Federal Reserve Bank of New York
Staff Reports

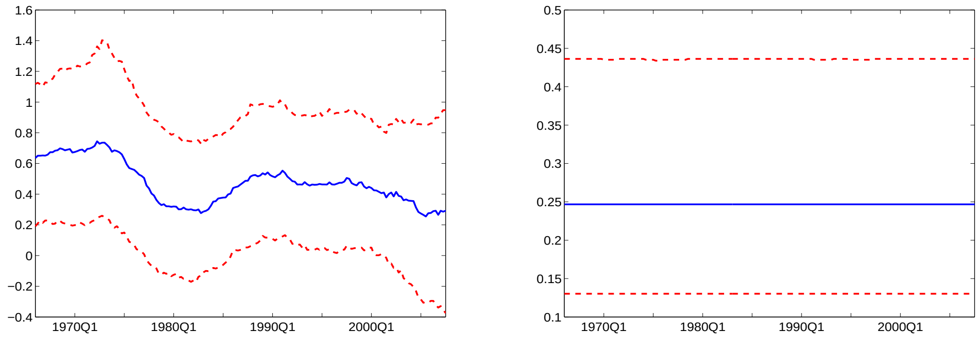
Real-Time Inflation Forecasting in a Changing World

Jan J. J. Groen
Richard Paap
Francesco Ravazzolo

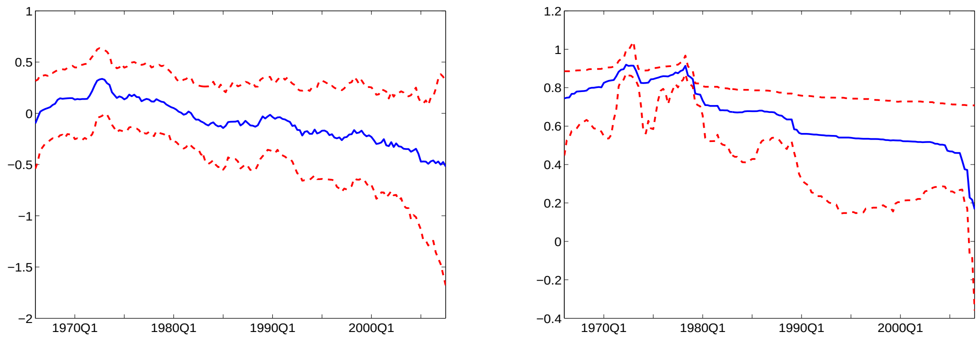
Staff Report no. 388
August 2009

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

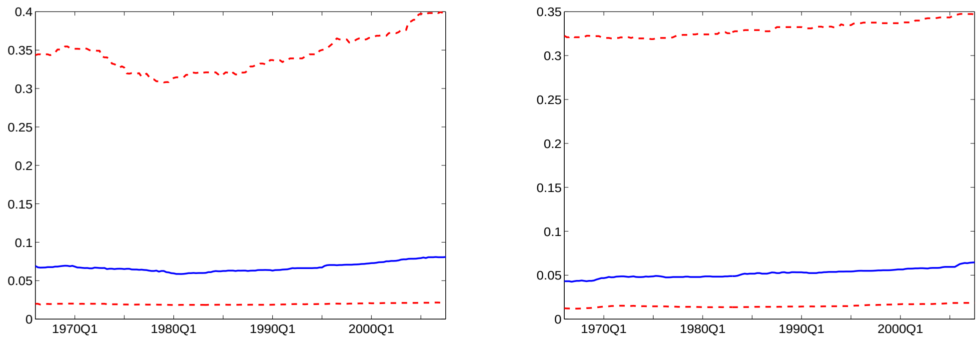
Figure 5: Posterior densities of the intercept, persistence and innovation variance in the TVP-AR model relative to BMASB for $h = 1$: PCE Deflator Inflation



(a) BMASB Intercept – TVP-AR Intercept



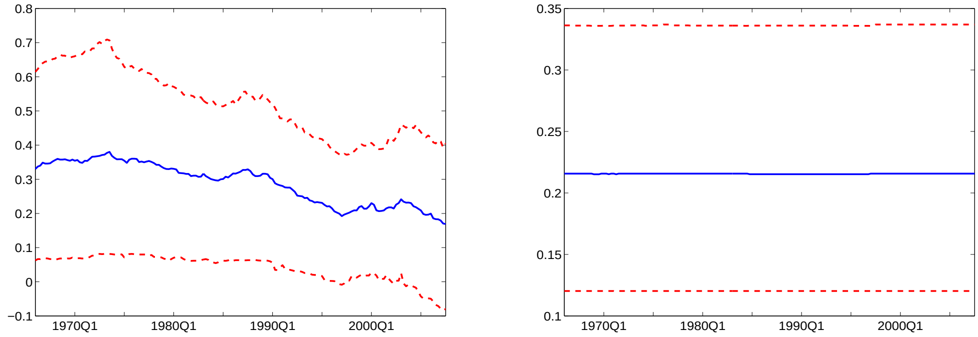
(b) BMASB Persistence – TVP-AR Persistence



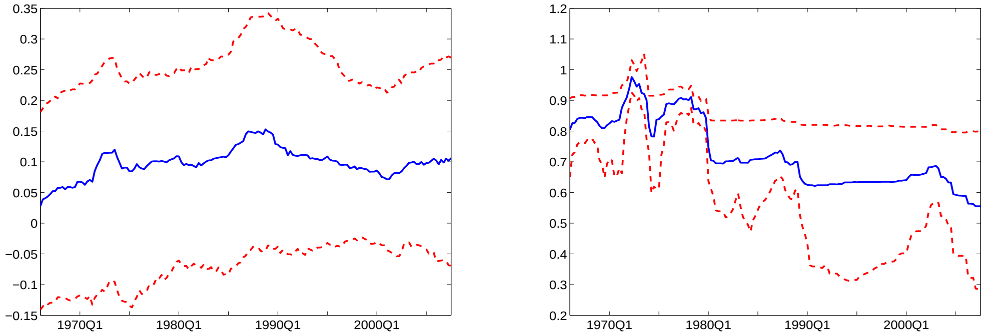
(c) BMASB σ_t^2 – TVP-AR σ_t^2

Note: The graphs in this figure show the posterior medians of the intercept, accumulated persistence and error variance in BMASB model (6) relative to the time-varying AR model (13) for PCE deflator inflation at $h = 1$. Persistence is computed by averaging the sum of the included autoregressive parameters across all model specifications using the posterior model probabilities. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

