Working Paper Series

Measurement Errors and Monetary Policy: Then and Now

WP 15-13

Pooyan Amir-Ahmadi
Goethe University Frankfurt

Christian Matthes
Federal Reserve Bank of Richmond

Mu-Chun Wang
University of Hamburg

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Measurement Errors and Monetary Policy: Then and Now

Pooyan Amir-Ahmadi*
Goethe University Frankfurt
Christian Matthes
Federal Reserve Bank of Richmond
Mu-Chun Wang
University of Hamburg

Working Paper No. 15-13

November 5, 2015

Abstract

Should policymakers and applied macroeconomists worry about the difference between real-time and final data? We tackle this question by using a VAR with time-varying parameters and stochastic volatility to show that the distinction between real-time data and final data matters for the impact of monetary policy shocks: The impact on final data is substantially and systematically different (in particular, larger in magnitude for different measures of real activity) from the impact on real-time data. These differences have persisted over the last 40 years and should be taken into account when conducting or studying monetary policy.

Keywords: real-time data, time-varying parameters, stochastic volatility, impulse responses

*We would like to thank Fabio Canova, Tim Cogley, and Dean Croushore as well as participants at the 2015 IAAE conference and the Cirano workshop on real-time data in Montreal for their comments. Miki Doan, Marisa Reed, and Daniel Tracht provided excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Introduction

When monetary policymakers evaluate the effects of their most recent policy decisions (say, to prepare for the next round of monetary policy decisions) they only have access to preliminary real-time estimates of macroeconomic data that have been collected after the policy decision they want to evaluate. Is the difference between real-time and final macroeconomic data important enough to be considered when analyzing and conducting monetary policy? We revisit this question asked by Croushore & Evans (2006) in light of recent evidence (Aruoba (2008)) that the measurement errors in macroeconomic data are far from satisfying the properties of classical measurement errors and evidence that there is substantial time variation in the dynamics of U.S. macroeconomic time series, as emphasized by Cogley & Sargent (2005) and Primiceri (2005).

We use a vector autoregression (VAR) with time-varying parameters and stochastic volatility estimated on data that includes real-time and final releases of macroeconomic data to uncover substantial time variation in the dynamics of measurement errors: We find that the measurement errors are significantly correlated for some variables, feature substantial changes in volatility and can be different from zero for long periods of time with magnitudes that are economically meaningful. We use a model with time-varying parameters and stochastic volatility because time variation in the dynamics and volatility of final data has been identified as important for (final) U.S. data by Cogley & Sargent (2005), Primiceri (2005), and Canova & Gambetti (2009), among others. Our paper shows that these features carry over to real-time data as well.

By using sign restrictions to identify a monetary policy shock, we establish that policymakers should indeed care about measurement errors. Differences between the impulse responses of real-time and final data on measures of real activity are significant and persist over time. As these differences are persistent over time, policymakers should take them into account.

Our work is related to the literature on time variation in macroeconomic dynamics such as Cogley & Sargent (2005), Primiceri (2005), and Gali & Gambetti (2009). The model we use to analyze time variation in our data is borrowed from those papers.[1] As pioneered by Canova & Nicolo (2002), Faust (1998) and Uhlig (2005), we use sign restrictions to identify monetary policy shocks. Canova & Gambetti (2009) use sign restrictions to identify monetary policy shocks in a VAR with time-varying parameters and stochastic volatility, but they do not consider real-time data.

Croushore & Evans (2006) tackle issues similar to ours, though in the context of a fixed coefficient VAR using either recursive or long-run restrictions. Their model of measurement error is less general than ours. For example, their models not only use fixed coefficient models, but also do not allow for biases (non-zero intercepts) in the relationship between different vintages of data. We find that these features matter, reinforcing the results by Aruoba (2008), and beyond that establish that measurement errors feature stochastic volatility and are correlated across variables.

In contrast to our work and Croushore & Evans (2006), the large majority of papers on real-time data focuses on statistical models of measurement error that do not identify effects of structural shocks. Jacobs & van Norden (2011) are motivated by the evidence in Aruoba (2008) and build a flexible model for a univariate measurement error series. In contrast to us, they model intermediate data releases, but do not consider the relationship of measurement errors across variables, time variation in the parameters, or stochastic volatility. Just as the model used in Jacobs & van Norden (2011), our model is general enough to allow for measurement errors that are correlated with either only final data ('news') or correlated only with the real-time data ('noise') as well as intermediate cases, as we show in the model section. Jacobs, Sarferaz, van Norden & Sturm (2013) build a multivariate version of Jacobs & van Norden (2011), but still abstract from stochastic volatility and time-varying parameters. Both Jacobs & van Norden (2011) and Jacobs et al. (2013) do not study the response of the economy to structural shocks, which is our main focus.

D'Agostino, Gambetti & Giannone (2013) use a VAR with time-varying parameters and stochastic volatility on real-time data to study the forecasting ability of models in this class.

[1] An overview of this literature is given in Koop & Korobilis (2010).

Fixed-coefficient VARs using various vintages of real-time data have previously been used to improve forecasting ability by Kishor & Koenig (2009) and Carriero, Clements & Galvao (2015), for example.

While we focus on post-WWII data, the issue of mismeasured data is of utmost importance when studying long-run historical data. While scholars using historical data usually do not have access to revised data for the entire sample, they sometimes explore overlapping data sources - Cogley & Sargent (2014) do this in a model for US inflation that features stochastic volatility for true data, but in contrast to our approach their model does not feature stochastic volatility for the measurement error. Croushore & Sill (2014) estimate a dynamic stochastic general equilibrium (DSGE) model on final data and then use the approach of Schorfheide, Sill & Kryshko (2010) to link real time data to the state variables of the estimated DSGE model. Similar to our findings, their findings show both that there are substantial differences between real-time and final data responses and that final data responses tend to be larger in absolute value.

In the next section we describe our model. We then show how it can be motivated by a DSGE model with asymmetric information before turning to results for our benchmark specification. Finally, we show that our findings are robust to alternative specifications: (i) using an alternative measure of real activity, employment growth, (ii) using an alternative identification scheme to identify monetary policy shocks, and (iii) using an alternative definition of final data.

2 The Model

We jointly model the dynamics of the first release of any data point published - we call this real-time data - and the latest vintage available at the time of the writing of this paper - which we use as a proxy for final data. Throughout this paper, we study the dynamics of vectors of the following form:

$$y_t = \begin{pmatrix} \pi_t^{real} \\ \pi_t^{final} \\ x_t^{real} \\ x_t^{final} \\ i_t \end{pmatrix} \tag{1}$$

where π_t denotes inflation, i_t the nominal interest rate, and x_t a measure of real activity. In our benchmark, x_t will be GDP growth, but we also study employment growth.[2] A superscript $real$ denotes real time data, whereas the superscript $final$ denotes final data. Throughout the paper real-time data refers to the first available release of a data point. We want to recover the joint dynamics of real-time and final data and ask what those dynamics tell us about the effects of monetary policy shocks on both real-time and final data. The dynamics of y_t are given by

$$y_t = \mu_t + \sum_{j=1}^{L} A_{j,t} y_{t-j} + e_t \tag{2}$$

where the intercepts μ_t, the coefficients on lagged observables $A_{j,t}$, and the covariance matrix Ω_t of e_t are allowed to vary over time. Following most of the literature that has used these models on quarterly data such as Del Negro & Primiceri (2015) and Amir-Ahmadi, Matthes & Wang (2014), we set the number of lags $L = 2$.

By writing down a model for real-time and final versions of the same data series, we have also implicitly defined a model of the measurement errors η_t^π and η_t^x:

$$\begin{pmatrix} \eta_t^\pi \\ \eta_t^x \end{pmatrix} = \begin{pmatrix} \pi_t^{real} \\ x_t^{real} \end{pmatrix} - \begin{pmatrix} \pi_t^{final} \\ x_t^{final} \end{pmatrix} = S y_t \tag{3}$$

[2] Jointly modeling the dynamics of real-time and final inflation, GDP growth, employment growth, and the nominal interest rate leads to issues of numerical instabilities in the Gibbs sampler we use to estimate the model. We thus study different variants of the model including one indicator of real activity at a time. We could have reduced the lag length, but that would have made our results less comparable to others in the literature. Similar issues are documented in Benati (2014), for example. For the same reason we also refrain from including intermediate data revisions as observables.

where S is a selection matrix.[3] We thus use a flexible time series model for the measurement errors that does not impose strong restrictions on the measurement errors - they can be correlated, have non-zero means, and feature substantial time variation in conditional moments.[4] This is important since Aruoba (2008) has found that data revisions are not necessarily well behaved.

To see that our model can capture measurement errors that feature both 'news' and 'noise' components (i.e. the measurement errors can be correlated with both final and real-time data), we can use a toy version of our model for a generic scalar variable c_t without time variation, stochastic volatility, or any dynamics:

$$\begin{bmatrix} c_t^{final} \\ c_t^{real} \end{bmatrix} = e_t = \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \tag{4}$$

where $e_t \sim N(0, \Omega^c)$. The measurement error $\eta_t^c = c_t^{real} - c_t^{final}$ is then defined as $e_{2,t} - e_{1,t}$. Consider as an example the following model for e_t:

$$e_{1,t} = w_t$$

and

$$e_{2,t} = w_t + v_t$$

where w_t and v_t are independent Gaussian random variables. Then we have $\eta_t^c = v_t$, which is independent of the final data $c_t^{final} = w_t$. Reversing the roles of c_t^{final} and c_t^{real} shows that measurement error can be independent of real-time data in our framework. To see an intermediate case where the measurement error is correlated with both real-time and final data, assume that $e_{1,t} = w_t$, as before, but now $e_{2,t} = 2w_t + v_t$ so that $\eta_t^c = w_t + v_t$, which is correlated with both real-time and final data.

