value of β_t at the beginning of the sample for the instrument 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).

The second period that our model highlights is in 2001, driven by two intermeeting rate changes in January and April of 2001.[25] The third period of high instrument relevance

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

[24] 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.

[25] The rate change around September 11th 2001 is not part of our instrument series as most financial markets were closed until the rate change on September 17th 2001, making it impossible to compute the changes in Fed Funds futures needed to construct the instrument.

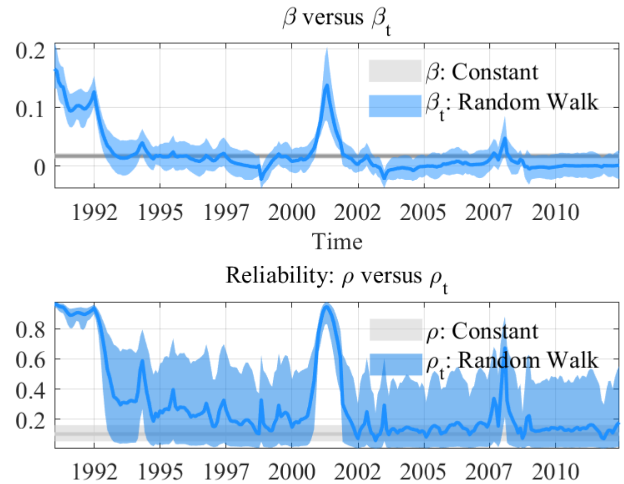


Figure 2: Posterior of β_t and ρ_t (median and 68 percent posterior bands).

is the Great Recession around 2008. Thus, our framework helps us to understand what information is contained in the instruments. We next examine whether this time variation in instrument relevance matters for impulse responses.

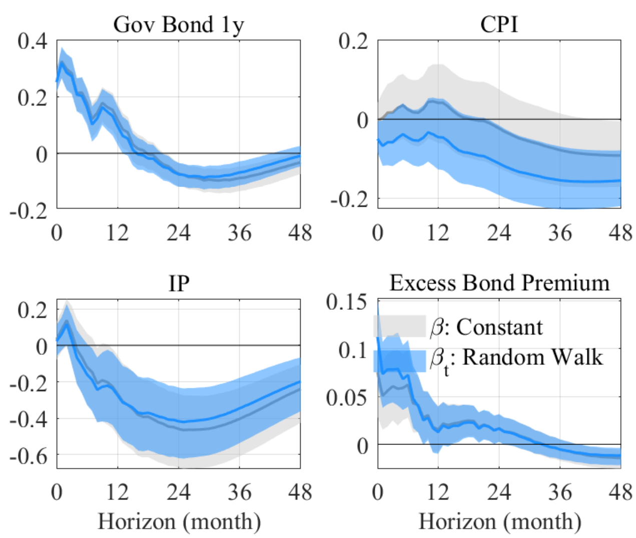


Figure 3: Impulse responses (median and 68 percent posterior bands) to a monetary policy shock, time variation in β.

Figure 3 shows the impulse to a monetary policy shock that raises interest rates by 25 basis points under the fixed coefficient (gray) and random walk (blue) specifications.[26]

[26]We maintain the normalization of a monetary shock to raise interest by 25 percent unless noted otherwise.

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 can hence lead to substantially different impulse responses. As mentioned before, in our example this does not come at a cost in terms of the width of the error bands.

We now contrast these results with the case of stochastic volatility in the noise term. Figure 4 shows that the same volatile periods that were previously identified as high β_t periods are now identified as periods with large noise volatility and basically zero reliability ($\rho_t = 0$). In light of the correlation structure with VAR forecast errors that we analyzed above, it is not surprising then that the resulting impulse responses, as displayed in Figure 5, show no meaningful response in prices or Industrial Production with the posterior median response for prices being positive for the first three years. In light of these findings, we will focus on time variation in β for the rest of the paper.

Modeling changes in β_t means that volatile periods of the instrument are interpreted as informative events; with stochastic volatility in the noise term, they are be interpreted as noise. In most of our applications, it turns out that the instruments are generally not very informative (low ρ_t) except possibly for clearly delineated short periods of high instrument volatility. Thus, using stochastic volatility in noise implies a prior that, in these specific applications, puts very little faith in the instruments. This stands in contrast to standard priors in the proxy VAR literature (Arias et al. 2021, Caldara & Herbst 2019) that imply that the instruments are indeed useful/reliable. Our preferred assumption of time-varying

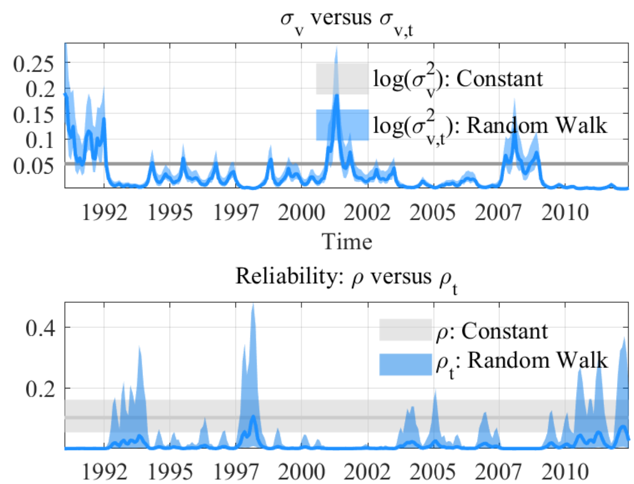


Figure 4: Posterior of $\sigma_{v,t}$ and ρ_t (median and 68 percent posterior bands).

Figure 5: Impulse responses (median and 68 percent posterior bands) to a monetary policy shock, time variation in σ_v.

parameters instead of time variation in the volatility of the noise term σ_v can thus also be seen as a context-specific prior choice that implies at least *some* instrument reliability, in line with the previous literature.

4.2 A Markov-Switching Alternative

Recalling Figure 2, one possible criticism that could be raised is that the estimated path for β_t might be better characterized by a Markov-switching model (Hamilton 1989, Sims & Zha 2006). 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. Nevertheless, we next estimate a two-state Markov-switching specification and show that it yields very similar results. 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 (9) and (12), respectively.

$$\text{Markov Switching in Equation (2): } \beta_t = \beta_{s_t}, Pr(s_t = i|s_{t-1} = j) = p_{ij}. \tag{12}$$

Figure 6 shows the impulse response of log CPI to a 25 basis point 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. This impulse response is very similar to the random walk specification.

Figure 7 illustrates 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 meaningful difference in impulse responses, as discussed above. The choice of 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 the default choice because of its flexibility.

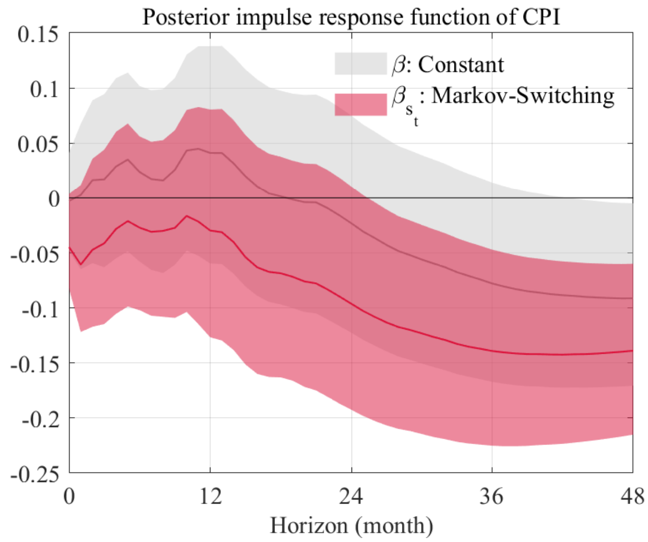


Figure 6: Impulse response of log CPI for Markov-switching specification (median and 68 percent posterior bands).

4.3 Shutting Down Periods Where the Instrument Is Informative/Uninformative

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)[27] 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, thus exacerbating weak identification problems. The threshold probability of 0.5 only selects the early 1990s and

[27]The approach in Koop et al. (2010)computes Bayesian model probabilities of both the unrestricted model and the restricted model with $\beta_t = 0$.