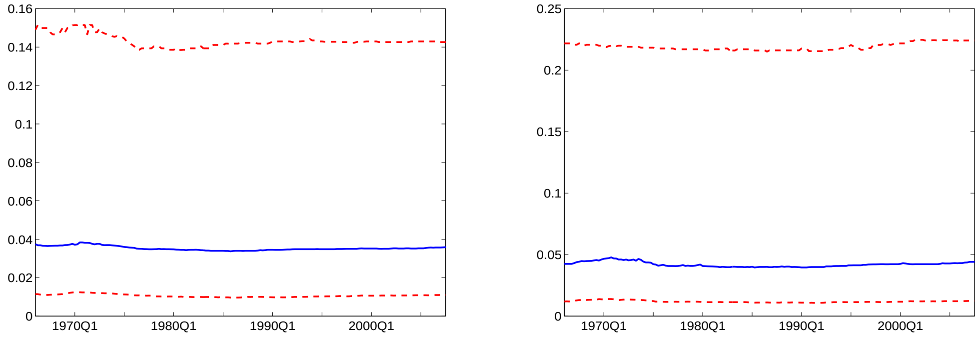
Figure 6: Posterior densities of the intercept, persistence and innovation variance in the TVP-AR model relative to BMASB for $h = 1$: GDP Deflator Inflation



(a) BMASB Intercept – TVP-AR Intercept



(b) BMASB Persistence – TVP-AR Persistence



(c) BMASB σ_t^2 – TVP-AR σ_t^2

Note: The graphs in this figure show the posterior medians of the intercept, accumulated persistence and error variance in BMASB model (6) relative to the time-varying AR model (13) for GDP deflator inflation at $h = 1$. Persistence is computed by averaging the sum of the included autoregressive parameters across all model specifications using the posterior model probabilities. The dashed lines in the graphs are the 25th and 75th percentiles of the posterior densities.

5 Real-Time Prediction of U.S. Inflation Rates

We will be focusing in this section on the out-of-sample forecasting performance of our BMASB Phillips curve model (6) relative to other, often very parsimonious, models that are frequently used for inflation forecasting. Section 5.1 provides an outline of our forecasting exercise, including a description of the alternative models. A discussion of the out-of-sample forecasting results follows in Section 5.2.

5.1 Forecasting procedure

In Section 4.2 we described the full sample developments in inflation dynamics for the U.S. PCE and GDP deflator series through the eyes of our BMASB Phillips curve model (4)–(6). However, the ultimate test for this model is how it competes with alternative specifications in a real-time, out-of-sample context. Hence, the forecasting exercise in this section.

The starting point of our forecasting exercise is the model in (4)–(6), which we have been referring to as the BMASB Phillips curve model. We use the model to obtain and evaluate one-quarter and one-year ahead forecasts for the quarter-on-quarter inflation rate of both the PCE deflator and the GDP deflator in the United States. For computational reasons we obtain the one-year ahead forecasts through direct forecasting.[9] Each forecast is based on a re-estimation of the model using an expanding window of historical data and the MCMC procedure outlined in Section 3.1. For example, suppose the first h-step ahead forecast is produced in quarter t_0 for $h = 1, 4$. As we want to evaluate the forecasts in real-time, we use the original vintage of data available at t_0 to re-estimate the BMASB Phillips curve model on the sample $t = 1, \dots, t_0$, with the forecast horizon $h = 1$ or 4. The resulting, direct, forecast using data on x_{jt} for $t = 1, \dots, t_0$ and the posterior draws from the estimation up to t_0 (see Section 3.2) is then evaluated against the vintage of inflation data that is available h quarters ahead, i.e., the vintage at $t_0 + h$. We repeat this process of re-estimation and forecast generation for $t_0 + 1, \dots, T - h$. This results in a time series of forecast errors for $t = t_0, \dots, T - h$, which we then use to compute the square root of mean squared forecast errors (RMSE).[10]

To assess how our BMASB Phillips curve model (6) performs in real-time, we need

[9]Whether an iterative procedure provides more accurate forecasts than a direct approach is a matter of ongoing debate, see the discussion in Marcellino *et al.* (2006).

[10]That is, if one defines the out-of-sample forecast error of a model for y_{t+h} as $\hat{\epsilon}_{t+h}$ then

$$\text{RMSE} = \sqrt{\frac{1}{T - t_0 - h} \sum_{s=t_0}^{T-h} \hat{\epsilon}_{s+h}^2}.$$

to compare the corresponding RMSE with those from viable alternative inflation forecast models. These alternatives include univariate models and multivariate models, where our univariate models are summarized in the first panel of Table 4. First amongst these univariate models is the random walk model, which since Atkeson and Ohanian (2001) is seen as one of the hardest models to beat when it comes to out-of-sample inflation prediction. Also, time-invariant autoregressive specifications for inflation, using lag orders between 1 and 4, are considered as parsimonious alternatives to (6).

The two models after those in Table 4 are variations on these parsimonious model specifications that incorporate time-variation in the model structure. The first of these is the TVP-AR model (13) with a maximum lag order of 4 quarters. In this specification we, firstly, allow the intercept, the autoregressive parameters as well as the error variance to break in the same manner as in our BMASB Phillips curve model and, then, construct a BMA across all possible lag order combinations. The other one is an inflation forecast model that has been successfully used by Stock and Watson (2007, 2008) to predict inflation. They propose an observed components model with stochastic volatility specifications for the unobserved component of inflation as well as the temporary deviation from it, i.e.,

$$\begin{aligned} y_t &= \beta_t + \sigma_t \varepsilon_t \\ \beta_t &= \beta_{t-1} + \omega_t \eta_t \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + u_{1t} \\ \ln \omega_t^2 &= \ln \omega_{t-1}^2 + u_{2t}, \end{aligned} \tag{14}$$

where $\varepsilon_t \sim \text{NID}(0,1)$, $\eta_t \sim \text{NID}(0,1)$ and $u_t = (u_{1t}\ u_{2t})' \sim \text{NID}(\mathbf{0}, \rho I_2)$ with ρ a scalar parameter controlling the smoothness of the stochastic volatility processes, and where ε_t, η_t and u_t are independent. We follow Stock and Watson (2007) and set in (14) $\rho = 0.04$; Stock and Watson (2007) motivate their choice for ρ based on the fit of (14) for U.S. inflation rates over the 1955-2004 sample.