To concisely describe the model we use to study time variation in the parameters

[3]

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}$$

[4] The measurement errors inherit these features from the variables in the VAR.

of the model, we define $X_t' \equiv I \otimes (1, y_{t-1}' ..., y_{t-L}')$ and rewrite (2):[5]:

$$y_t = X_t' \theta_t + e_t \tag{5}$$

$$\theta_t = \theta_{t-1} + u_t \tag{6}$$

Following Primiceri (2005), it is convenient to break the covariance matrix of the reduced-form residuals into two parts as implied by the following equation:

$$e_t = \Lambda_t^{-1} \Sigma_t \varepsilon_t \tag{7}$$

where ε_t is a vector of independently and identically distributed (iid) Gaussian innovations with mean 0 and covariance matrix I. Λ_t is a lower triangular matrix with ones on the main diagonal and representative non-fixed element λ_t^i. Σ_t is a diagonal matrix with representative non-fixed element σ_t^j. Those elements vary over time according to:

$$\lambda_t^i = \lambda_{t-1}^i + \zeta_t^i \tag{8}$$

$$\log \sigma_t^j = \log \sigma_{t-1}^j + \nu_t^j \tag{9}$$

All innovations are normally distributed with covariance matrix V, which, following Primiceri (2005), we restrict as follows:

$$V = Var \left[\begin{pmatrix} \varepsilon_t \\ u_t \\ \zeta_t \\ \nu_t \end{pmatrix} \right] = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & T & 0 \\ 0 & 0 & 0 & W \end{pmatrix} \tag{10}$$

T is further restricted to be block diagonal, which simplifies inference. ζ_t and ν_t are vectors that collect the corresponding scalar innovations described above. We estimate this model using the Gibbs sampling algorithm described in Del Negro &

[5] I denotes the identity matrix.

Primiceri (2015)[6].

We follow Primiceri's choice of priors, adjusted for the size of our training sample. The Gibbs sampler we use is outlined in detail in Del Negro & Primiceri (2015). In contrast to Cogley & Sargent (2005), we do not impose any restrictions on the eigenvalues of the companion form matrix of the VAR. We do so both on empirical grounds (in Amir-Ahmadi et al. (2014) we show that there is a substantial probability of temporarily explosive dynamics in US data) and theoretical grounds (Cogley, Matthes & Sbordone (2015) show that temporarily explosive dynamics can emerge naturally in micro-founded dynamic equilibrium models when agents are learning).

In order to ascertain whether or not monetary policy shocks affect real-time and final data differently and if those effects have changed over time, we identify monetary policy shocks using our VAR models. As our benchmark, we use sign restrictions. An identification scheme of this sort has been used in time-varying parameter VARs with stochastic volatility by Benati & Lubik (2014), Canova & Gambetti (2009), and Amir-Ahmadi et al. (2014), among others.

Structural models used by macroeconomists give us a good sense of the signs of the effects of monetary policy shocks on final data. The corresponding effects on real-time data are less clear, and depend on the specifics of any particular DSGE model with both real-time and final data (we present one such model in the next section). This consideration leads us to only use sign restrictions on final data, not on real-time data. We are thus not imposing any restrictions on the impulse response functions of real-time data. We restrict the nominal interest rate to not decrease after a positive monetary policy shock and both final inflation and final GDP growth to not increase after a positive monetary policy shock. We impose those restrictions on impact and for the first two periods after impact - this is the same number of periods as chosen by Benati (2010), for example. For the model with employment growth, we impose that, in addition to the restrictions on inflation and the nominal interest rate, final employment growth can not increase after a positive monetary policy shock.

[6]We use 250,000 posterior draws, out of 200,000 are used as burn-in. We then keep every 10th draw of the remaining 50,000, resulting in 5000 stored draws. We have assessed and ensured convergence of the Markov Chain using the standard diagnostics.

The equation for the nominal interest rate that is recovered using our identification scheme gives, by construction, the nominal interest rate as a function of lagged real-time and final data. The lagged final data is not directly observable by the central bank when it makes its decisions every period. As such, we do not directly interpret the nominal interest rate equation as a monetary policy rule (in contrast to Canova & Gambetti (2009)), but instead interpret it as the central bank responding to observables such as survey and forecast data, which in turn depend on both real-time and final data. This assumption can be justified by referring to micro-founded structural models where the private sector has an informational advantage and thus knows the final data before the central bank does. The next section describes one DSGE model with these features. That model shares features with work by Aoki (2003), Nimark (2008), Lubik & Matthes (2014), and Svensson & Woodford (2004), for example.[7] Given that, for computational reasons, we can not include additional observables such as intermediate data releases, the only viable alternative would have been to restrict the central bank to only react to lagged real-time (i.e. first release) observables. This approach would have substantially underestimated the information available to the central bank. We think our approach better approximates the actual (large) information sets considered by central banks when making their decisions.[8]

3 The Choice of Variables for Our VAR - A DSGE-based Motivation

In this section, we describe one (admittedly, along many dimensions, very simple) DSGE model that delivers, as its reduced form, a VAR in the same state variables

[7]To keep their models tractable, those papers either assume relatively simple stochastic processes for the measurement error that can not match our findings on the properties of measurement errors, or they assume that the true realization of the data is observed by the central bank, but only with a lag.

[8]In future work, we plan to relax this assumption and instead of using 'final' data use the latest available vintage of data each quarter as new data becomes available. We view this as a separate project since in such a VAR we could no longer study the joint dynamics of real time and final data, which is our main goal in the current paper.

that we use in our empirical analysis. We do this to motivate our choice of variables, but also to highlight that we can indeed identify monetary policy shocks using the observables described above.[9] The DSGE model does not feature time-varying dynamics or stochastic volatility - those features could be added by introducing learning along the lines of Cogley et al. (2015), for example. For simplicity, the DSGE model presented here has a VAR of order 1 as its reduced form, whereas we work with a VAR of order 2 in the empirical analysis. Additional lags could easily be introduced in the DSGE model, but would not add any insight to the exposition. We directly present the linearized version of the model. The first two equations give the dynamics of our real variable x_t and inflation π_t, conditional on iid shocks z_t and g_t as well as the nominal interest rate i_t:

$$x_t = a_x E_t x_{t+1} + b_x(i_t - E_t \pi_{t+1}) + g_t + c_x x_{t-1} \tag{11}$$

$$\pi_t = a_\pi E_t \pi_{t+1} + b_\pi x_t + z_t + c_\pi \pi_{t-1} \tag{12}$$

Following the literature cited in the previous section, the private sector in this model has access to final data and thus the IS and Phillips curves are straight out of standard New Keynesian models. The variable x_t in micro-founded DSGE models is usually the output gap. Real-time measures of the output gap are unfortunately chronically unreliable (Orphanides & van Norden (2002)), so we choose to use alternative measures of real activity in our empirical analysis instead. This lack of reliability does not come from the real-time nature of output data, but rather from the large uncertainty surrounding trend estimates of GDP in real time. Thus, it seems likely that real time measures of the output gap would not be used in monetary policy decisions.

[9]This is not immediately obvious since our sign restrictions imply that the nominal interest rate is the monetary policy instrument. The nominal interest rate in our VAR reacts to final data lagged once and twice, which is not in any central bank's information set. The DSGE model in this section shows how our approach can be justified.

Next, we implicitly define the measurement errors:

$$x_t^{real} = x_t + \eta_t^y \tag{13}$$

$$\pi_t^{real} = \pi_t + \eta_t^\pi \tag{14}$$

The two non-standard features of this economy are the monetary policy rule and the dynamics of the measurement error, which are substantially more general than what is usually assumed in the literature:

$$i_t = a_i x_{t-1}^{real} + b_i \pi_{t-1}^{real} + c_i i_{t-1} + d_i E_{t-1} x_t + e_i E_{t-1} \pi_t + \varepsilon_t^i \tag{15}$$

$$\begin{aligned} \eta_t &= [\eta_t^x \ \eta_t^\pi]' \\ &= M_1 [y_t \ \pi_t \ i_t]' + M_2 [x_{t-1} \ \pi_{t-1} \ i_{t-1}]' + \rho \eta_{t-1} + M_3 \varepsilon_t^{real} \end{aligned} \tag{16}$$

Note that the nominal interest rate only reacts to real-time data as well as survey measures of expectations gathered in the previous period. The monetary policy shock is denoted ε_t^i. While this is certainly restrictive (central banks have access to revised data for past periods), this assumption respects the information constraints of actual central banks in that the central bank can not react to final data for recent periods. Data is revised for multiple years (as mentioned before, Aruoba (2008) uses a window of three years to define final data for the United States, for example), whereas there are multiple policy meetings per year for all major central banks.