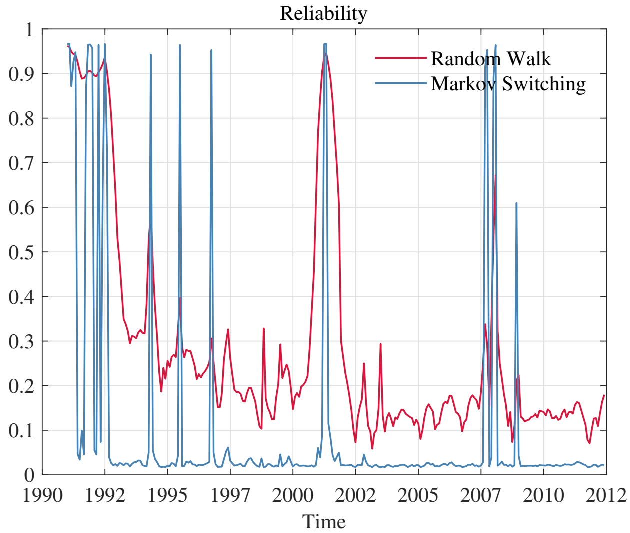


Figure 7: Posterior reliability for Markov-switching and random walk specifications (posterior median).

2001 as informative periods.

Figure 8 shows the results when we use $\tilde{m}_t$ as our instrument. For 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.[28] 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 spikes in the instrument relevance we document, the procedure in this section would have led to a meaningful loss of information.

Figure 9 shows 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, implying that identification is weak throughout the sample. This is true even though we keep 90 percent of the observations from the original sample because there is little information contained in those observations.

[28]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.

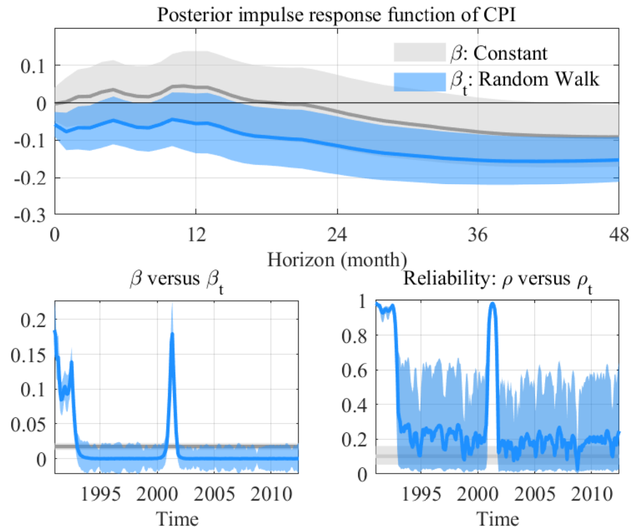


Figure 8: Results for $\tilde{m}_t$ (median and 68 percent posterior bands). The fixed coefficient VAR is based on the original m_t instrument.

In the Appendix (Figures A-16 and A-17), 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 to allow for time variation so that the instrument is not used for identification when it is not informative.

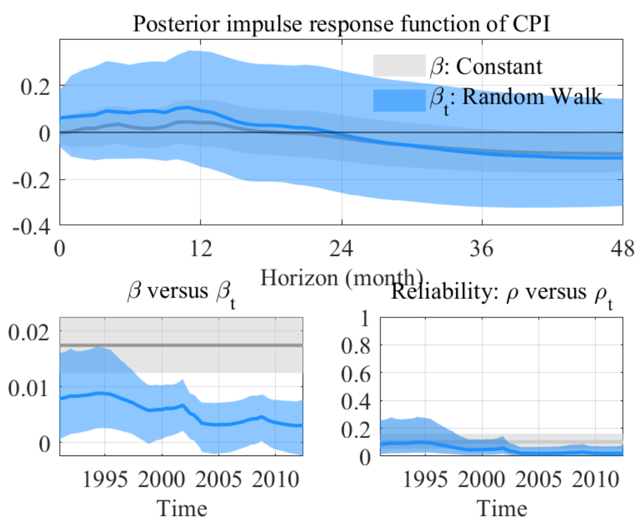


Figure 9: Results for $\overline{m}_t$ (median and 68 percent posterior bands). The fixed coefficient VAR is based on the original m_t instrument.

4.4 Alternative Instruments

Our various observation equations that link m_t and the unobserved monetary shock all imply that m_t is iid, borrowing from Caldara & Herbst (2019). Other papers in the literature (Arias et al. 2021, Plagborg-Møller & Wolf 2021) 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 $\mathbf{y}_t$ in the VAR. The key results are summarized in Figure 10 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 G, including estimates of the reliability and β_t when we use a more general measurement equation and directly include lags of the VAR variables in the measurement equation - results are basically unchanged there as well).

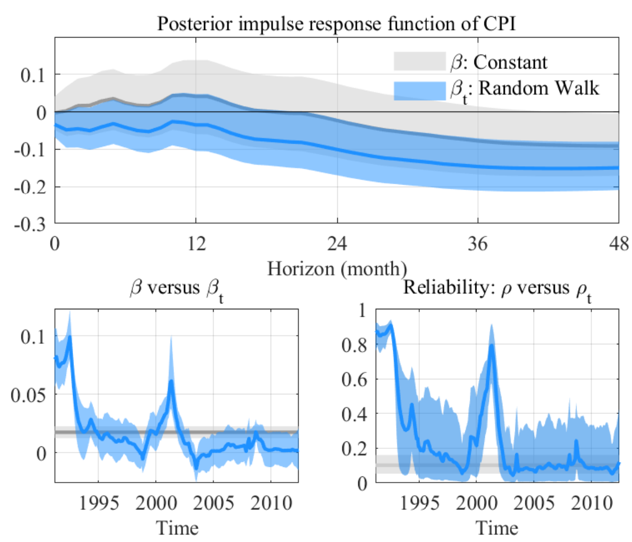


Figure 10: Results for the case of the modified instrument (median and 68 percent posterior bands).

In Appendix G, we also show results from our approach when using the instruments by Miranda-Agrippino & Ricco (2020), Bauer & Swanson (2022).

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, which helps 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 removing periods that are not informative will generally not help, unless a researcher is willing to model time variation in instrument relevance.

Although we focus on monetary policy shocks in our application, 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

Abadie, A., Gu, J. & Shen, S. (2023), 'Instrumental variable estimation with first-stage heterogeneity', *Journal of Econometrics* .

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.

Antolín-Díaz, J. & Rubio-Ramírez, J. F. (2018), 'Narrative sign restrictions for svars', *American Economic Review* **108**(10), 2802–29.

Arias, J. E., Rubio-Ramírez, J. F. & Waggoner, D. F. (2021), 'Inference in Bayesian Proxy-SVARs', *Journal of Econometrics* **225**(1), 88–106.

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.

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.

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.

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).

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.

Drautzburg, T. (2020), 'A narrative approach to a fiscal DSGE model', *Quantitative Economics* **11**(2), 801–837.

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.

Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, Springer New York.

Gertler, M. & Karadi, P. (2015), 'Monetary Policy Surprises, Credit Costs, and Economic Activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.

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.

Gilchrist, S. & Zakrajsek, E. (2012), 'Credit Spreads and Business Cycle Fluctuations', *American Economic Review* **102**(4), 1692–1720.

Gürkaynak, R. S. & Wright, J. H. (2013), 'Identification and Inference Using Event Studies', *Manchester School* **81**, 48–65.

Hamilton, J. D. (1989), 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle', *Econometrica* **57**(2), 357–384.

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

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

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

Kim, S., Shephard, N. & Chib, S. (1998), 'Stochastic volatility: Likelihood inference and comparison with arch models', *The Review of Economic Studies* **65**(3), 361–393.

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.

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.

Känzig, D. R. (2021), 'The macroeconomic effects of oil supply news: Evidence from OPEC announcements', *American Economic Review* **111**(4), 1092–1125.

Liao, W., Ma, J. & Zhang, C. (2023), 'Identifying exchange rate effects and spillovers of us monetary policy shocks in the presence of time-varying instrument relevance', *Journal of Applied Econometrics* **38**(7), 989–1006.

URL: *https://onlinelibrary-wiley-com.proxyiub.uits.iu.edu/doi/abs/10.1002/jae.2998*

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.