The remaining models in Table 4 all incorporate information from a range of additional regressors. These encompass a simple linear regression of the quarter-to-quarter inflation rate h quarters ahead on all 14 predictor variables described in the previous section plus four inflation lags, that is,

$$y_{t+h} = X_t' \beta + \sigma \varepsilon_t, \tag{15}$$

where $X_t = (1, x_{1t}, \ldots, x_{kt})'$ and $\varepsilon_t \sim \text{NID}(0,1)$. To deal with the curse of dimensionality in such a regression, we also estimate the β parameter in this regression using a ridge regression (shrinkage) estimator:

$$\hat{\beta} = \left(\sum_{t=1}^{T-h} X_t X_t' + \lambda I \right)^{-1} \left(\sum_{t=1}^{T-h} X_t y_{t+h}' \right) \tag{16}$$

and the scalar shrinkage parameter λ. For the latter we choose $\lambda = 10$, as De Mol *et al.* (2006) show that the degree of shrinkage should be proportional to the number of regressors to achieve the best forecasting performance in a data-rich context.

Further, we employ a version of our BMASB Phillips curve model *without* time-varying parameters and variance, i.e, (2) where we construct a Bayesian model average (BMA) across the different selected regressor combinations. Finally, we consider a bivariate VAR model of the inflation rate and real output growth as well as a Bayesian model average across all possible bivariate VAR models of inflation and one of our 14 economic predictor variables.

5.2 Out-of-Sample Results

All of the models discussed in Section 5.1 are used to generate quarter-on-quarter PCE deflator and GDP deflator inflation forecasts one-quarter ahead ($h = 1$) and one-year ahead ($h = 4$). These are evaluated by computing the corresponding RMSEs across three periods: 1980Q1-2008Q4, 1980Q1-1994Q4 and 1995Q1-2008Q4. The evaluation samples span a number of large events that potentially could have caused time-variation in the dynamics of inflation rates, like, e.g., the 'monetarist experiment' by the Federal Reserve under Volcker, the 'Great Moderation' in the mid-1980s and the 9/11 catastrophe in 2001. The forecasts of these models are based on posterior results of the model parameters and, if relevant, the latent variables computed using an expanding window of data, starting with 1960Q1-1979Q4 based on the original data vintages starting from 1979Q4. Finally, the resulting RMSEs based on the corresponding forecast errors are used to compute RMSE ratios relative to our BMASB Phillips curve model to asses how well, or not, they are doing in a real-time out-of-sample setting *vis-à-vis* our model, where a ratio smaller than 1 indicates that a model outperforms (4)–(6) and *vice versa*.

Tables 5 and 6 report in the first line the RMSEs for our BMASB Phillips curve model-based forecasts in case of the PCE and the GDP deflator inflation measures, respectively. Below that line, both tables report the ratio of the RMSE for each of the competing models as discussed in Table 4 relative to the RMSE of our BMASB Phillips curve model. When we focus on PCE deflator inflation first, see Table 5, it becomes quite striking how successful the BMASB Phillips curve forecasts are in comparison with the other models. Over the full 1980-2008 evaluation sample and the first sub-sample none of these can beat our Phillips curve specification (6) at the one-quarter and one-year ahead forecast. In the final sub-sample, only the USCV Stock and Watson (2007) model performs better at both forecast horizons although the difference in performance at one-quarter ahead forecasting is very small. The purely autoregressive and random walk specifications are not performing

Table 1: Marginal posterior probabilities of predictor variable selection

	PCE Deflator Inflation		GDP Deflator Inflation	
	$h = 1$	$h = 4$	$h = 1$	$h = 4$
INFL_t	0.095	0.000	0.000	0.106
INFL_{t-1}	0.638	0.441	0.063	0.194
INFL_{t-2}	0.103	0.132	0.236	0.131
INFL_{t-3}	0.349	0.192	0.225	0.000
ROUTP_t	0.263	0.559	0.000	0.000
RCONS_t	0.125	0.446	0.155	0.357
RINVR_t	0.273	0.397	0.229	0.275
PIMP_t	0.078	0.000	0.000	0.000
UNEMPL_t	0.270	0.288	0.197	0.202
HSTS_t	0.107	0.330	0.289	0.265
NFPR_t	0.000	0.771	0.358	0.113
OIL_t	0.286	0.048	0.367	0.158
FOOD_t	0.409	0.346	0.366	0.218
RAW_t	0.526	0.181	0.368	0.292
M2_t	0.347	0.202	0.005	0.272
TS_t	0.047	0.145	0.000	0.314
YL_t	0.000	0.000	0.154	0.033
MS_t	0.948	0.898	1.000	1.000

Note: The table presents the marginal posterior inclusion probabilities in the predictive regression model (6) for $h = 1$ and $h = 4$ over the full sample, 1960Q1 – 2008Q4.

Variable mnemonics: INFL - PCE or GDP Deflator inflation; ROUTP - percentage quarterly change real GDP; RCONS- percentage quarterly change real personal consumption expenditures; RINVR - percentage quarterly change real residential investment; PIMP - percentage quarterly change import price deflator; UNEMPL - unemployment rate (% labor force); HSTS - log level housing starts; NFPR - percentage quarterly change non-farm payrolls; OIL - percentage quarterly change real oil spot price; FOOD - percentage quarterly change real food commodities price index; RAW - percentage quarterly change real raw materials price index; M2 - percentage quarterly change M2 monetary aggregate; TS - slope term structure level; YL - level term structure factor; MS - one-year ahead inflation expectations from the Michigan Consumer survey.