The definition of the measurement error dynamics (we denote the vector of measurement errors by η_t) allows for dependence on lagged measurement errors as well as lagged and contemporaneous final data. We do not state a theory that delivers these dynamics, but instead want to show that even with general dynamics like these, the dynamics are fully captured by the variables we use in our VAR.

To solve the model, we can first reduce the system and use the implicit definition of the measurement errors to eliminate η. We are then left with a system that does include $E_{t-1}\pi_t$ and $E_{t-1}x_t$ as state variables. Solving this model using standard methods for linear rational expectations models such as Gensys (Sims (2002)), we get the following reduced form:

$$\begin{bmatrix} x_t \\ \pi_t \\ x_t^{real} \\ \pi_t^{real} \\ i_t \\ E_t x_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = A \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \\ x_{t-1}^{real} \\ \pi_{t-1}^{real} \\ i_{t-1} \\ E_{t-1} x_t \\ E_{t-1} \pi_t \end{bmatrix} + B \begin{bmatrix} g_t \\ z_t \\ \varepsilon_t^{real} \\ \varepsilon_t^{i} \end{bmatrix} \tag{17}$$

where A and B are matrices that are returned by the solution algorithm for linear rational expectations models. If we take time $t-1$ conditional expectations of this system, the resulting first two equations of that system give $E_{t-1}x_t$ and $E_{t-1}\pi_t$ as a linear function of x_{t-1}, π_{t-1}, x_{t-1}^{real}, π_{t-1}^{real}, i_{t-1}, and $E_{t-1}x_t$ and $E_{t-1}\pi_t$. Those two equations can thus be solved for $E_{t-1}x_t$ and $E_{t-1}\pi_t$ as a function of x_{t-1}, π_{t-1}, x_{t-1}^{real}, π_{t-1}^{real} and i_{t-1}. Doing this and replacing the expectation terms in the system above gives the reduced form dynamics in terms of the variables that we use in our VAR:

$$\begin{bmatrix} y_t \\ \pi_t \\ y_t^{real} \\ \pi_t^{real} \\ i_t \end{bmatrix} = A_2 \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \\ y_{t-1}^{real} \\ \pi_{t-1}^{real} \\ i_{t-1} \end{bmatrix} + B_2 \begin{bmatrix} g_t \\ z_t \\ \varepsilon_t^{real} \\ \varepsilon_t^{i} \end{bmatrix} \tag{18}$$

where A_2 and $B2$ are the matrices corresponding to A and B in the original solution after the lagged expectations terms have been substituted out.

In this model, the forecast error in the nominal interest rate equation is the monetary policy shock (all other right-hand side variables in the monetary policy rule are predetermined), even though some of the right hand-side variables in the reduced form nominal interest rate equation are not in the central bank's information set. If we took this model literally, we could be tempted to use this insight to directly estimate the monetary policy error. Instead, we see this DSGE model as one possible model that delivers a reduced form in line with our empirical specification. Thus, we choose to use sign restrictions to identify the monetary policy shock instead.

4 Data

As our benchmark, we use the Philadelphia Fed's real-time database (Croushore & Stark (2001)) to construct a sample of annualized quarterly real-time and final inflation (based on the GNP/GDP deflator) and annualized quarterly real-time and final real GNP/GDP growth[10]. As a robustness check, we also use annualized quarterly real-time and final employment growth (based on nonfarm payroll employment).[11] Real-time growth rates are calculated using all available data when an estimate of the latest level of the corresponding series is first available - the growth rate of GDP over the last year at any point in time is defined as the ratio of the latest real GDP release to the most current available vintage of real GDP one quarter earlier, for example.

As a proxy for final data, we use the most recent vintage available to us. Other approaches are certainly possible - Aruoba (2008) defines the final data as the vintages available after a fixed lag (for most variables 3 years). In our section on robustness checks we present additional results that use this alternative definition of final data. The real-time data is available starting in the fourth quarter of 1965. The last vintage we use is from the second quarter of 2014 (incorporating data up to and including the first quarter of 2014). We use 40 observations to initialize the prior for our time-varying-VAR model. For the nominal interest rate (which is measured without error), we use the average effective Federal Funds rate over each quarter.

Figure 1 plots the real-time data and the measurement error as defined in the previ-

[10]From now on we will refer to this variable as GDP growth.

[11]The employment series is originally monthly. We define employment within a quarter as the average employment over the three months belonging to that quarter. The real-time (or first available) estimate is defined as the estimate available in the middle month of the following quarter (in line with the definition of the other quarterly variables). We also estimated versions of our model using the real-time and final unemployment rate, but there are substantially fewer revisions in the unemployment rate by the very nature in which the data for the unemployment rate is collected - it is a survey-based measure. More details on the difference between real-time and final data for a broader set of variables can be found in Aruoba (2008).

ous section. To get the final data, we have to subtract the measurement error from the real-time data - positive measurement error implies that the real-time measurement is higher than the final data. To convince yourself that the difference between real-time and final data can be meaningful, it suffices to look at the mid-1970s: Real-time GDP growth and employment growth actually were lower than in the most recent recession, but there were substantial revisions to that data later on - the measurement errors associated with those errors is negative, meaning that data was revised upward substantially.[12] In the following section we will analyze the time series properties of the measurement errors and check how their behavior has changed over time. We will see that it is indeed important to allow for time variation in the dynamics of these series.

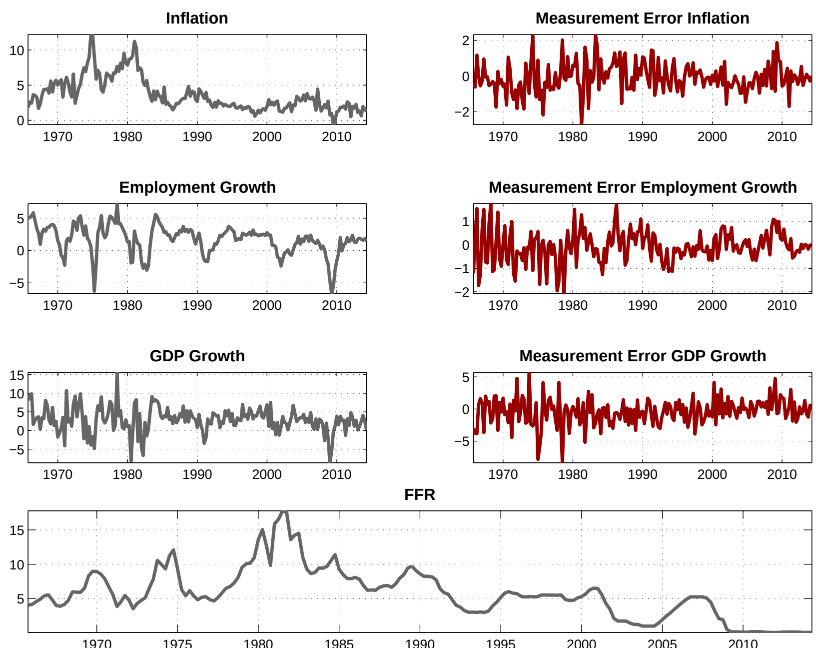

Figure 1: Real-time data and total revision

[12]Lubik & Matthes (2014) use a learning model to model the choices of a central bank that only has access to real-time data as it makes its decisions. Just as Orphanides (2002), they highlight that mis-measured data had a big impact on U.S. monetary policy in the 1970s. In the current paper, we instead focus on the impact of monetary policy shocks on both real-time and final data.

5 Results

5.1 The Time-Varying Properties of Measurement Errors

First, we want to describe how the properties of the measurement errors in inflation and GDP growth have changed over time. Cogley & Sargent (2005) have pioneered the use of 'local to time t' moments to study changes in the dynamics of VARs with time-varying parameters and stochastic volatility. In short, they calculate (unconditional) moments of the data governed by equation 2 at each point in time assuming that the coefficients will remain fixed over time. That way they recover a sequence of moments over time. This is feasible in their setup because they impose restrictions on the eigenvalues of the companion form matrix of the VAR. We, on the other hand, do not impose any such restrictions for the reasons mentioned before. Instead, we study forecasts from our model based on smoothed or full-sample parameter estimates (assuming, similar to Cogley & Sargent (2005), that the coefficients will not change in the future) and calculate the moments of forecasts of the measurements errors. It is important to emphasize that we use these moments of forecasts as low-dimensional summary statistics that capture the dynamics of our model. We can think of these moments as finite horizon versions of the summary statistics used by Cogley & Sargent (2005). We focus here on one-year ahead forecasts. Increasing the forecast horizon substantially would increase the uncertainty surrounding the estimated forecast moments exactly because we do not impose any restrictions on the dynamics of the VAR.

Figure 2 plots the median and 68 % posterior bands for the one-year ahead forecasts of the measurement error based on the model with GDP growth.[13] Our model predicts substantial measurement errors one year in advance. We can think of these one-year ahead forecasts as a proxy for trends or more generally the persistent parts of the measurement errors.[14] Our results confirm those in Aruoba (2008), who finds that measurement errors in many variables do not have a mean of zero. Throughout

[13]The date on the x-axis represents the date of the conditioning information.

[14]Interpreting forecasts as trends has a long tradition in empirical macroeconomics going back to Beveridge & Nelson (1981).