Miranda-Agrippino, S. & Ricco, G. (2020), 'The Transmission of Monetary Policy Shocks', *Anerican 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.

Petrova, K. (2022), 'Asymptotically valid bayesian inference in the presence of distributional misspecification in var models', *Journal of Econometrics* **230**(1), 154–182.

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

Primiceri, G. E. (2005), 'Time Varying Structural Vector Autoregressions and Monetary Policy', *Review of Economic Studies* **72**(3), 821–852.

Rambachan, A. & Shephard, N. (2021), 'When do common time series estimands have nonparametric causal meaning?'.

Rigobon, R. (2003), 'Identification through heteroskedasticity', *The Review of Economics and Statistics* **85**(4), 777–792.

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

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.

Sims, C. A. & Zha, T. (2006), 'Were There Regime Switches in U.S. Monetary Policy?', *American Economic Review* **96**(1), 54–81.

Smets, F. & Wouters, R. (2007), 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach', *American Economic Review* **97**(3), 586–606.

Stock, J. H. (2008), 'Recent developments in structural VAR modeling', Lecture slides, NBER Summer Institute Methods Lecture "What's New in Econometrics: Time Series".

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.

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.

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.

Wolf, C. K. (2020), 'SVAR (mis)identification and the real effects of monetary policy shocks', *American Economic Journal: Macroeconomics* **12**(4), 1–32.

Online Appendix for

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

A Algorithms and Priors

A.1 Time Varying Parameter

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 $\mathbf{U}_t \sim_{iid} N(\mathbf{0}, \check{B})$.

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

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

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

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

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

$\check{B}^i = \check{B}^{i-1}$, defined as

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

2. Draw $\Omega \mid \mathbf{Y}_{1:T}, \mathbf{M}_t, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, \sigma_w^{i-1}, q^{i-1}$. 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

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

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

4. Draw $\beta_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, \sigma_w^{i-1}, q^{i-1}$. 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 $\sigma_w^2 \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, \sigma_w^{i-1}, q^{i-1}$. 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 $q \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, \sigma_w^{i-1}, q^{i-1}$. 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 the 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 $\beta_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. 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 $p_{11}, p_{22} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. 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 $s_{1:T} \mid \mathbf{Y}_{1:T}, \mathbf{M}_{1:T}, \check{B}, \Phi, \Omega^{i-1}, \beta_{1:T}^{i-1}, \sigma_v^{i-1}, p_{11}^{i-1}, p_{22}^{i-1}, s_{1:T}^{i-1}$. Sample $s_{1:T}$ using the Multi-Move sampler outlined in Frühwirth-Schnatter (2006), algorithm 11.5.

A.3 Stochastic Volatility

We assume that the variance of the measurement error innovation follows a random walk

$$\log\left(\sigma_{v,t}^{2}\right)=\log\left(\sigma_{v,t-1}^{2}\right)+w_{t}, w_{t}\overset{iid}{\sim}N\left(0,\sigma_{u}^{2}\right).$$

In addition, we assume the following priors:

$$\begin{aligned} p\left(\sigma_{u}^{2}\right) &\sim IG(\tau/2,\tau r/2). \\ p(\log(\sigma_{v,0}^{2})) &\sim N(v_{0},W_{0}). \end{aligned}$$

Similar to the case of time varying β, the scale parameter r of the IG prior is crucial for controlling the stochastic volatility. We follow the procedure outlined in Amir-Ahmadi et al. (2020) to estimate this parameter.

In practice, we use the same Gibbs steps to draw $\beta_{1:T}$ but set q the hyperparameter controlling the time variation to a very small number, i.e. 10^{-4}. The step 3 of algorithm 1 is then replaced by

3a Draw $\sigma_{v,1:T}^{2}\mid\mathbf{Y}_{1:T},\mathbf{M}_{1:T},\check{B},\Phi,\Omega^{i-1},\beta_{1:T}^{i-1},\sigma_{v,1:T}^{i-1},\sigma_{w}^{i-1},q^{i-1},\sigma_{u}^{i-1},r^{i-1}$. The sampler drawing $\log\left(\sigma_{v,t}^{2}\right)$ is based on Kim et al. (1998) who approximate the distribution of $\log\left(\sigma_{v,t}^{2}\right)$ by mixtures of normal distributions.

3b Draw $\sigma_{u}\mid\mathbf{Y}_{1:T},\mathbf{M}_{1:T},\check{B},\Phi,\Omega^{i-1},\beta_{1:T}^{i-1},\sigma_{v,1:T}^{i-1},\sigma_{w}^{i-1},q^{i-1},\sigma_{u}^{i-1},r^{i-1}$. Sample σ_{u}^{i-1} from $IG\left(\bar{u}_{1}/2,\bar{u}_{2}/2\right)$, the known conditional posterior distribution associated with σ_{u}^{2}.

3c Draw $r\mid\mathbf{Y}_{1:T},\mathbf{M}_{1:T},\check{B},\Phi,\Omega^{i-1},\beta_{1:T}^{i-1},\sigma_{v}^{i-1},\sigma_{w}^{i-1},q^{i-1},\sigma_{u}^{i-1},r^{i-1}$. The scale parameter is sampled with a MH step outlined in Amir-Ahmadi et al. (2020).

A.4 More on Priors

We use a 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 OLS estimates of an AR(1) model based on 12 observations of pre-sample data. We use a fairly uninformative prior for the residual variance $\sigma_v^2 \sim IG(s_1/2, (s_1 s_2^2/2)$. For the estimation of the parameter 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

In case of stochastic volatility in the measurement error, we use a very similar prior setting as in the case of time varying parameter. In particular, we also estimate the hyperparameter r which controls the degree of stochastic volatility base on Amir-Ahmadi et al. (2020). We assume again a half-Cauchy prior with scale parameter θ. As mentioned in the description of the sampler, we set q the hyperparameter controlling the time variation to 10^{-4}. The other prior hyperparameters are summarized in the following table:

Table A-3: Stochastic Volatility Prior Hyperparameters

v_0	W_0	b_0	V_0	τ	θ
$\log(0.2^2)$	1	0	1	2	0.01

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

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

B What Happens When Our Model Is Misspecified?

We next turn to discussing possible sources and consequences of misspecification: What happens if the true data-generating process features changes in volatility in either the noise term v_t or the monetary policy shock, but we estimate a model with time variation in β? We first present some general deliberations for the case of stochastic volatility in monetary policy shocks before studying these issues in further detail using Monte Carlo experiments. One possible source of confusion in the discussion of stochastic volatility in structural shocks is that we can always rewrite the measurement equation linking the instrument and the monetary policy shock as featuring time-varying parameters when the true data-generating process features changes in the volatility of monetary policy shocks. This does *not* mean that the two specifications are equivalent - a homoskedastic VAR specification will be misspecified when the monetary policy shock features stochastic volatility. This leads to different implications for estimated parameter paths of β_t, allowing us to argue that stochastic volatility in monetary policy shocks is not what drives our results. To see this, assume that the true data generating process for the instrument is

$$m_t = \overline{\beta} e_t^{MP} + \sigma_v v_t \tag{A-2}$$

where the true monetary policy shock e_t^{MP} features changes in volatility σ_t^{MP} and $\overline{\beta}$ is a fixed parameter. We now derive an equivalent representation of the instrument equation that features changes in β_t:

$$m_t = \beta_t \overline{e}_t + \sigma_v v_t \tag{A-3}$$

where $\overline{e}_t$ is the standardized monetary policy shock that has fixed variance 1 (as in our empirical model assumes). β_t in this case equals $\overline{\beta}\sigma_t^{MP}$.[2] With knowledge of all parameters in the stochastic volatility case (including the entire time path of shock volatility) we can, thus, derive this equivalent representation of the measurement equation. When there are volatility changes in the data-generating process, we now show that our approach will not recover the unit variance shock $\overline{e}_t$. Therefore, our approach will not recover the implied $\beta_t = \overline{\beta}\sigma_t^{MP}$. The estimated path of β_t will, instead, be more muted.