Table 2: Posterior model probabilities: PCE deflator inflation

Model	Prob. %
Forecast horizon: $h = 1$	
$\text{INFL}_{t-1}, \text{HSTS}_t, \text{TS}_t, \text{RAW}_t, \text{MS}_t$	2.27
$\text{INFL}_{t-1}, \text{RAW}_t, \text{MS}_t$	1.87
$\text{INFL}_{t-1}, \text{RAW}_t$	1.53
$\text{INFL}_{t-1}, \text{INFL}_{t-3}, \text{RINVR}_t, \text{FOOD}_t, \text{RAW}_t, \text{MS}_t$	1.47
$\text{INFL}_{t-1}, \text{INFL}_{t-3}, \text{FOOD}_t, \text{RAW}_t, \text{MS}_t$	1.40
$\text{INFL}_t, \text{INFL}_{t-1}, \text{M2}_t, \text{OIL}_t, \text{FOOD}_t, \text{RAW}_t, \text{MS}_t$	1.40
$\text{INFL}_t, \text{INFL}_{t-1}, \text{PIMP}_t, \text{M2}_t, \text{OIL}_t, \text{FOOD}_t, \text{MS}_t$	1.33
$\text{INFL}_{t-2}, \text{INFL}_{t-3}, \text{ROUTP}_t, \text{OIL}_t, \text{FOOD}_t, \text{MS}_t$	1.27
$\text{INFL}_{t-1}, \text{INFL}_{t-3}, \text{RCONS}_t, \text{UNEMPL}_t, \text{MS}_t$	1.20
$\text{INFL}_{t-2}, \text{INFL}_{t-3}, \text{OIL}_t, \text{MS}_t$	1.20
Forecast horizon: $h = 4$	
$\text{INFL}_{t-3}, \text{ROUTP}_t, \text{RCONS}_t, \text{RINVR}_t, \text{HSTS}_t, \text{NFPR}_t, \text{MS}_t$	3.70
$\text{INFL}_{t-1}, \text{ROUTP}_t, \text{RINVR}_t, \text{NFPR}_t$	3.30
$\text{INFL}_{t-1}\ \text{ROUTP}_t, \text{RCONS}_t, \text{NFPR}_t, \text{MS}_t$	3.20
$\text{INFL}_{t-1}, \text{ROUTP}_t, \text{UNEMPL}_t, \text{NFPR}_t, \text{MS}_t$	2.10
$\text{INFL}_{t-1}, \text{RCONS}_t, \text{UNEMPL}_t, \text{NFPR}_t, \text{FOOD}_t, \text{MS}_t$	2.10
$\text{INFL}_{t-2}, \text{INFL}_{t-3}, \text{ROUTP}_t, \text{RCONS}_t, \text{HSTS}_t, \text{NFPR}_t, \text{TS}_t, \text{FOOD}_t, \text{MS}_t$	2.00
$\text{INFL}_{t-2}, \text{ROUTP}_t, \text{RINVR}_t, \text{UNEMPL}_t, \text{NFPR}_t, \text{MS}_t$	1.80
$\text{INFL}_{t-3}, \text{ROUTP}_t, \text{NFPR}_t, \text{TS}_t, \text{MS}_t$	1.80
$\text{INFL}_{t-3}, \text{ROUTP}_t, \text{NFPR}_t, \text{TS}_t, \text{FOOD}_t, \text{MS}_t$	1.80
$\text{ROUTP}_t, \text{RINVR}_t, \text{HSTS}_t, \text{OIL}_t, \text{FOOD}_t, \text{MS}_t$	1.80

Note: The table lists the ten models with the highest posterior probabilities and their posterior probabilities (%) for the quarterly PCE series, 1960Q1 – 2008Q4. See Table 1 for a description of the predictor variables.

Table 3: Posterior model probabilities: GDP deflator inflation

Model	Prob. %
Forecast horizon: $h = 1$	
$\text{INFL}_{t-3}, \text{ROUTP}_t, \text{RCONS}_t, \text{MS}_t$	4.00
$\text{NFPR}_t, \text{OIL}_t, \text{MS}_t$	3.13
$\text{HSTS}_t, \text{FOOD}_t, \text{MS}_t$	2.53
$\text{FOOD}_t, \text{RAW}_t, \text{MS}_t$	2.40
$\text{NFPR}_t, \text{MS}_t$	2.33
$\text{HSTS}_t, \text{MS}_t$	2.00
$\text{HSTS}_t, \text{NFPR}_t, \text{MS}_t$	1.73
$\text{FOOD}_t, \text{RAW}_t, \text{MS}_t$	1.47
$\text{OIL}_t, \text{MS}_t$	1.40
$\text{UNEMPL}_t, \text{MS}_t$	1.40
Forecast horizon: $h = 4$	
$\text{RINVR}_t, \text{MS}_t$	5.30
MS_t	3.70
$\text{OIL}_t, \text{MS}_t$	3.70
$\text{RINVR}_t, \text{TS}_t, \text{MS}_t$	3.00
$\text{INFL}_{t-1}, \text{INFL}_{t-2}, \text{RAW}_t, \text{MS}_t$	3.00
$\text{HSTS}_t, \text{MS}_t$	2.70
$\text{INFL}_{t-1}, \text{INFL}_{t-2}, \text{M2}_t, \text{MS}_t$	2.00
$\text{RCONS}_t, \text{UNEMPL}_t, \text{HSTS}_t, \text{TS}_t, \text{MS}_t$	1.80
$\text{INFL}_{t-1}, \text{MS}_t$	1.70
$\text{INFL}_{t-1}, \text{INFL}_{t-2}, \text{MS}_t$	1.70

Note: The table lists the ten models with the highest posterior probabilities and their probabilities (%) for quarterly GDP deflator series, 1960Q1 – 2008Q4. See Table 1 for a description of the predictor variables.