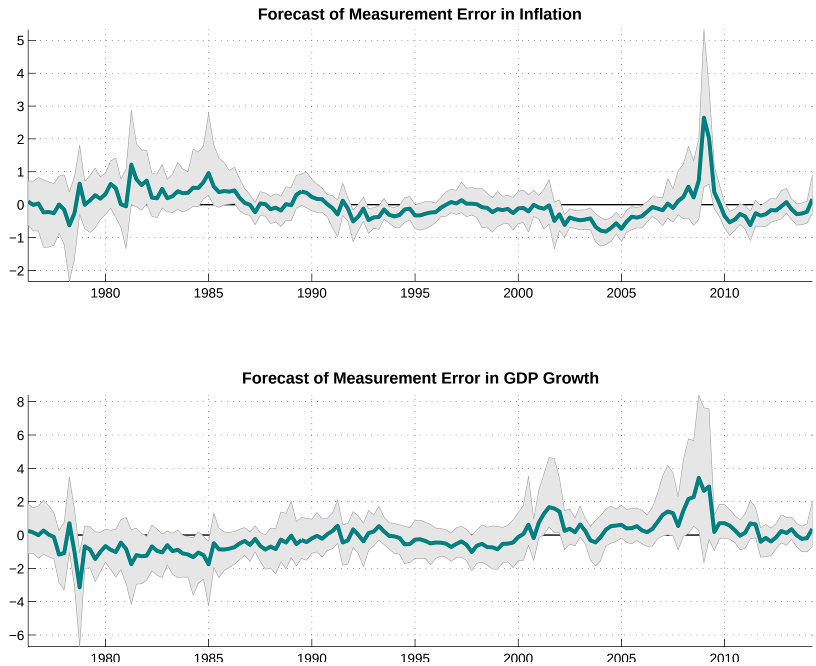

Figure 2: One-year ahead forecasts of measurement errors

most the 1980s the one-year ahead forecast in the measurement error of inflation is positive, of an economically meaningful size, and borderline statistically significant - inflation was initially overestimated during that period. During the 1990s and up to the financial crisis, inflation instead tended to be initially underestimated. During the financial crisis, inflation was substantially overestimated initially.

The measurement error in real GDP growth is negative during the 1980s (meaning that GDP growth was initially estimated to be lower than the final data suggests), before turning statistically insignificant during the 1990s. From 2000 to the financial crisis we see an initial overestimation of GDP growth (with a substantial overestimation during the financial crisis).

We now turn to higher moments of the forecasts. Figure 3 plots the volatilities of the one-year ahead forecasts of measurement errors and the associated correlation between the forecasted measurement errors. Both volatilities share a similar pattern[15]

[15] It is common in models of the type we use here, used in conjunction with the type of data that we analyze, that the volatilities of the variables in the VAR share a similar pattern - see for example Del Negro & Primiceri (2015).

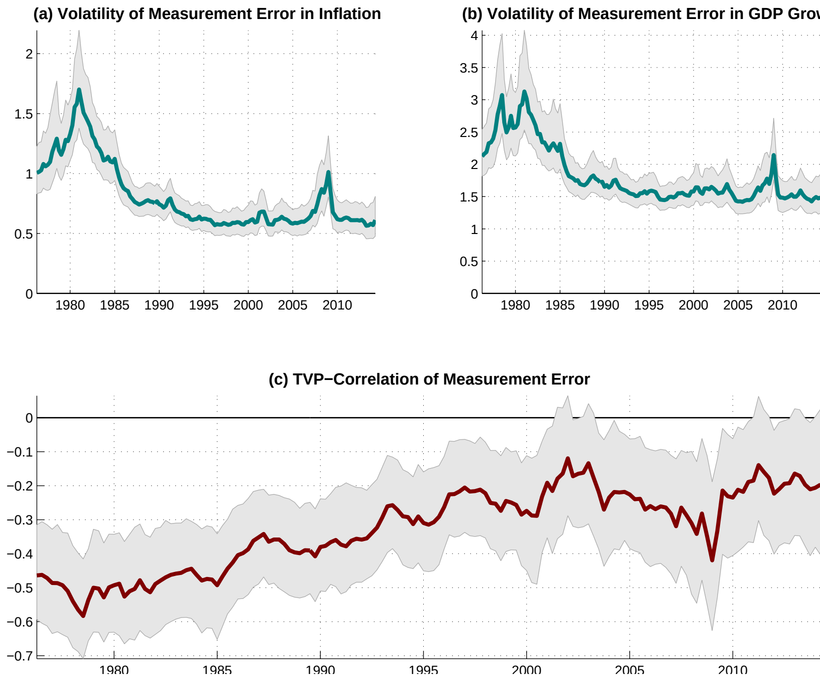

Figure 3: Volatility and correlation of forecasted measurement errors

- high volatility in the 1970s and early 1980s, a decline afterward and a noticeable uptick in volatility during the recent financial crisis. Interestingly, the correlation between the measurement errors is significantly negative throughout our sample, but has an upward trend for most of our sample that is only broken during the early 2000s. A negative correlation implies that an increase in the measurement error of GDP growth (real-time GDP growth becomes larger relative to final GDP growth) is associated with a decrease in the measurement error in inflation (final inflation becomes larger relative to the real-time measurement), so that an initial overestimation of GDP growth tends to be associated with an underestimation of inflation. Since the magnitude of the correlation decreased substantially over time, this pattern has become weaker over time. To summarize, a simple model of measurement errors that models them as being independent across variables and having constant innovation variance can miss important features of observed measurement errors.

5.2 The Effects of Monetary Policy Shocks Over Time

We first show impulse responses for different periods. We follow the standard approach in the literature to construct these impulse responses: For each time period, we draw parameters from the posterior distribution for that period and then keep these coefficients fixed as we trace out the effects of a monetary policy shock. We focus on impulse responses at short horizons because that is where we find the largest difference between real-time and final data. Since we are interested in the differences between the effects on real-time and final data (rather than changes in the impulse responses functions over time *per se*), we use one standard deviation shocks, where the standard deviation changes over time. This will give us a sense of how the impact of a usual shock has changed over time. Figure 4 plots the evolution of the nominal interest rate to such a shock. The black line gives the pointwise median response and the gray bands cover the area from the 15th to the 85th percentile of the response with each of the 5 shades of gray covering the same probability. We can see that there are differences on impact over time (in particular, the standard deviation of monetary policy shocks decreases), but the overall median pattern remains stable over time. In contrast, there is substantial time variation in the uncertainty surrounding the median response. At some points in time there are some draws that imply explosive behavior of the nominal interest rate.

Figure 5 plots the responses of real-time and final inflation to the same monetary policy shock. The black line and gray areas correspond to the median and the 15th to 85th percentiles of real-time data responses, whereas the red lines represent the responses of final data. The bold red line is the median and the outer dashed red bands correspond to the same percentiles as the outermost error bands for real-time data (the 15th and 85th percentiles). For the most part the responses of real-time and final inflation are very similar, especially after 4 to 5 periods. The sign restrictions are mostly satisfied by responses of real-time inflation even though we do not impose those restrictions. Nonetheless, we do find significant differences. For example, in 1979 the median impact response of real-time inflation is twice as large as that of final data. Broadly speaking, we see a larger difference (on impact) for the first part

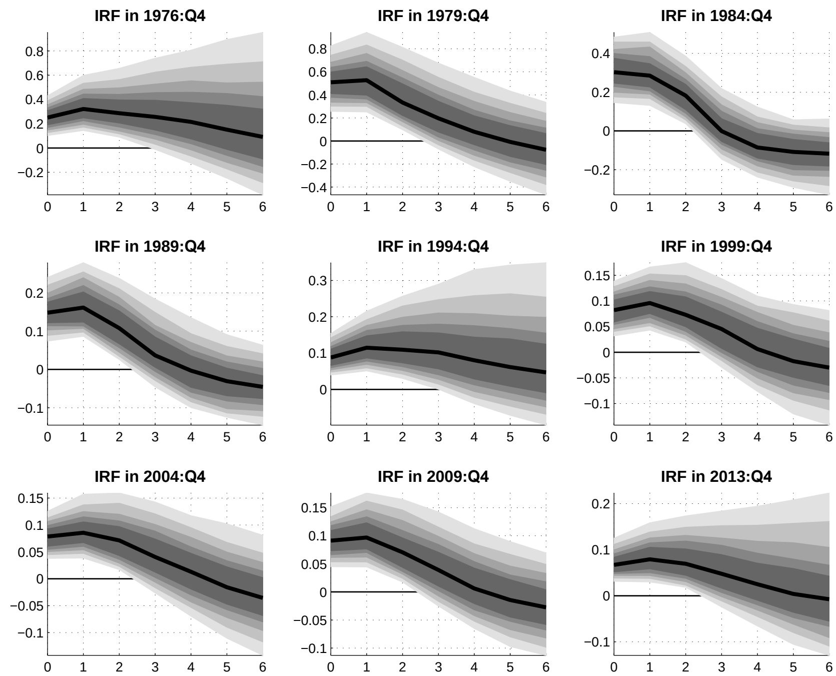

Figure 4: Impulse response functions for the nominal interest rate to a one standard deviation monetary policy shock.

of our sample (through the 1980s). Substantial differences in the responses between real-time and final inflation are present in the late 1970s to the late 1980s.