To see this, we can rewrite the VAR equation (1) in slightly more compact notation as

$$\mathbf{y}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{e}_t \tag{A-4}$$

where $\mathbf{x}_{t-1}$ collects the relevant lags of $\mathbf{y}_t$ as well as the constant term. Let us assume that the true data-generating process also takes this VAR form with parameter matrices $\mathbf{A}^*$ and $\mathbf{B}^*$, but with one element of the true structural shocks $\mathbf{e}_t^*$ (namely the true monetary policy shock e_t^{MP}) having nonconstant variance σ_t^{MP}. We now study how this affects our estimates of the structural shocks when the estimated model does not feature changes in volatility. Suppose that we have parameter values $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ (we can think about a draw from the posterior or a point estimate). Then the implied estimate of the vector of structural shocks $\tilde{\mathbf{e}}_t$ is

$$\tilde{\mathbf{e}}_t = \tilde{\mathbf{B}}^{-1}([\mathbf{A}^* - \tilde{\mathbf{A}}]\mathbf{x_{t-1}} + \mathbf{B}^*\mathbf{e}_t^*) \tag{A-5}$$

It is known that misspecification of the distribution of the structural shocks in linear VARs does not asymptotically bias the estimates of the dynamics of the VAR (Petrova 2022), so asymptotically under standard conditions all elements of $\mathbf{A}^* - \tilde{\mathbf{A}}$ will be zero except possibly for those related to the constant term. In large samples, variation in $\tilde{\mathbf{e}}_t$ is, thus,

[2] As a side note, in order to get a value of β_t near 0 in this specification (as in our estimated results for US data) we would need the volatility $\overline{\beta}\sigma_t^{MP}$ to go to zero.

driven by variation in the true shocks $\mathbf{e}_t^*$, and the estimated shocks $\tilde{\mathbf{e}}_t$ will inherit stochastic volatility from the true shocks $\mathbf{e}_t^*$.

Let us now consider an, admittedly, unlikely special case, namely $\mathbf{B}^* - \tilde{\mathbf{B}} = \mathbf{0}$. In that case, we recover the true structural shocks with stochastic volatility and as a result in large enough samples our estimates of β_t will converge to a neighborhood of $\overline{\beta}$ because we have the true structural shock as a right hand-side variable in the step of the Gibbs Sampler that estimates the path of β_t (this is the aforementioned muted response relative to the implied value of β_t that we derived at the beginning of the section). Why? Asymptotically, we recover the true time-invariant $\overline{\beta}$ because we are running a regression of m_t on the true shocks, which, as we just discussed, we can recover in this special case. Therefore, stochastic volatility in the true structural shock of interest has different implications for parameter estimates in our model compared to a data-generating process with changes in β_t that do not come from stochastic volatility, even if these assumptions seem similar at first sight.

Importantly, with estimation error or bias in $\tilde{\mathbf{B}}$ and, due to the sample size, possibly estimation error in $\tilde{\mathbf{A}}$, this argument will not hold *exactly*, but, as we show in our Monte Carlo study below, holds *approximately* for a realistic data-generating process and sample size.[3]

We also show by means of these 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. However, those specifications for the data-generating process lead to estimated paths of β_t that are inconsistent with our findings based on US data, thus providing evidence for our modeling assumptions. While we

[3] As far as there is meaningful time variation in the posterior of β_t in this scenario, our estimation will put more weight on periods with high estimated β_t to infer the effects of monetary policy shocks — periods that in this scenario are actually periods with high volatility of the monetary policy shock—, while discounting periods with low estimated β_t — periods with low volatility of the monetary policy shock.

could, in theory, use marginal likelihoods to compare those different model specifications, we instead choose to use Monte Carlo-based evidence because marginal likelihoods can be substantially influenced by tail behavior of priors even if that tail behavior is inconsequential for most objects of interests, such as posterior error bands for impulse responses. This dependence on priors becomes particularly pronounced when priors are relatively loose, prompting us to instead focus on Monte Carlo experiments.

B.1 Monte Carlo Studies

We follow Wolf (2020) and use the Smets & Wouters (2007) model as a laboratory.[4] 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 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_v.[5]

The first Monte Carlo experiment uses time variation in β_t that broadly matches the number of informative periods from our empirical estimates. 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,[6] we set $\beta_t = 1$ in the data-generating process; otherwise it is zero (hence the instrument is just noise). Figure A-1 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.

[4] 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. Details on the exact calibration of the data-generating processes can be found in Appendix C.

[5] 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.

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

Our results confirm those of the population analysis in Wolf (2020):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.

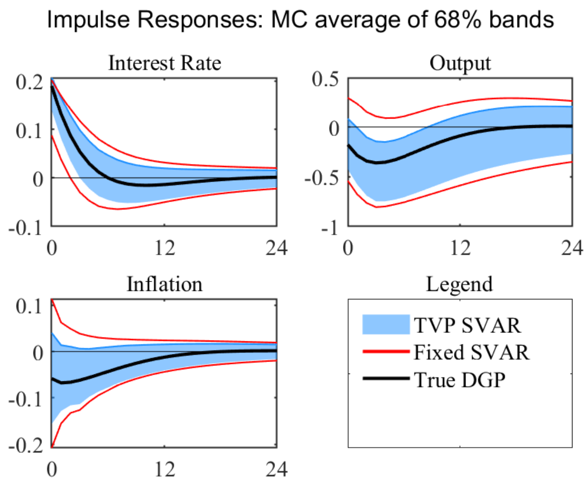


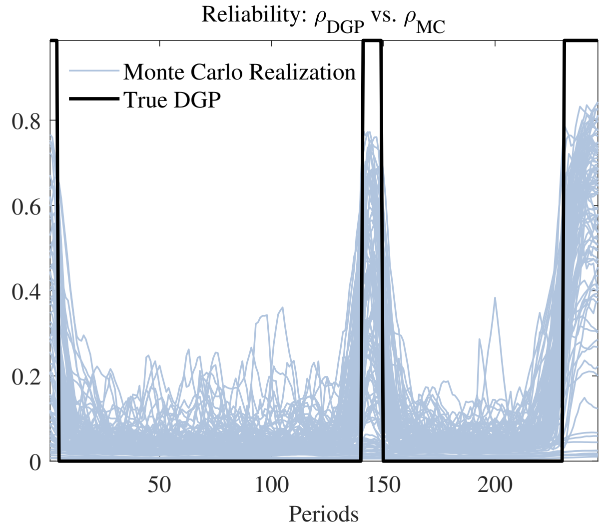
Figure A-1: Impulse responses for the data-generating process with time-varying β_t and the Monte Carlo replications, first Monte Carlo experiment.

Figure A-2a 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 bars). Our approach captures the changes in instrument reliability well.

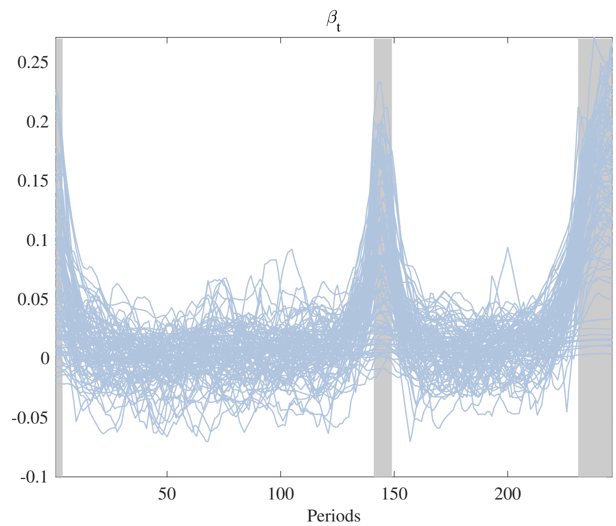
It is instructive to directly study the posterior median paths of β_t. As we discuss in more detail in Appendix C, we must rescale the true β_t values described there by the standard deviation of the monetary shock to make them comparable to our estimation results since our estimated model assumes monetary policy shocks with unit variance, whereas the monetary shock in the Smets-Wouters model has non-unit variance.[7] 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

[7]All scale effects in our estimated model are captured by $\mathbf{B}$.

shock with non-unit variance were 0 and 1).



(a) Reliability (posterior medians).



(b) Posterior medians of β_t.

Figure A-2: Reliability and posterior paths of β_t, first Monte Carlo experiment.