Table 4: Alternative univariate and multivariate models for forecasting inflation

name	description	specification
	univariate models	
RW	Random walk	$y_t = y_{t-1} + \varepsilon_t$
AR(1)	Autoregressive model of order 1	$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t$
AR(2)	Autoregressive model of order 2	$y_t = \mu + \sum_{i=1}^{2} \phi_i y_{t-i} + \varepsilon_t$
AR(3)	Autoregressive model of order 3	$y_t = \mu + \sum_{i=1}^{3} \phi_i y_{t-i} + \varepsilon_t$
AR(4)	Autoregressive model of order 2	$y_t = \mu + \sum_{i=1}^{4} \phi_i y_{t-i} + \varepsilon_t$
TVP-AR(4)	AR(4) with structural instability and BMA	see (13)
UCSV	Unobserved component model with SV	see (14)
Linear	Linear regression with all predictors	$y_{t+h} = X_t'\beta + \sigma\varepsilon_t$
Ridge	Ridge regression with $\lambda = 10$	see (16)
BMA	Bayesian model averaging	(6) with $\beta_t = \beta$, $\sigma_t = \sigma$ $\forall t$
	multivariate models	
VAR(4)	Bivariate VAR(4) with inflation & output growth	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-i} + \epsilon_t$
BMA-VAR(4)	BMA of all possible bivariate VAR(4)	$Y_t = \mu + \sum_{i=1}^{4} \Phi_i Y_{t-1} + \epsilon_t$

Table 5: RMSE - PCE

	Horizon: $h = 1$			Horizon: $h = 4$		
	F	I	II	F	I	II
BMASB	0.39	0.37	0.41	0.43	0.42	0.44
			univariate models			
RW	1.23	1.26	1.18	1.20	1.29	1.12
AR(1)	1.19	1.20	1.17	1.12	1.14	1.10
AR(2)	1.15	1.17	1.14	1.10	1.14	1.07
AR(3)	1.10	1.12	1.09	1.10	1.13	1.07
AR(4)	1.10	1.12	1.09	1.10	1.13	1.07
TVP-AR(4)	1.14	1.16	1.12	1.20	1.29	1.11
UCSV	1.07	1.17	**0.99**	1.07	1.22	**0.95**
Linear	1.04	1.09	1.00	1.18	1.28	1.08
Ridge	1.06	1.06	1.06	1.03	1.06	1.00
BMA	1.05	1.07	1.04	1.26	1.41	1.09
			multivariate models			
VAR(4)	1.08	1.10	1.07	1.05	1.03	1.07
VAR-BMA(4)	1.05	1.05	1.06	1.11	1.20	1.01

Note: The table presents root mean square prediction error (RMSE) of the BMASB Phillips curve-type model (6), the first line, as well as RMSE ratios relative to it for different univariate and multivariate models, see Table 4, for the full 1980Q1-2008Q4 evaluation sample (F) and two subsamples (I: 1980Q1-1994Q4, II: 1995Q1-2008Q4) at one-quarter ($h = 1$) and one-year ($h = 4$) ahead forecasting horizons for inflation. **Bold** indicates when our BMASB Phillips curve forecasts are outperformed by any of the forecasts from the competing models.

Table 6: RMSE - GDP deflator

	Horizon: $h = 1$			Horizon: $h = 4$		
	F	I	II	F	I	II
BMASB	0.27	0.31	0.22	0.32	0.36	0.27
			univariate models			
RW	1.25	1.19	1.36	1.10	1.11	1.08
AR(1)	1.21	1.15	1.34	1.10	1.09	1.13
AR(2)	1.15	1.11	1.23	1.09	1.09	1.10
AR(3)	1.08	1.04	1.17	1.09	1.10	1.09
AR(4)	1.09	1.04	1.18	1.09	1.11	1.07
TVP-AR(4)	1.11	1.09	1.14	1.22	1.09	1.41
UCSV	1.11	1.16	1.03	1.10	1.23	**0.90**
Linear	1.09	1.04	1.19	1.23	1.26	1.18
Ridge	1.04	1.01	1.12	1.01	1.04	**0.97**
BMA	1.13	1.08	1.23	1.30	1.36	1.20
			multivariate models			
VAR(4)	1.07	1.05	1.13	1.02	1.02	1.02
VAR-BMA(4)	1.14	1.11	1.21	1.16	1.17	1.13

Note: See the notes for Table 5.

well, as our model clearly outperforms these in terms of RMSE. Furthermore, we see that models containing explanatory variables perform in general better than models which only use lagged inflation information for prediction.

The results for GDP deflator inflation are quite similar to those for PCE deflator inflation, see Table 6. BMASB Phillips curve model (4)–(6) is only outperformed by two model specification in the final sub-sample for one-year ahead forecasts. These two model specifications are the Ridge estimator approach and again the USCV model.

The general conclusion from Tables 5 and 6 is that the BMASB Phillips curve-type model (4)–(6) does really well for predicting different inflation series at different forecasts horizons. Only in the sample 1995-2008 the model is outperformed by the USCV specification of Stock and Watson (2007) at one-year ahead forecasting but in the sample 1980-1994 the BMASB Phillips curve-type model performs clearly better than this model. Therefore, our BMASB Phillips curve-type specification does capture very well the time-variation in both the correlation between inflation and activity measures (the 'Phillips curve correlation') as well as inflation dynamics itself. Several studies have shown that by 'sucking' in a lot of data in an efficient way, model averaging and ridge regression can be simple and effective ways to face future instability of unknown form. Our forecasting results, however, indicate that accounting for structural instability may improve forecast performance. The BMASB model (4)–(6), which allows simultaneously for model uncertainty and structural instability, overall has the best out-of-sample performance, stressing the roles of both kinds of uncertainty.

6 Conclusion

Forecasting inflation has become much more difficult over the last decades. As a consequence, Phillips curve forecasts, i.e., inflation forecasts using an economic activity variable, have not fared well in several empirical studies and hardly ever improve upon simple univariate forecasts. Nonetheless, Phillips curve-type of relationships remain the backbone of many macroeconomic models and are important to understand policy discussions about the business cycle and inflation.

The failure of Phillips curve forecasts has several sources. Firstly, there is uncertainty about which set of activity measures best describes the Phillips correlation at a particular time. Also, inflation dynamics have changed over time resulting in breaks in the mean and variance of inflation, which in return would have caused breaks in Phillips curve-type relationships. In this paper we have introduced a generalized, reduced form Phillips curve-type model that attempts to incorporate uncertainty about the above two elements. It allows for uncertainty in the inclusion of relevant predictor variables (model uncertainty), the estimation uncertainty in the model parameters (parameter uncertainty) and finally the stability in the value of the model parameters (structural instability).