The impulse responses for GDP growth in figure 6 show a different pattern with more pronounced differences. On impact and for the first few periods after the shock hits, final GDP growth is lower than real-time GDP growth. This pattern is most pronounced in 1984 and 1989, but persists throughout our sample. The magnitude of those differences is economically significant - it matters if the response to a contractionary monetary policy shock on impact is a reduction of 0.25 percentage points in annualized GDP growth or 0.75 percentage points (these are roughly the magnitudes in 1984:Q4). We can also see that the sign restrictions we impose on final data are also met for most draws of the real-time data response.

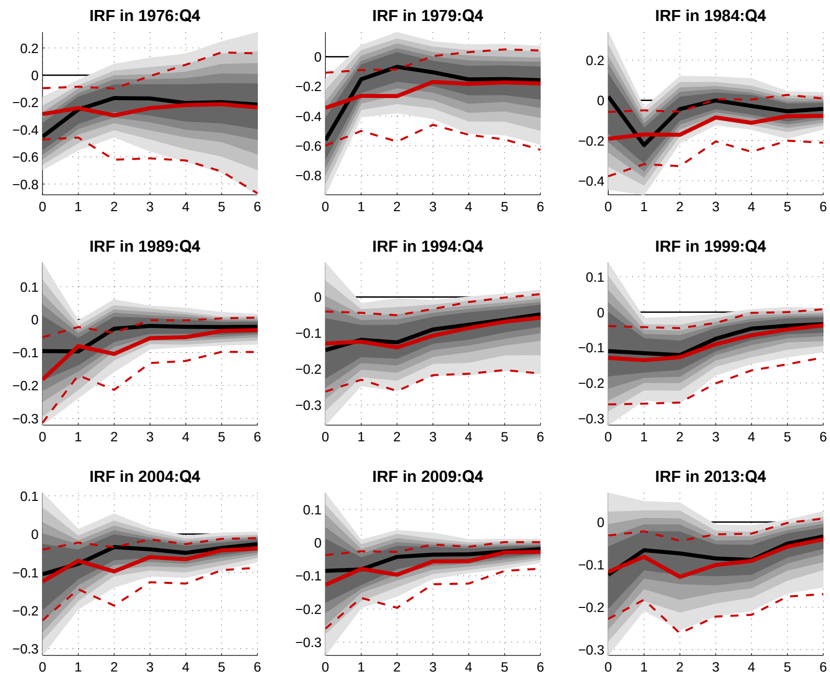

Figure 5: Impulse response functions for real-time (gray/black) and final (red) inflation to a one standard deviation monetary policy shock.

So far we have studied the marginal distributions of the impulse responses to real-time and final data and compared them to each other. We are also interested in the evolution of the joint distribution of impulse responses across real-time and final data. Our estimation algorithm allows us to study the joint posterior of impulse responses for a given horizon at each point in time. For each of those time/horizon pairs, we calculate an estimate of the joint posterior of real-time and final impulse responses (for each horizon and date this can be thought of as a scatterplot). We call $r_t^{real,i}(j)$ the impulse response at horizon j of real-time variable i ($i \in \{\pi, GDP, emp\}$) calculated using a draw of VAR coefficients at time t and $r_t^{final,i}(j)$ the response of the corresponding final variable (both calculated using the same parameter draw).

We first plot the median and the 15th and 85th percentile bands for the difference

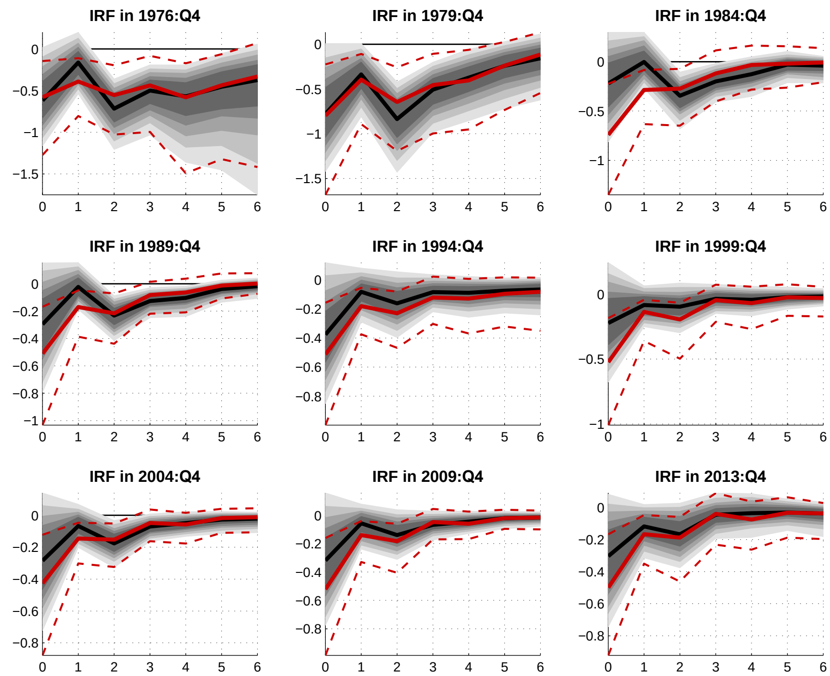

Figure 6: Impulse response functions for real-time (gray/black) and final (red) real GDP growth to a one standard deviation monetary policy shock.

between final data and real-time impulse responses of GDP growth[16] on impact (i.e. at horizon 0): $r_t^{final,GDP}(0) - r_t^{real,GDP}(0)$. A negative number means that the final data response is smaller than the corresponding real-time response. Figure 7 reveals that the median difference has been negative throughout our sample with a maximum of -0.1 and a minimum of -0.5 percentage points,[17] meaning that the response of final data is larger in magnitude than the response of real-time data as the final response is restricted to be negative on impact and the real-time response is negative for most draws. This again emphasizes that the differences are economically meaningful - central banks would care about these magnitudes. The 85th percentile of the differ-

[16]The median difference for inflation is centered at 0 for most of the sample, so we omit it here. This finding is also evident from figure 8.

[17]Remember that we use annualized values throughout this paper.

ence hovers around 0. Thus, there is a positive probability that the difference is 0 at any point in our sample as can be seen by the point-wise error bands[18]. However, the fact that the median difference is negative throughout and of a economically significant magnitude leads us to believe that it is indeed important to take the difference between real-time and final data seriously. Furthermore, policymakers regularly worry about worst case outcomes. We can see that the difference between the impact response for final and real-time data could be substantially larger in magnitude than what is suggested by the median numbers.

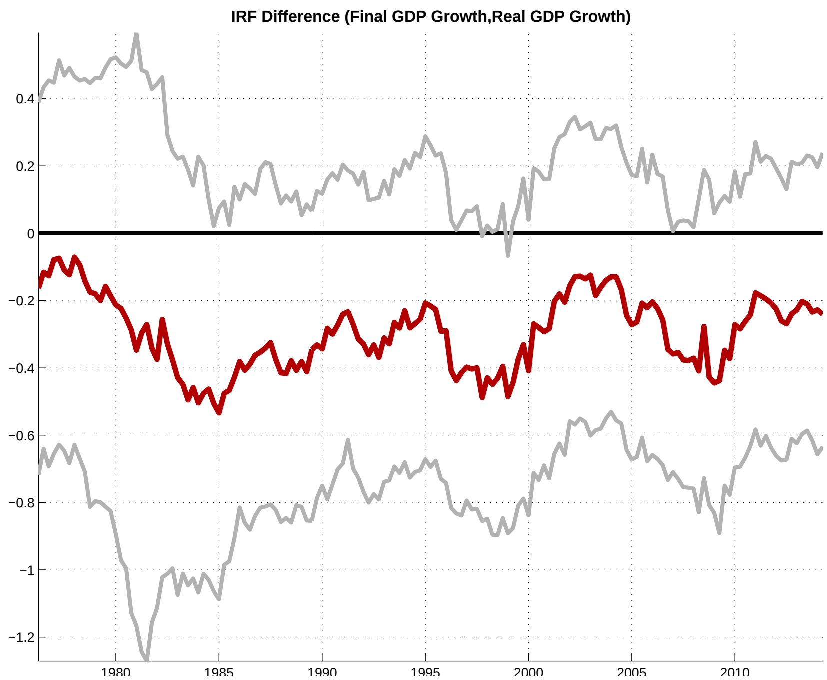

Figure 7: Differences between GDP growth impulse responses on impact: The distribution of $r_t^{final,GDP}(0) - r_t^{real,GDP}(0)$ over time.

For each period in our sample, we then regress the real-time responses at that

[18]Note that the error bands are calculated based on the marginal distribution of the differences each period. They do not directly take into account information about the difference in the proceeding and following periods (i.e. the joint distribution of the difference across periods). The bands based on the marginal distribution only take into account information about other periods in an indirect fashion since they are based on smoothed (full sample) parameter estimates. For fixed coefficient VARs with sign restrictions, issues with pointwise error bands have been highlighted by Inoue & Kilian (2013), for example. It is not clear how to extend their methods to VARs with time-varying coefficients and stochastic volatility in general and to our question at hand in particular.

point in time on the final responses at the same point in time and a constant:

$$r_t^{real,i}(0) = \alpha_t^i + \beta_t^i r_t^{final,i}(0) + u_t \tag{19}$$

Thus, each hypothetical scatterplot is summarized by two numbers, the constant α_t^i and the slope β_t^i. We focus on the contemporaneous response since the differences are largest for small horizons. Since the sample size for each regression is given by the number of draws we use to calculate the impulse responses, we do not report standard errors for the coefficients - these standard errors would be tiny. If responses based on real-time data are just a noisy version of the responses based on final data, we would expect the intercept α_t^i to be zero and the coefficient on the responses for final data β_t^i to be 1. Otherwise there is a bias in the real-time data responses relative to the responses based on final data that economists are actually interested in. Figure 8 shows how the intercept and the coefficient on the final-data response vary over time for the case of the contemporaneous response to a monetary policy shock. The gray line represents the slope of the regression β_t^i (right axis) and the red line represents the intercept α_t^i (left axis). Both paths show a similar pattern: Until 1980 there is a clear bias. After 1980 the coefficients quickly move toward values that imply no bias.