The posterior paths shown in Figure A-2b 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 in which our specification of the measurement equation for the instrument is more severely misspecified. Here, we focus 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 v_t. 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 C for details). Figure A-3 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.

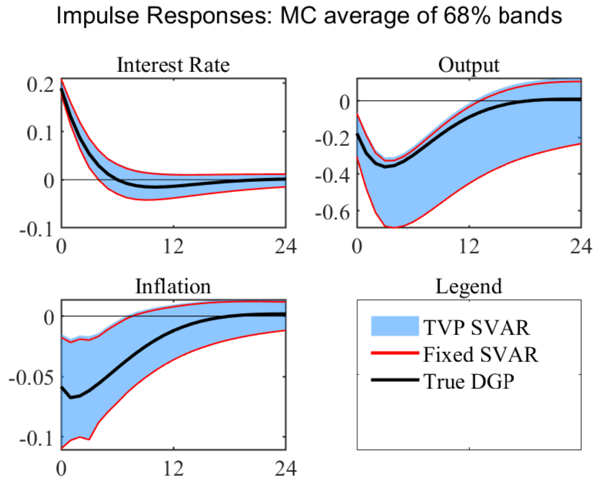


Figure A-3: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in measurement error.

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 A-4 are remarkably stable and generally do not hover around zero, but instead are close to the true value of 0.12, a stark contrast to the posterior paths obtained using US data.

Finally, we ask how our approach performs when confronted with data where there is stochastic volatility in the true monetary policy shock.[8] We again keep the paths of the instrument's volatility the same as in our benchmark specification. Figure A-5 shows that our approach again performs as well as the fixed coefficient version.[9]

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 A-6), but in line with our previous discussion. Due to the relative stability, it is then also not surprising that our approach does as well as the fixed coefficient approach in this

[8]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).

[9]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, while the estimated impulse responses are still one-standard deviation shocks according to the estimated standard deviation.

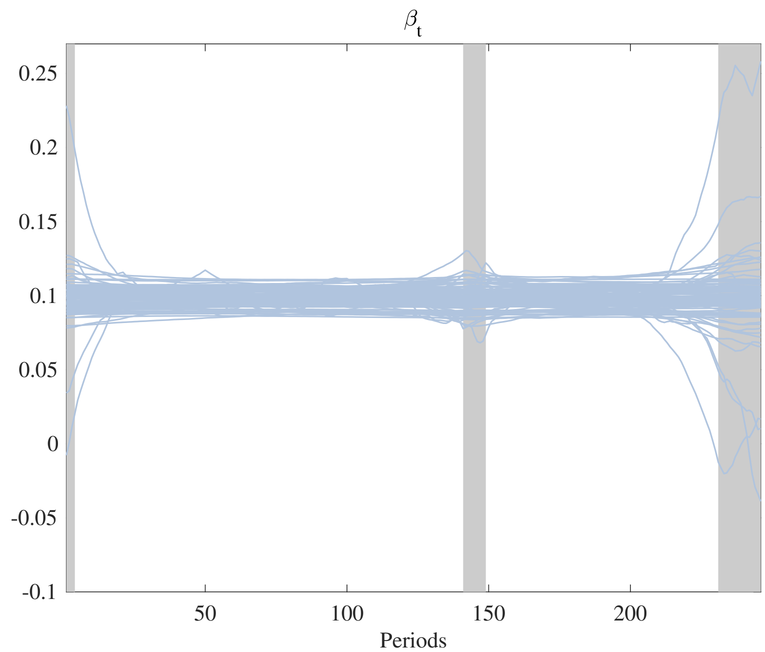


Figure A-4: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of v_t.

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 experiment involves one subtlety that we alluded to previously and 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. As we have shown above, such a transformation is always possible in models with stochastic volatility. However, this does 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 the latter 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 variance) inherits changes in volatility via

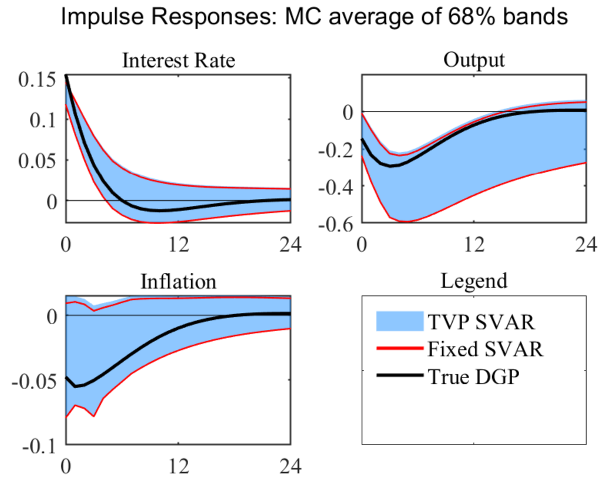


Figure A-5: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in the monetary policy shock.

the VAR that is estimated to have fixed coefficients and fixed forecast error variance.[10] Hence the estimated movements in β_t are 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. In summary, 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.

C 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.

[10]To convince readers that the issue of rescaling the monetary policy shock has no impact on our results, we show in Appendix C.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.

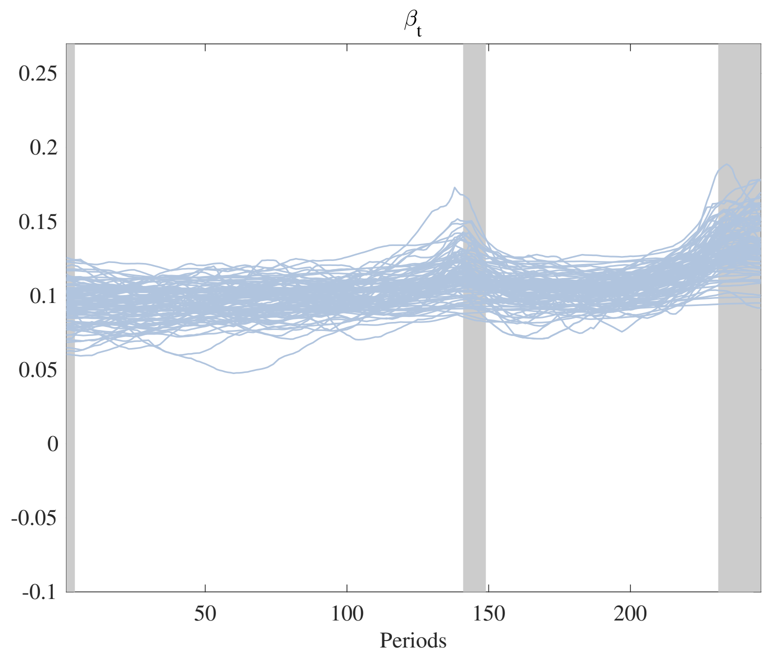


Figure A-6: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of the monetary policy shock.

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.[11] Note that in contrast to our estimated model (where e_t^{MP} is assumed to have unit variance and all scaling is captured in the impact matrix $\mathbf{B}$), 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

[11]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.

draws. The prior specification is exactly the same as in the empirical estimation.

C.1 Benchmark

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

$$m_t = e_t + v_t, Var(m_t) = (1+\kappa)\sigma_e^2 \tag{A-6}$$

$$m_t = v_t, Var(m_t) = \kappa\sigma_e^2 \tag{A-7}$$

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.

C.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:

$$m_t = \overline{\beta} e_t + v_t, Var(m_t) = \overline{\beta}^2 \sigma_e^2 + \sigma_{v,1}^2 \tag{A-8}$$

$$m_t = \overline{\beta} e_t + v_t, Var(m_t) = \overline{\beta}^2 \sigma_e^2 + \sigma_{v,2}^2 \tag{A-9}$$

We now need to solve the following two equations:

$$\overline{\beta}^2 \sigma_e^2 + \sigma_{v,1}^2 = (1+\kappa)\sigma_e^2 \tag{A-10}$$

$$\overline{\beta}^2 \sigma_e^2 + \sigma_{v,2} = \kappa\sigma_e^2 \tag{A-11}$$

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, $\sigma_{v,1}^2 = \sigma_e^2$.