We apply our approach to model and forecast PCE and GDP deflator inflation in the U.S. between 1960 and 2008, where the forecasts are for two forecast horizons, one-quarter ahead and one-year ahead. When we use our framework to model the post-WWII inflation dynamics in the U.S. we do find some interesting empirical facts. First, over the period 1960-2008 several structural breaks occurred in the relationship between US inflation and predictor variables which include its own lags, real activity and cost measures, and other macroeconomic indicators. These changes appear to coincide with important events such as the oil crises in the 1970s, changes in the monetary policy regime, and the economic recession at the beginning of 1990s. Next, we find less evidence for exogenous breaks in the variance of inflation than what usually is found in the literature. And by conditioning on a vast range of potential combinations of activity measures, our framework finds substantially lower degrees of, time-varying, persistence in the inflation deviations from its mean than in other studies.

Finally, we find that allowing for model uncertainty and structural breaks at the same time results in superior inflation forecasts. Our Phillips curve-type specification provides very accurate forecasts of U.S. inflation for the 1980-2008 period compared to a set of competing linear models and nonlinear models including the random walk. Only in the latter half of our forecast evaluation period, i.e., 1995-2008, the UCSV model of Stock and Watson (2007) seems to be a good alternative for one-year ahead forecasting.

References

Ang, A., G. Bekaert and M. Wei, 2007, Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?, *Journal of Monetary Economics* **54**, 1163–1212.

Ang, A., M. Piazzesi and M. Wei, 2006, What Does the Yield Curve Tell us about GDP Growth?, *Journal of Econometrics* **131**, 359–403.

Atkeson, A. and L. E. Ohanian, 2001, Are Phillips Curves Useful for Forecasting Inflation?, *Federal Reserve Bank of Minneapolis Quarterly Review* **25**, 2–11.

Benati, L., 2004, Evolving Post-World War II U.K. Economic Performance, *Journal of Money, Credit, and Banking* **36**, 691–718.

Carter, C. and R. Kohn, 1994, On Gibbs Sampling for State-Space Models, *Biometrika* **81**, 541–553.

Carter, C. and R. Kohn, 1997, Semiparametric Bayesian Inference for Time Series with Mixed Spectra, *Journal of the Royal Statistical Society, Series B* **B**, 255–268.

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

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

Clark, T. and M. W. McCracken, 2008, Averaging Forecasts from VARs with Uncertainty Instabilities, *Journal of Applied Econometrics* . Forthcoming.

Cogley, T. and A. M. Sbordone, 2008, Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve, *American Economic Review* **98**, 2101–2126.

Cogley, T., G. E. Primiceri and T. J. Sargent, 2009, Inflation-Gap Persistence in the U.S., *American Economic Journal: Macroeconomics* . Forthcoming.

Cogley, T. and T. J. Sargent, 2002, Evolving Post World War II U.S. Inflation Dynamics, *in* B. S. Bernanke and K. Rogoff (editors), *NBER Macroeconomics Annual 2001*, MIT Press, Cambridge, U.S.

Cogley, T. and T. J. Sargent, 2005, Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S., *Review of Economic Dynamics* **8**, 262–302.

De Mol, C., D. Giannone and L. Reichlin, 2006, Forecasting with a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?, *Discussion Paper 5829*, Centre for Economic Policy Research.

Diebold, F. X., G. D. Rudebusch and B. Aruoba, 2006, The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach, *Journal of Econometrics* **131**, 309–338.

Estrella, A. and G. A. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance* **46**, 555–576.

Fama, E. F. and R. R. Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* **77**, 680–692.

Galí, J. and M. Gertler, 1999, Inflation Dynamics: A Structural Econometric Analysis, *Journal of Monetary Economics* **44**, 195–222.

Geman, S. and D. Geman, 1984, Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**, 721–741.

George, E. I. and R. E. McCulloch, 1993, Variable Selection Via Gibbs Sampling, *Journal of the American Statistical Association* **88**, 881–889.

Gerlach, R., C. Carter and R. Kohn, 2000, Efficient Bayesian Inference for Dynamic Mixture Models, *Journal of the American Statistical Association* **95**, 819–828.

Giordani, P. and R. Kohn, 2007, Efficient Bayesian Inference for Multiple Change-Point and Mixture Innovation Models, *Sveriges Riksbank Working paper* .

Giordani, P., R. Kohn and D. van Dijk, 2007, A Unified Approach to Nonlinearity, Outliers and Structural Breaks, *Journal of Econometrics* **137**, 112–137.

Gordon, R. J., 1997, The Time-Varying NAIRU and its Implications for Economic Policy, *Journal of Economic Perspectives* **11**, 11–32.

Groen, J. J. J. and H. Mumtaz, 2008, Investigating the Structural Stability of the Phillips Curve Relationship, *Working Paper 350*, Bank of England.

Haldane, A. and D. Quah, 1999, UK Phillips Curves and Monetary Policy, *Journal of Monetary Economics* **44**, 259–278.

Kim, S., S. Shephard and S. Chib, 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies* **65**, 361–393.

Koop, G. and S. Potter, 2007, Estimation and forecasting in models with multiple breaks, *Review of Economic Studies* pp. 763–789.

Kuo, L. and B. Mallick, 1998, Variable Selection for Regression Models, *The Indian Journal of Statistics* **60**(Series B), 65–81.

Levin, A. T. and J. M. Piger, 2004, Is Inflation Persistence Intrinsic in Industrial Economies?, *Working Paper 334*, European Central Bank.

Marcellino, M., J. H. Stock and M. W. Watson, 2006, A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series, *Journal of Econometrics* **127**(1-2), 499–526.

Pesaran, M.H., D. Pettenuzzo and A. Timmermann, 2006, Forecasting Time Series Subject to Multiple Structural Breaks, *Review of Economic Studies* **73(4)**, 1057–1084.

Sbordone, A. M., 2002, Prices and Unit Labor Costs: A New Test of Price Stickiness, *Journal of Monetary Economics* **49**, 265–292.

Sensier, M. and D. J. C. van Dijk, 2004, Testing for Volatility Changes in U.S. Macroeconomic Time Series, *Review of Economics and Statistics* **86**, 833–839.