Figure 9 shows the results for the same regressions in the case of the contemporaneous response of real-time and final GDP growth. We see a broadly similar pattern for the intercept that moves toward zero after 1980. There is no substantial shift in the behavior of the slope, though. The slope is never as small as the minimum slope for inflation, but it also does not substantially move toward 1 after 1980. Real-time GDP growth responds differently than final GDP growth to a monetary policy shock on impact in systematic fashion throughout our sample. We think of these results as a cautionary tale about the information content of real-time data releases of GDP growth. It is important to remember here that we try to recover the true response of real-time data to a monetary policy shock, not the response to a monetary policy shock that can be recovered in real-time.

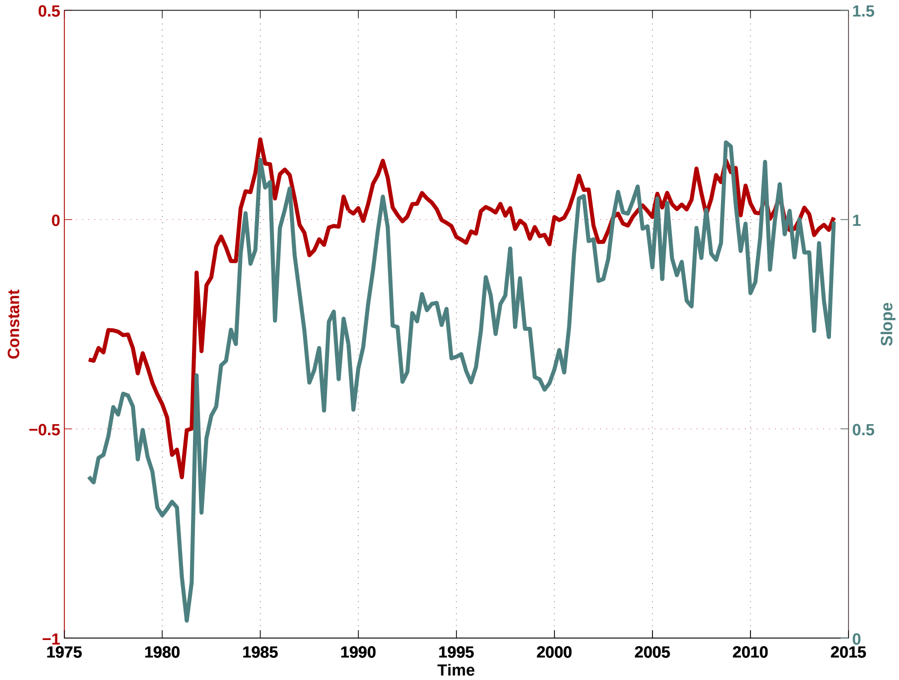

Figure 8: Relationship between inflation real-time and final data based impact impulse responses over time. Intercept α_t^π in red and slope β_t^π in gray.

6 Robustness Checks

In this section, we present additional results: (i) a model with employment growth instead of GDP growth, (ii) impulse responses identified using a recursive identification scheme, and (iii) a model with an alternative definition of final data, following Aruoba (2008).

6.1 Results for the Model With Employment Growth

Turning to a model with employment growth rather than GDP growth, we find that the forecasted measurement error in inflation is very similar across the two specifications, as shown in figure 10. The error bands for the forecast error in employment growth do not contain 0 for more periods than in the GDP growth case. The case for non-zero measurement errors is thus at least as strong for the employment growth case as for the GDP growth case. Turning to second moments in figure 11, we see

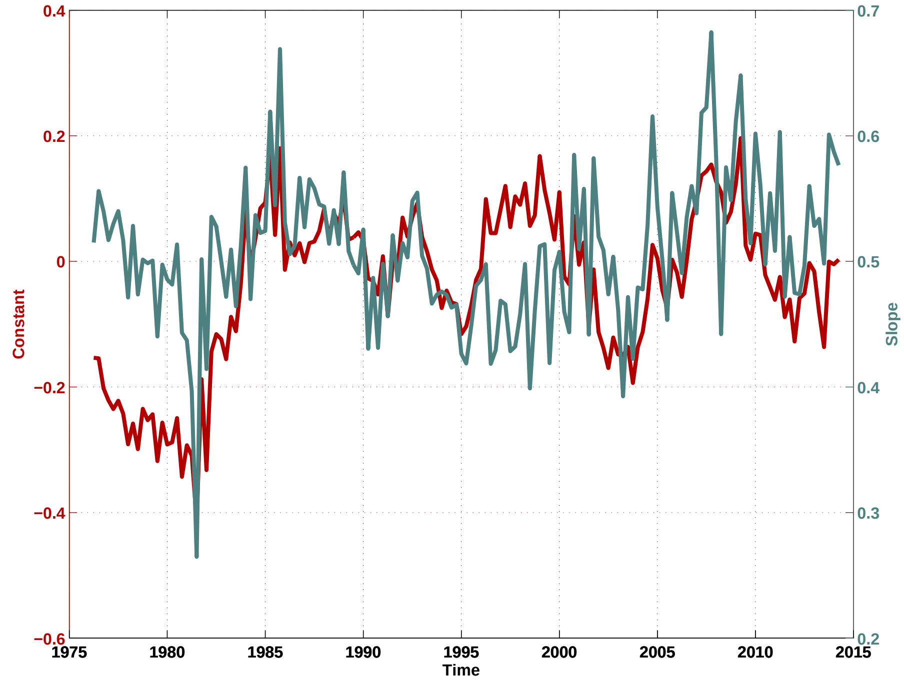

Figure 9: Relationship between GDP growth real-time and final data based impact impulse responses over time. Intercept α_t^{GDP} in red and slope β_t^{GDP} in gray.

the same pattern for the volatility of the inflation measurement error as in the case of the model with GDP growth. The evolution of volatility of the measurement error in employment growth is broadly similar to that of GDP growth. The correlation structure, on the other hand, is quite different from the case of the GDP growth VAR. There is no substantial trend in the correlation; the correlation is smaller in magnitude and is very close to zero for substantial periods of time. While the assumption of uncorrelated measurement errors is less of a problem when using employment and inflation data, our results for this case still show that the dynamics of measurement errors call for a more complex model than what is often assumed. We find that stochastic volatility and a bias in the measurement errors are present.

We next turn to studying the effects of monetary policy shocks on employment growth data. The responses of the nominal interest rate as well as real-time and final inflation are very similar to the benchmark model, so we omit them here.

For the impulse response of employment growth to a monetary policy shock, a

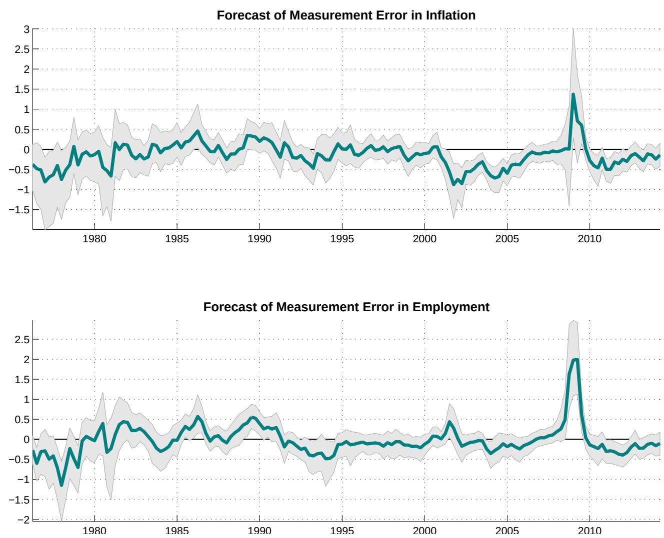

Figure 10: Forecasted measurement errors for employment growth model

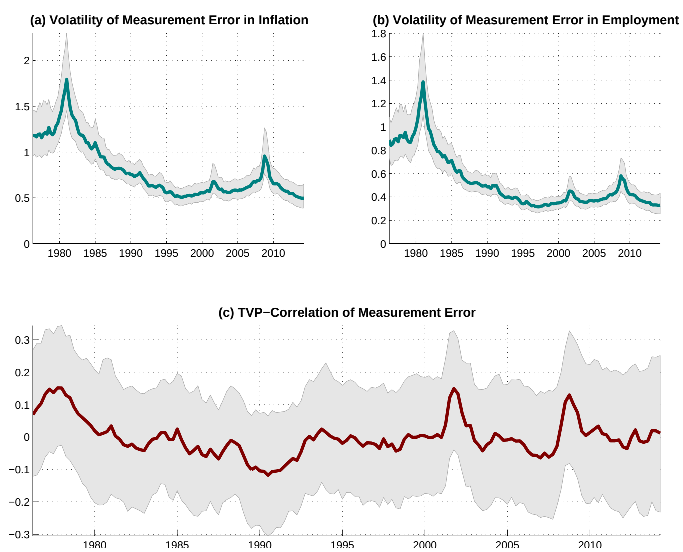

Figure 11: Volatility and correlation of forecasted measurement errors for employment growth model

qualitatively similar picture to real GDP growth emerges in figure 12. The largest differences appear on impact. Those differences are economically significant and the final data responses are larger in absolute value than those of the real-time data.