C.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

$$m_t = \tilde{\beta} e_t + v_t, Var(m_t) = \tilde{\beta}^2 \sigma_{e,1}^2 + \tilde{\sigma}_v^2 \tag{A-12}$$

$$m_t = \tilde{\beta} e_t + v_t, Var(m_t) = \tilde{\beta}^2 \sigma_{e,2}^2 + \tilde{\sigma}_v^2 \tag{A-13}$$

The equations we need to solve are:

$$\tilde{\beta}^2 \sigma_{e,1}^2 + \tilde{\sigma}_v^2 = (1+\kappa)\sigma_e^2 \tag{A-14}$$

$$\tilde{\beta}^2 \sigma_{e,2}^2 + \tilde{\sigma}_v^2 = \kappa\sigma_e^2 \tag{A-15}$$

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, $\sigma_{e,2}^2 = (1+\kappa)\sigma_e^2$.

C.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

$$m_t = \bar{e}_t + v_t, \tag{A-16}$$

where $\bar{e}_t$ 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-7 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-7 with Figure A-5. The posterior median paths of β_t are now flat (Figure A-8).

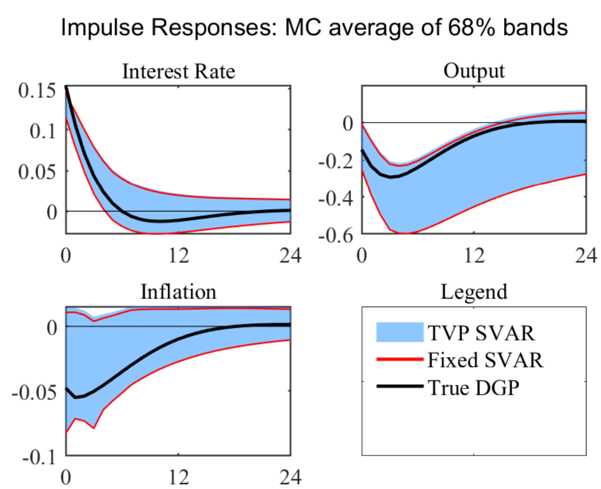


Figure A-7: Impulse responses for the data-generating process and the Monte Carlo replications: Stochastic volatility in the monetary policy shock, alternative instrument.

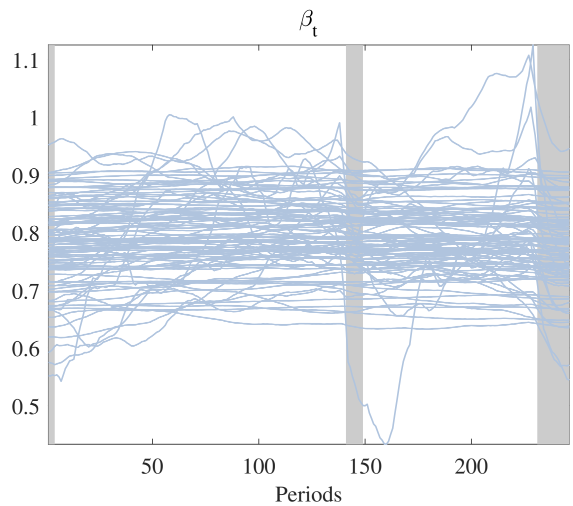


Figure A-8: Estimated paths of β_t in the Monte Carlo runs (posterior medians) when the data-generating process features time variation in the volatility of the monetary policy shock, alternative instrument.

D 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/`.

E Stochastic Volatility in the Monetary Policy Shock

To highlight that our specification delivers meaningfully different results from a specification with stochastic volatility in the monetary policy shock, we now estimate a structural VAR with stochastic volatility in the structural shocks. The stochastic volatility follows an independent log random walk specification for each shock. We include our instrument in our VAR as the first variable and use a recursive identification scheme along the lines of Plagborg-Møller & Wolf (2021). The stochastic volatility specification we use here is silent on where stochastic volatility comes from - measurement error or the true monetary shock. We choose to interpret the results *as if* there is no measurement error and we directly observe the monetary shock. In that sense, these results impose restrictions that are diametrically opposite to our benchmark specification.

Priors are comparable to our benchmark specification. The model we estimate is

$$z_t = \begin{bmatrix} m_t \\ \mathbf{y}_t \end{bmatrix} = \mathbf{c} + \mathbf{A}(L) z_{t-1} + \mathbf{B}\mathbf{u}_t$$

where $\mathbf{B}$ is lower triangular and $\mathbf{u}_t \sim N(\mathbf{0}, \Omega_t)$ with Ω_t being diagonal. Since our instrument is only available from 1990m1 to 2012m6, we also re-estimate our benchmark model using this exact sample. Figure A-9 shows that the assumption of stochastic volatility of the monetary policy shock delivers a much more muted response of the price level compared to our benchmark.

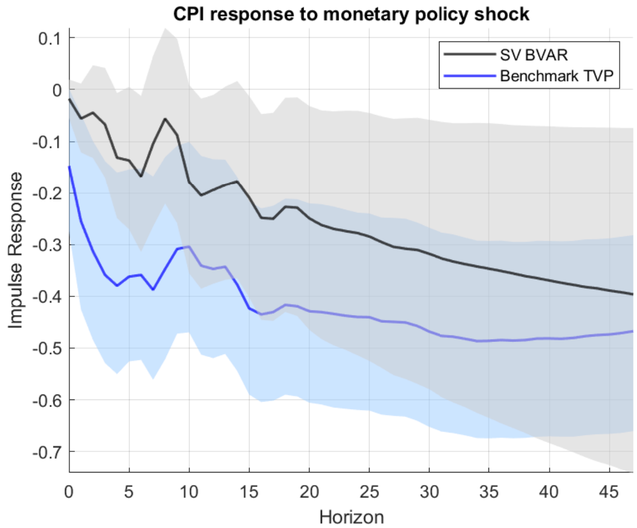


Figure A-9: Impulse responses (median and 68 percent posterior bands) to a 25 basis point monetary policy shock - stochastic volatility ion monetary policy shock vs time-variation in β_t.

F Varying the Amount of Information in the Sample

In this section, we carry out various Monte Carlo exercises to assess what happens when only a relatively small fraction of the sample comes with an informative instrument. The DGP is a bi-variate VAR(1) with coefficient matrix $A = [0.95\ 0; 0\ 0.9]$ and impact matrix $B = [1\ \ -0.5; 0.8\ 1]$. The structural shocks are iid N(0,1). The instrument is set to the true structural shock (we only consider the first shock) in the last 1, 10, or 50 percent of the sample. Otherwise, the instrument is drawn iid from a N(0,1) distribution. Each MC exercise is repeated 500 times. We consider sample sizes of $T \in \{200, 500, 1000\}$ and report the RMSE of the impulse responses at different horizons $h \in \{1, 6, 12, 24\}$. The relative RMSE is calculated as the benchmark model RMSE divided by the RMSE of the fixed coefficient model based on the posterior medians of the estimated impulse responses. As we can see, except with only one percent of observations featuring an informative instrument, our approach clearly outperforms the fixed coefficient alternative.[12]

[12]We obtain qualitatively the same results if we use a model where β_t evolves according to a two-state Markov chain.

Table A-4: Relative root mean squared error for first variable.

	T=200			T=500			T=1000		
	1%	10%	50%	1%	10%	50%	1%	10%	50%
$h = 1$	1.0081	0.7124	0.7141	1.0880	0.2978	0.6595	0.9585	0.1340	0.6365
$h = 6$	1.0114	0.8778	0.9417	1.1000	0.5649	0.8743	0.9701	0.3720	0.8541
$h = 12$	1.0060	0.9547	0.9455	1.0915	0.7626	0.9119	0.9827	0.5804	0.8869
$h = 24$	1.0007	0.9992	0.9550	1.0574	0.8819	0.9319	0.9991	0.7695	0.8990

Table A-5: Relative root mean squared error for second variable.