Shephard, N.E., 1994, Partial non-Gaussian state-space models, *Biometrika* **81**, 115–131.

Sims, C. A. and T. Zha, 2006, Were there Regime Switches in U.S. Monetary Policy?, *American Economic Review* **96**, 54–81.

Stock, J.H. and M.W. Watson, 1999, Forecasting Inflation, *Journal of Monetary Economics* **44**, 293–335.

Stock, J.H. and M.W. Watson, 2007, Has Inflation Become Harder to Forecast?, *Journal of Money, Credit and Banking* **39**, 3–33.

Stock, J.H. and M.W. Watson, 2008, Phillips Curve Inflation Forecasts, *Working Paper 14322*, National Bureau of Economic Research.

Tanner, M.A. and W.H. Wong, 1987, The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association* **82**, 528–550.

Wright, J. H., 2003, Forecasting U.S. Inflation by Bayesian Model Averaging, *International Finance Discussion Paper 780*, Board of Governors of the Federal Reserve System.

Appendices

A Gibbs Sampling Algorithm

Following the scheme in Section 3.1, the Gibbs sampler for BMASB Phillips Curve model (4)-(6) sequentially goes through the following steps:

Step 1: Sampling the variable selection parameters in D

We follow Kuo and Mallick (1998), which is a simplified version of the George and McCulloch (1993) algorithm. Starting from the previous iteration, the variable D is drawn from its full conditional posterior distribution. We compute the value of the posterior density (11) for $\delta_j = 0$ and $\delta_j = 1$ given the value of the other parameters which results in p_{j0} and p_{j1}, respectively. The full conditional posterior is then given by

$$\Pr[\delta_j = 1|\bar{\theta}, B, S, K, D_{-j}, y, x] = \frac{p_{j1}}{p_{j0} + p_{j1}}, \tag{A.1}$$

for $j = 1, \ldots, k$, where $D_{-j} = (\delta_1, \ldots, \delta_{j-1}, \delta_{j+1}, \ldots, \delta_k)'$.

Step 2: Sampling K_β

The (occasional) structural breaks in the regression parameters B, measured by the latent variable κ_{jt}, are drawn using the algorithm of Gerlach *et al.* (2000, Section 3), which derives its efficiency from generating κ_{jt} without conditioning on the states β_{jt}. The conditional posterior density for κ_{jt}, $t = 1, \ldots, T$, $j = 0, \ldots, k$ unconditional on B is

$$\begin{aligned} & p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, y, x) \\ & \quad \propto p(y|K, S, \theta, x) p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x) \\ & \quad \propto p(y_{t+h+1}, \ldots, y_{T-h}|y_{h+1}, \ldots, y_{t+h}, K, S, \theta, x) \\ & \qquad\quad p(y_{t+h}|y_{h+1}, \ldots, y_{t+h-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, \theta, x) p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x), \end{aligned} \tag{A.2}$$

where $K_{\beta,-t} = \{\{\kappa_{js}\}_{j=0}^{k}\}_{s=1, s\neq t}^{T-h}$. The density $p(\kappa_{0t}, \ldots, \kappa_{kt}|K_{\beta,-t}, K_\sigma, S, \theta, x)$ is equal to $\prod_{j=0}^{k} \pi_j^{\kappa_{jt}}(1-\pi_j)^{1-\kappa_{jt}}$ since κ_{jt} does not depend on δ_j. The two remaining densities $p(y_{t+h+1}, \ldots, y_{T-h}|y_{h+1}, \ldots, y_{t+h}, K, S, \theta, x)$ and $p(y_{t+h}|y_{h+1}, \ldots, y_{t+h-1}, \kappa_1, \ldots, \kappa_t, K_\sigma, S, \theta, x)$ can easily be evaluated as shown in Gerlach *et al.* (2000, Section 3). Because κ_t can take a finite number of values, the integrating constant can easily be computed by normalization.

Step 3: Sampling the regression parameters in B

The full conditional posterior density for the latent regression parameters B is computed using a simulation smoother. We follow Carter and Kohn (1994). The Kalman smoother is applied to derive the conditional mean and variance of the latent factors; for the initial value a multivariate normal prior with mean 0 is chosen. Note that in case the variable x_j is not selected, the full conditional distributions of κ_{jt} and β_{jt} for $t = 1, \dots, T-h$ do not depend on the data y and x. Hence, in this case we sample unconditionally from (4) and (5).

Steps 4 and 5: Sampling the variance parameters K_σ and S

To draw K_σ and S we want to follow a similar approach as above. As the model for $\ln \sigma_t^2$ does not result in a linear state space model the Kalman filter cannot be applied. Therefore, we apply the approach of Giordani and Kohn (2007) and rewrite the model (5)–(6) as

$$\begin{aligned} \ln(y_{t+h} - \beta_{0t} - \sum_{j=1}^{k} \delta_j \beta_{jt} x_{jt})^2 &= \ln \sigma_t^2 + u_t \\ \ln \sigma_t^2 &= \ln \sigma_{t-1}^2 + \kappa_{k+1,t} \eta_{k+1,t}, \end{aligned} \quad \text{(A.3)}$$

where $u_t = \ln \varepsilon_t^2$ has a log χ^2 distribution with 1 degree of freedom. We follow Carter and Kohn (1994, 1997), Shephard (1994) and Kim *et al.* (1998) and approximate the $\ln \chi^2(1)$ distribution by a finite mixture of normal distributions. We consider a mixture of five normal distributions such that the density of u_t is given by

$$f(u_t) = \sum_{s=1}^{5} \varphi_s \frac{1}{\omega_s} \phi((u_t - \mu_s)/\omega_s) \quad \text{(A.4)}$$

with $\sum_{s=1}^{5} \varphi_s = 1$. The appropriate values for μ_s, ω_s^2 and φ_s can be found in Carter and Kohn (1997, Table 1). In each step of the Gibbs sampler we simulate a component of the mixture distribution from the distribution of the mixing distribution. Given the value of the mixture component we can apply standard Kalman filter techniques. Hence, the variables K_σ and S can be sampled in a similar way as K_β and B in step 2 and 3.