Plotting the differences between contemporaneous responses of real-time and final

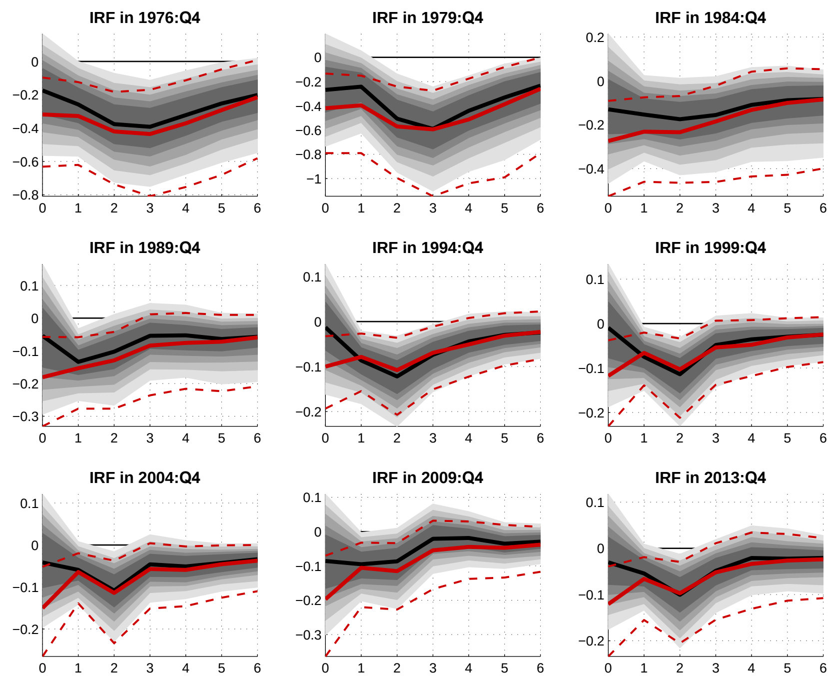

Figure 12: Impulse response functions for real time (gray/black) and final (red) employment growth to a one standard deviation monetary policy shock.

employment growth over time, a similar picture to real GDP growth emerges. Figure 13 shows that the median difference is negative throughout and the 85th percentile band is hovering around 0, which implies that there is a substantial probability at any point in time that the difference in responses is negative and that the values of that difference are economically meaningful.

Finally, we return to using regressions on draws for the real-time and final data contemporaneous responses. Since the responses to inflation turned out to be similar to those obtained using the VAR with GDP growth (with even smaller differences between real-time and final data responses), we focus on the responses of employment growth.

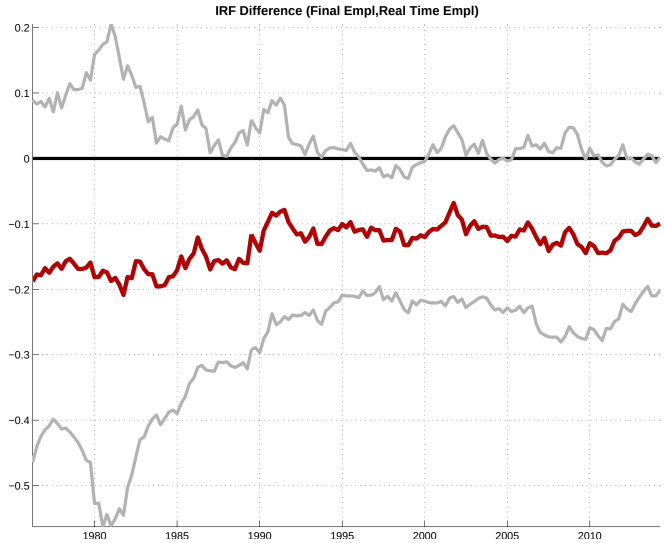

Figure 13: Differences between employment growth impulse responses on impact.

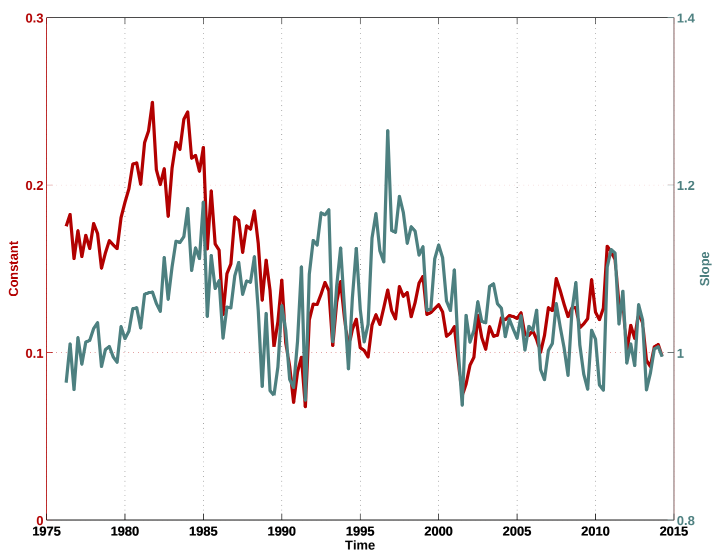

Figure 14: Relationship between real-time and final data employment growth based impact impulse responses over time. Intercept α_t^{emp} in red and slope β_t^{emp} in gray.

Figure 14 shows that the movements in the slope and intercept coefficients are smaller than for GDP growth. The slope coefficient is not too far from 1, but the intercept is always larger than 0, so there is still a bias in the relationship of real-time and final data responses on impact. There is a break in the 1980s toward less bias in the relationship between the responses, but the break happens later than in the case of the GDP growth VAR (around 1985).

Most importantly, the responses of final employment growth are larger in magnitude than those of the real-time data, just as we found in the model with GDP growth.

6.2 A Recursive Identification Scheme

The results so far all use sign restrictions to identify monetary policy shocks. To check whether or not our results are robust to other identification schemes (especially identification schemes that impose the same restriction on real-time and final data), the same exercise is carried out using a recursive identification scheme with the nominal interest rate ordered last. Conditional on reduced-form parameter estimates, the exact ordering of the other variables does not matter for the impulse response to a monetary policy shock (Christiano, Eichenbaum & Evans (1999)).[19]. Thus, this recursive identification scheme imposes the same restrictions on the responses of real-time and final data - they are ordered before the monetary policy variable. For the sake of brevity, we focus on the response to real GDP growth. It should be noted, though, that in the case of the recursive identification scheme the response of inflation displays a price puzzle, which is one reason why we prefer the sign restriction approach. Figure 15 displays the responses of real GDP growth. We still see substantial differences between real-time and final data responses (with final data responses showing some erratic behavior in 1985 and 2000). While we do not put a lot of faith in a literal interpretation of the results of this recursive identification scheme, it is nonetheless important to out that just as in the case with sign restrictions, the differences between real-time and final data responses are still

[19]The ordering of variables can in theory matter for the estimation of the reduced-form parameters in the class of models we use - see Primiceri (2005) for a discussion.

largest on impact (because of the recursive identification scheme, a monetary policy shock only impacts the other variables at horizon 1) and the response of final data is smaller on impact (except for the first two time periods shown, where the impact responses are similar). Thus, the difference on impact in the sign restriction case thus does not seem to be an artifact of imposing restrictions on only final data responses.

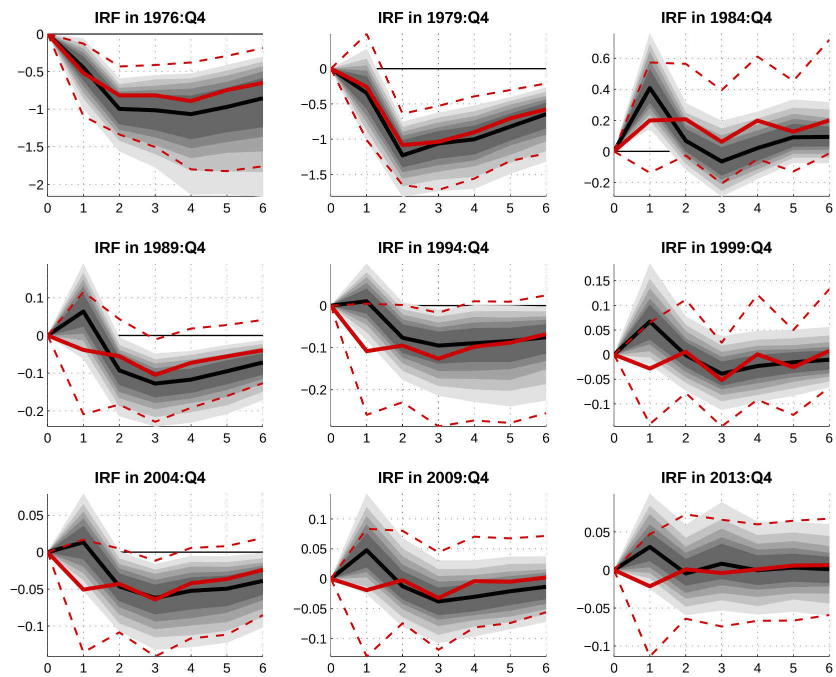

Figure 15: Impulse response functions for real-time (gray/black) and final (red) real GDP growth to a one standard deviation monetary policy shock, using the recursive identification scheme.