	T=200			T=500			T=1000		
	1%	10%	50%	1%	10%	50%	1%	10%	50%
$h = 1$	1.0318	0.7081	0.6034	0.9834	0.2086	0.6497	0.9530	0.2266	0.5426
$h = 6$	1.0086	0.7151	0.8343	0.9872	0.3320	0.9180	0.9553	0.4064	0.7884
$h = 12$	0.9981	0.7856	0.9217	0.9937	0.5301	0.9547	0.9582	0.6365	0.8914
$h = 24$	1.0017	0.9239	0.9973	1.0107	0.8421	0.9603	0.9688	0.9130	0.9447

G Additional Figures

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

G.1 Markov Switching

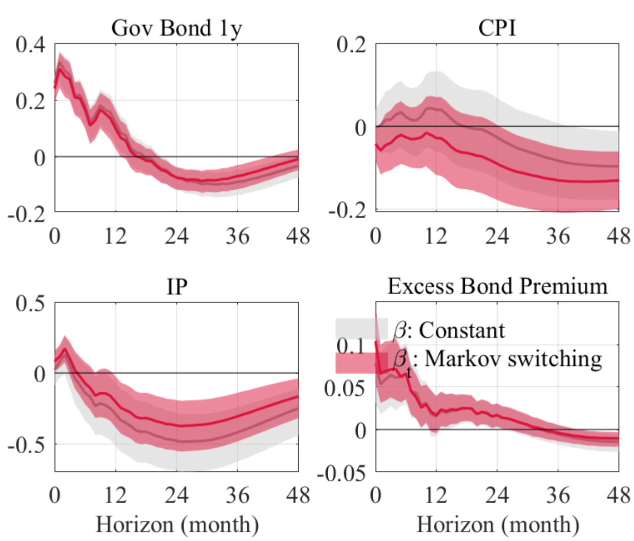


Figure A-10: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - Markov-switching specification.

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

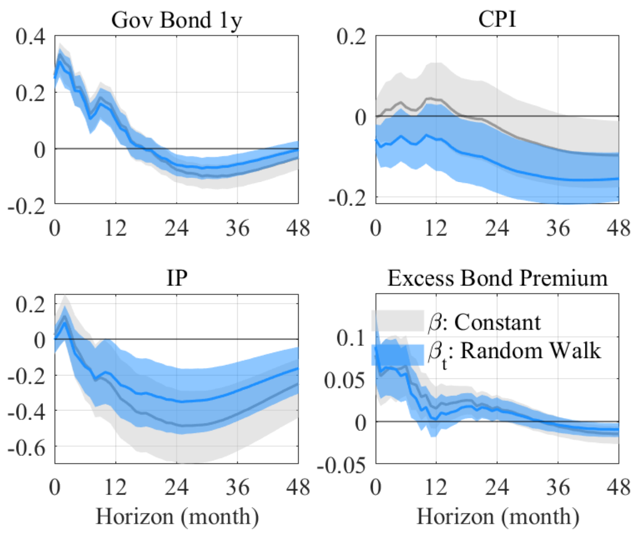


Figure A-11: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - $\tilde{m}_t$.

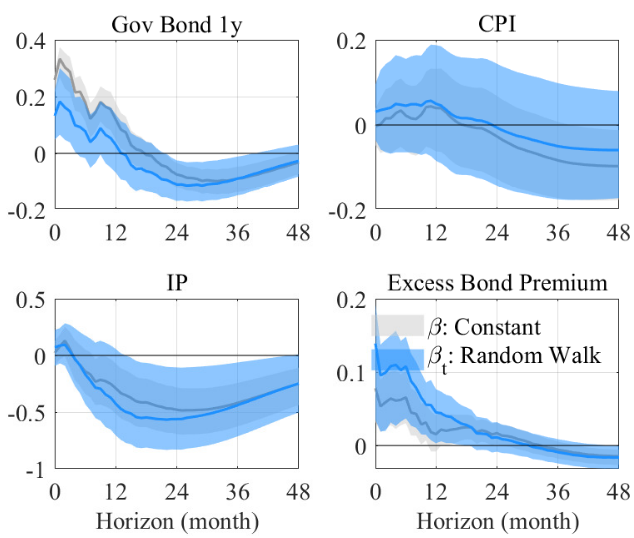


Figure A-12: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - $\overline{m}_t$.

G.3 Alternative Instruments

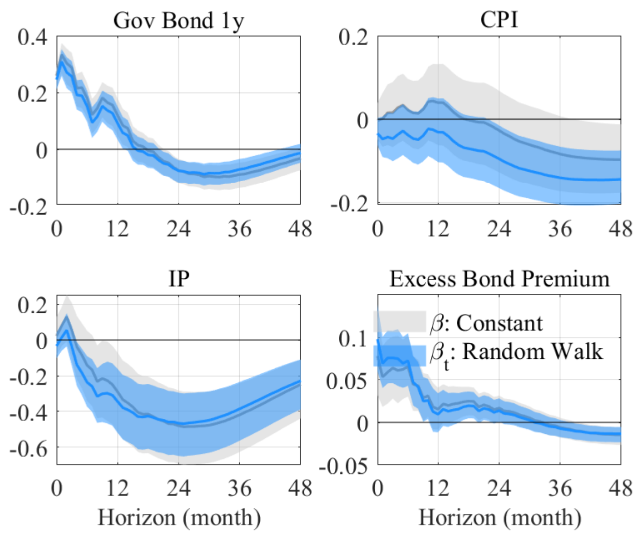


Figure A-13: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - modified instrument.

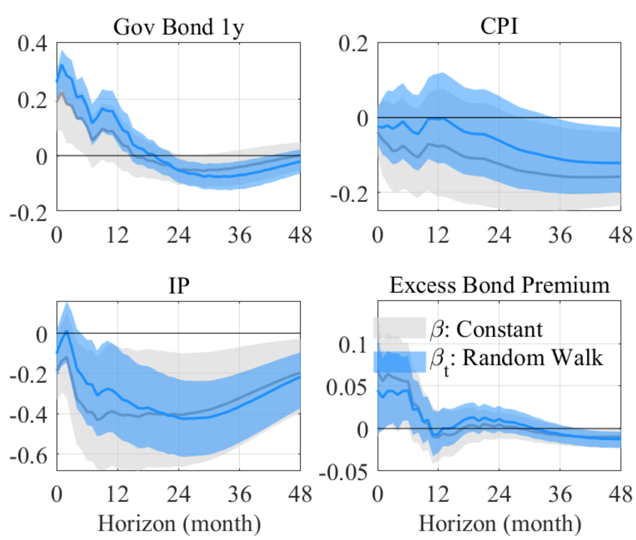


Figure A-14: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - Miranda-Agrippino & Ricco (2020) instrument.

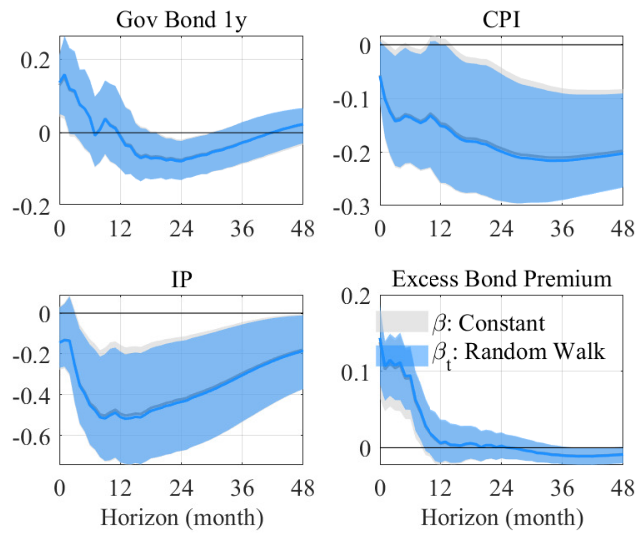


Figure A-15: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock - Bauer & Swanson (2022) instrument.