Step 6: Sampling $\bar{\theta}$

Finally, to sample the parameters $\bar{\theta}$ we can use standard results in Bayesian inference. Hence, the probabilities π_j are sampled from Beta distributions and the variance parameters q_j^2 are sampled from inverted Gamma-2 distributions.

B Data Sources and Construction

Inflation rates

Our two dependent variables are inflation rates based on the gross domestic product (GDP) deflator as well as the personal consumption expenditures (PCE) deflator. Both measures get revised on a regular basis and we therefore do not retrieve our data from the usual data sources. Instead, we get the original vintages of the underlying data from the 'Real-Time Data Set for Macroeconomists' (RTDSM) at the Federal Reserve Bank of Philadelphia (http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data). The RTDSM proxies the original vintages for each quarter by selecting the data that was originally available around the middle of that quarter (as close as possible to the 15th day of the middle month of a quarter). Vintages of inflation rates are then constructed as the percentage quarterly changes of the respective deflator series.[11]

Explanatory variables

We use in this paper an extensive set of activity and expectations measures to model inflation dynamics. Like the aforementioned inflation rates, the bulk of these variables gets revised so we strive to use as much as possible the original vintages of underlying data. Some of the measures can be directly retrieved from the respective real-time databases, others need to be constructed.

Real output growth - ROUTP We take the original quarterly data vintages for GDP in volume terms from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real output growth rates, i.e., the percentage quarterly change in real GDP.

Real consumption growth - RCONS We take the original quarterly data vintages for real personal consumption expenditures (PCE) from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real consumption growth rates, i.e., percentage quarterly change in real PCE.

Real residential investment growth - RINVR We take the original quarterly data vintages for real residential investment from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct real residential investment growth rates, i.e., percentage quarterly change in the real residential investment level.

Import price inflation - PIMP We take the original quarterly data vintages for the imports deflator from the RTDSM at the Federal Reserve Bank of Philadelphia. Based on these we construct import price inflation, i.e., percentage quarterly change in the imports deflator.

Non-farm payrolls growth rate - NFPR From the ALFRED® real-time database, we take as quarterly vintages those monthly data vintages of non-farm payrolls employment

[11] We define percentage quarterly change as 100 times the quarterly change of the logarithm of the original series.

that are closest to the middle of quarter. Then, we transform these data to the quarterly frequency through averaging; finally, the non-farm payrolls growth rate is constructed as the percentage quarterly change in non-farm payrolls.

Housing starts growth rate - HSTS We take the original quarterly data vintages of monthly housing starts from the RTDSM at the Federal Reserve Bank of Philadelphia. Then, we transform these data to the quarterly frequency through averaging; finally, the housing starts growth rate is constructed as the percentage quarterly change in housing starts.

M2 growth rate - M2 From the ALFRED® real-time database at the Federal Reserve Bank of St. Louis, we take as quarterly vintages those monthly data vintages of the M2 monetary aggregate that are closest to the middle of quarter. Then, we transform these data to the quarterly frequency through averaging; finally, the M2 growth rate is constructed as the percentage quarterly change in the M2 level.

Unemployment ratio - UNEMPL We take the original quarterly data vintages for unemployment as a percentage of the labor force (UNEMPL) from the RTDSM at the Federal Reserve Bank of Philadelphia.

Level term structure factor - YL This is a proxy for the level factor describing the dynamics in the term structure of interest rates. The term structure is approximated by seven interest rates: the 3-month Treasury bill rate, the 6-month Treasury bill rate, both from Global Financial Data (https://www.globalfinancialdata.com/), as well as the Fama and Bliss (1987) 1-year, 2-year, 3-year, 4-year and 5-year zero-coupon bond yields from the CRSP database at Wharton Research Data Services. These are monthly data, which are not revised as they are financial data. In order to get quarterly data we select the aforementioned interest rates at the end of the first month of a quarter. The level term structure factor equals the cross-sectional average across the above seven interest rates for each quarter.

Slope term structure factor - TS This is a proxy for the slope factor describing the dynamics in the term structure of interest rates. We use the same interest rates as for the level term structure factor - see above. These are monthly data; in order to get quarterly data we select the aforementioned interest rates at the end of the first month of a quarter. The slope term structure factor equals the spread between the 5-year zero-coupon bond yield and the 3-month T-bill rate for each quarter.

Real oil price inflation - OIL To construct real oil prices, we first retrieve nominal oil prices - for this we use the West Texas Intermediate oil spot price from Global Financial Data. Quarterly observations result by selecting in each quarter the observed oil spot price closest to the middle of the quarter; as these data are market prices they are not prone to revisions. Quarterly data vintages of real oil prices are then constructed by deflating the aforementioned oil spot price, which is unrevised, by either the GDP deflator or PCE deflator for that vintage, depending on which inflation rate one wants to model. Vintages of real oil price inflation are then equal to the percentage quarterly change in the constructed real oil price level.

Real food commodities inflation - FOOD Vintages of real food commodities inflation are constructed in a similar manner as those for real oil price inflation - see above. Only now the construction is based on the Commodities Research Bureau (CRB) Index of Foodstuffs commodity prices, which is based on the spot prices for butter, cocoa beans, corn, cottonseed oil, hogs, lard, steers, sugar and wheat. The CRB Foodstuffs price index is acquired through Global Financial Data.

Real raw industrial commodities inflation - RAW Vintages of real raw industrial commodities inflation are constructed in a similar manner as those for real oil price inflation - see above. Only now the construction is based on the CRB Index of Raw Industrials commodity prices, which is based on the spot prices for burlap, copper scrap, cotton, hides, lead scrap, print cloth, rosin, rubber, steel scrap tallow, tin, wool tops and zinc. The CRB Raw Industrials price index is acquired through Global Financial Data.

Reuters/University of Michigan Survey of Consumers' inflation expectations - MS The Reuters/University of Michigan Survey of Consumers asks members of the general public, amongst other, to give a quantitative assessment of expected inflation in a year's time. As this is a one-year ahead measure, we lag these series, which are never revised, with four-quarters as to make them properly real-time. The quarterly data are retrieved from http://www.sca.isr.umich.edu/main.php at the University of Michigan.