6.3 An Alternative Definition of Final Data

Aruoba (2008) argued for a definition of final data that uses data published after a fixed lag (roughly 3 years for most of his variables). We now follow this procedure (with a lag of 3 years) and replicate our benchmark analysis. For the sake of brevity,

we focus on the response to monetary policy shocks (the reduced form evidence on the dynamics of measurement errors is in line with the benchmark case).
The responses of both the nominal interest rate as well as both the real-time and final inflation rates are also similar to the benchmark case and are thus omitted here.

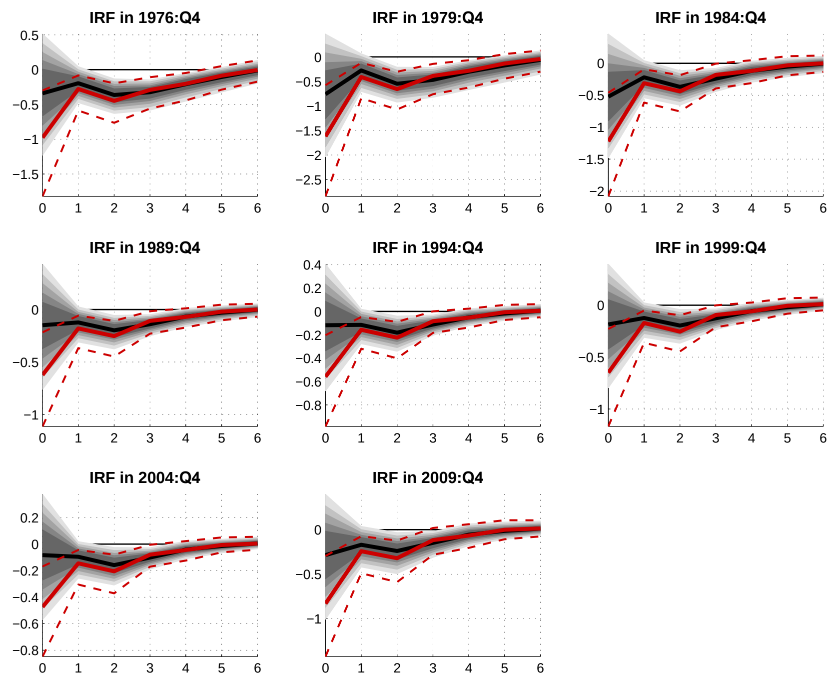

Figure 16: Impulse response functions for the real GDP growth rate to a one standard deviation monetary policy shock, using the alternative definition of final data.

Importantly, the response of real GDP growth (figure 16) shows the same bias as in the benchmark case: On impact, the response of final data is substantially larger in absolute magnitude. The differences in this case are more short-lived, but also more pronounced on impact relative to the benchmark.

7 Conclusion

Measurement errors are pervasive in real-time macroeconomic data. We extend the insights of Aruoba (2008) to incorporate time varying dynamics and document that these measurement errors feature substantial time-varying volatility, can be correlated with a time-varying correlation, and are not centered around zero. Thus, modeling real-time data as the sum of the final data and a simple independent noise process can miss important features of the data.

We show that these facts are not a curiosity, but have policy implications: (i) These differences between real-time data and final data manifest themselves in the substantially different ways that real-time and final data respond to monetary policy shocks, and (ii) the real-time responses can be substantially biased. As such, the responses of various measures of real activity are larger in magnitude for final data.

References

Amir-Ahmadi, P., Matthes, C. & Wang, M.-C. (2014), Drifts, Volatilities, and Impulse Responses Over the Last Century, Working Paper 14-10, Federal Reserve Bank of Richmond.

Aoki, K. (2003), 'On the optimal monetary policy response to noisy indicators', *Journal of Monetary Economics* **50**(3), 501–523.

Aruoba, S. B. (2008), 'Data Revisions Are Not Well Behaved', *Journal of Money, Credit and Banking* **40**(2-3), 319–340.

Benati, L. (2010), Evolving Phillips trade-off, Working Paper Series 1176, European Central Bank.

Benati, L. (2014), Economic Policy Uncertainty and the Great Recession, Working paper, University of Bern.

Benati, L. & Lubik, T. (2014), 'Sales, inventories, and real interest rates: A century of stylized facts', *Journal of Applied Econometrics* p. forthcoming.

Beveridge, S. & Nelson, C. R. (1981), 'A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'', *Journal of Monetary Economics* **7**(2), 151–174.

Canova, F. & Gambetti, L. (2009), 'Structural changes in the US economy: Is there a role for monetary policy?', *Journal of Economic Dynamics and Control* **33**(2), 477–490.

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

Carriero, A., Clements, M. & Galvao, A. B. (2015), 'Forecasting with Bayesian Multivariate vintage-based VARs', *International Journal of Forecasting* **31**, 757768.

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

Cogley, T., Matthes, C. & Sbordone, A. M. (2015), 'Optimized Taylor rules for disinflation when agents are learning', *Journal of Monetary Economics* **72**(C), 131–147.

Cogley, T. & Sargent, T. J. (2005), 'Drift and volatilities: Monetary policies and outcomes in the post WWII U.S.', *Review of Economic Dynamics* **8**(2), 262–302.

Cogley, T. & Sargent, T. J. (2014), Measuring price-level uncertainty and instability in the U.S., 1850-2012, Technical report.

Croushore, D. & Evans, C. L. (2006), 'Data revisions and the identification of monetary policy shocks', *Journal of Monetary Economics* **53**(6), 1135–1160.

Croushore, D. & Sill, K. (2014), Analyzing Data Revisions With a Dynamic Stochastic General Equilibrium Model, Working Paper 14-29, Federal Reserve Bank of Philadelphia.

Croushore, D. & Stark, T. (2001), 'A real-time data set for macroeconomists', *Journal of Econometrics* **105**(1), 111–130.

D'Agostino, A., Gambetti, L. & Giannone, D. (2013), 'Macroeconomic forecasting and structural change', *Journal of Applied Econometrics* **28**(1), 82–101.

Del Negro, M. & Primiceri, G. (2015), 'Time-varying structural vector autoregressions and monetary policy: a corrigendum', *Review of Economic Studies* .

Faust, J. (1998), 'The robustness of identified VAR conclusions about money', *Carnegie-Rochester Conference Series in Public Policy* **49**, 207–244.

Gali, J. & Gambetti, L. (2009), 'On the sources of the great moderation', *American Economic Journal: Macroeconomics* **1**(1), 26–57.

Inoue, A. & Kilian, L. (2013), 'Inference on impulse response functions in structural VAR models', *Journal of Econometrics* **177**(1), 1–13.

Jacobs, J. P. A. M., Sarferaz, S., van Norden, S. & Sturm, J.-E. (2013), Modeling Multivariate Data Revisions, CIRANO Working Papers 2013s-44, CIRANO.

Jacobs, J. P. & van Norden, S. (2011), 'Modeling data revisions: Measurement error and dynamics of true values', *Journal of Econometrics* **161**(2), 101–109.

Kishor, N. K. & Koenig, E. F. (2009), 'VAR Estimation and Forecasting When Data Are Subject to Revision', *Journal of Business & Economic Statistics* **30**(2), 181–190.

Koop, G. & Korobilis, D. (2010), Bayesian multivariate time series methods for empirical macroeconomics, Working paper, University of Strathclyde.

Lubik, T. A. & Matthes, C. (2014), Indeterminacy and Learning: An Analysis of Monetary Policy in the Great Inflation, Working Paper 14-2, Federal Reserve Bank of Richmond.

Nimark, K. (2008), 'Monetary policy with signal extraction from the bond market', *Journal of Monetary Economics* **55**(8), 1389–1400.

Orphanides, A. (2002), 'Monetary-Policy Rules and the Great Inflation', *American Economic Review* **92**(2), 115–120.

Orphanides, A. & van Norden, S. (2002), 'The Unreliability of Output-Gap Estimates in Real Time', *The Review of Economics and Statistics* **84**(4), 569–583.

Primiceri, G. (2005), 'Time varying structural vector autoregressions and monetary policy', *Review of Economic Studies* **72**(3), 821–852.

Schorfheide, F., Sill, K. & Kryshko, M. (2010), 'DSGE model-based forecasting of non-modelled variables', *International Journal of Forecasting* **26**(2), 348–373.

Sims, C. A. (2002), 'Solving Linear Rational Expectations Models', *Computational Economics* **20**(1-2), 1–20.

Svensson, L. E. O. & Woodford, M. (2004), 'Indicator variables for optimal policy under asymmetric information', *Journal of Economic Dynamics and Control* **28**(4), 661–690.

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