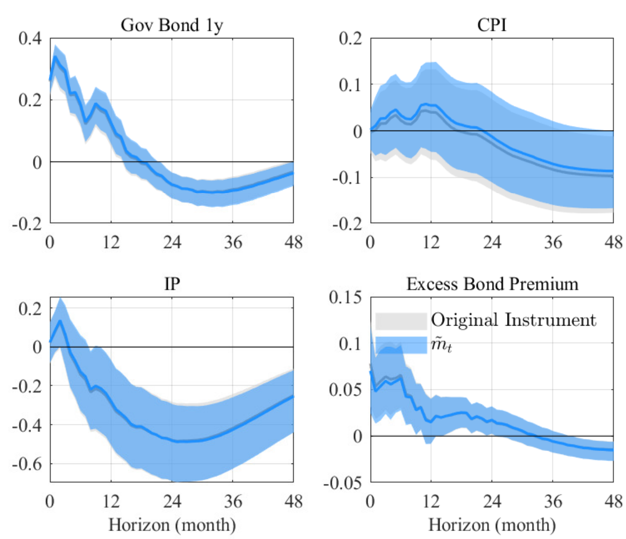


Figure A-16: Impulse responses (median and 68 percent posterior bands) based on fixed coefficient VAR for benchmark instrument and $\tilde{m}_t$

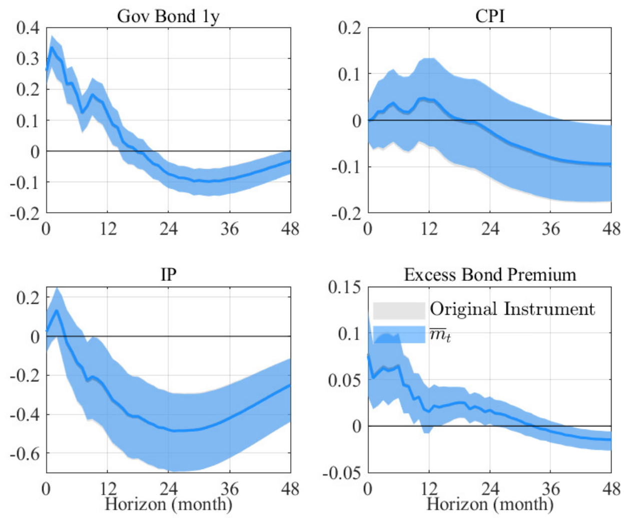


Figure A-17: Impulse responses (median and 68 percent posterior bands) based on fixed coefficient VAR for benchmark instrument and $\overline{m}_t$

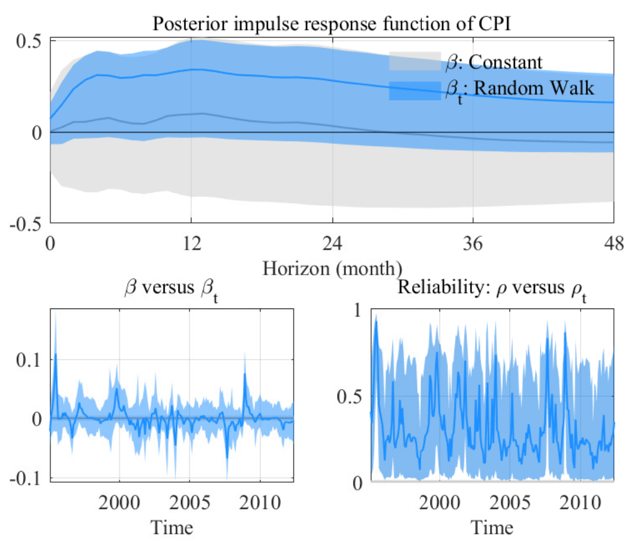


Figure A-18: Impulse responses (median and 68 percent posterior bands) based on Nakamura & Steinsson (2018) instrument.

G.4 More General Measurement Equation

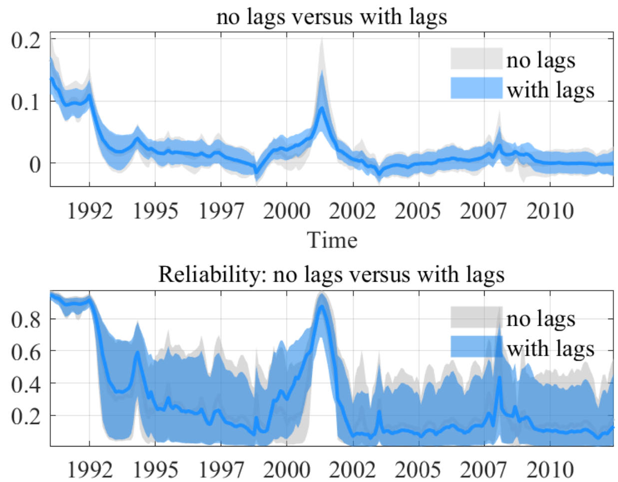


Figure A-19: β_t and ρ_t, one lag of VAR variables included in measurement equation of the instrument

G.5 Alternative paths for β_t in the DSGE Monte Carlo

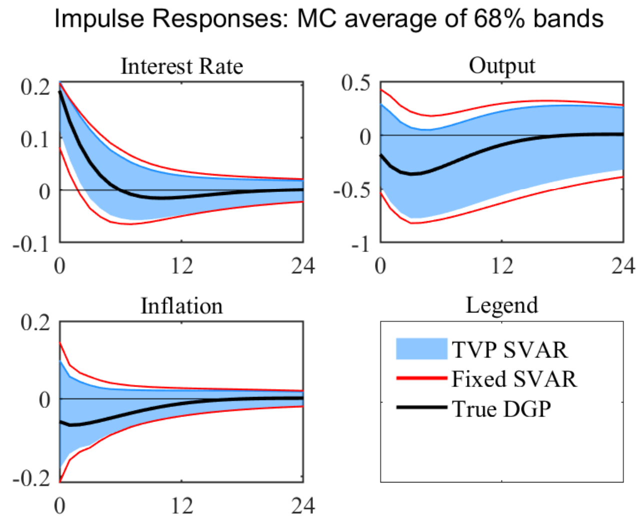


Figure A-20: Impulse responses for the data-generating process with time-varying β_t and the Monte Carlo replications, random transition paths for true β_t based on two-state Markov chain with transition probabilities matching transition used in main text.

H 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). We focus on the specification with a random walk specification for β_t.

Figure A-21 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-22: 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.

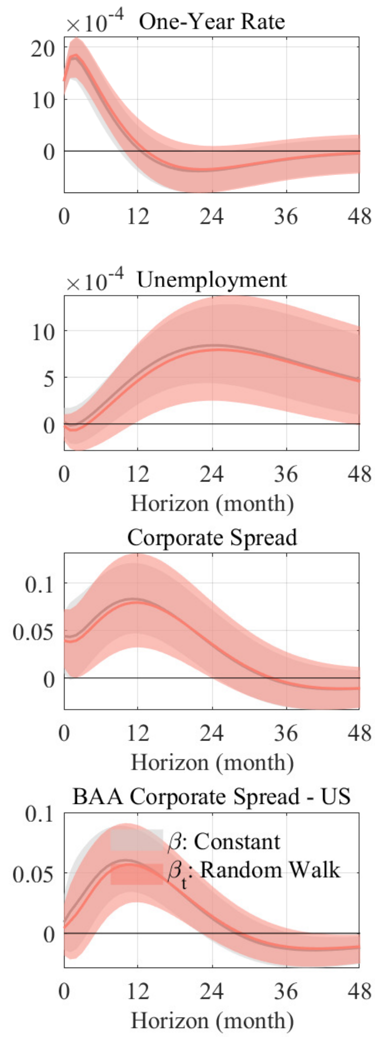


Figure A-21: Impulse responses (median and 68 percent posterior bands) to a one standard deviation monetary policy shock, UK.

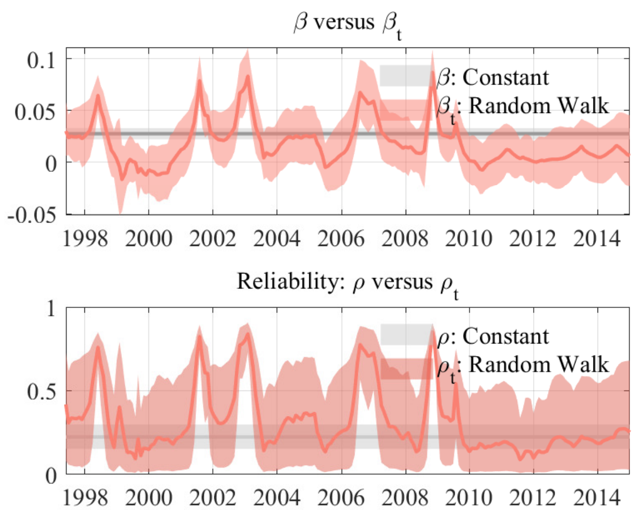


Figure A-22: Posterior of β_t and ρ_t, UK (median and 68 percent posterior bands).