Scenario generation vs. forecasting: predictive performance criteria and the role of vague priors

Andrei Sarychev†

March 13, 2014

1 Introduction

In many financial regulators and central banks around the world stress testing is used to assess portfolio and business model vulnerabilities. A scenario (stated in terms of macroeconomic, market, or portfolio variables) is in most cases an integral part of the stress-testing process, as it defines the exogenous shocks affecting risky exposures. The regulatory directives relevant to stress testing usually only stipulate that scenarios should be both severe and plausible, but provide little formal guidance on what this implies. Intuitively, one could argue that scenarios should be at the tails of the distribution of all possible realisations. This paper addresses the question how to operationalise this intuition. It proposes a set of criteria that an ideal scenario-generating framework should satisfy and the statistical techniques that render it possible.

Since the first version of the paper was written for internal business purposes of a financial regulator, it was intended to be read by non-specialists. Consequently, section 2 provides an overview of the chief difficulties encountered by empirical macroeconomists and the modelling options available to them. A sophisticated reader may skip it and proceed straight to Section 3 which explains shortcomings of the traditional approach to evaluating forecasting performance in the context of stress scenarios. Section 4 describes a technique for generating variable projections at a given percentile and implements a number of model specifications in the context of the UK macroeconomic dataset. It compares their predictive performance and the ability to generate meaningful scenarios at a

*Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee or Financial Policy Committee.

†andrei.sarychev@bankofengland.co.uk, andrei@sarychev.info.
distribution tail. It is found that the factor-augmented VAR with hierarchical SSVS pri-
ors dominates the competing alternatives. Section 5 makes a shot at formalising the issue
of incorporating laymen’s priors, which often assume the form of a vague scenario nar-
rative. This analysis speaks to the broader challenge facing the econometrics profession, as
formulated by Geweke [2001]:

Addressing the relation between Bayesian investigators, who can only re-
port a limited number of distributions for a finite set of models, and remote
clients, whose models, priors, functions of interest, and loss functions may be
unknown to the investigator.

Section 6 briefly concludes discussing directions for further improvements.

2 Desirable features of a scenario generation frame-
work

To design scenarios that are a priori plausible, a statistical framework utilising all relevant
quantitative information is required. This objective data-driven framework should allow
the user to project the joint evolution of large sets of variables together with the infor-
mation on the empirical likelihood of possible projections. Furthermore, in order to tell
whether a scenario is severe it is imperative to have some reference points for comparison.
This would enable the user to select scenario paths that are appropriately severe along
desired dimensions, focusing on tail outcomes but within the realm of plausible. To do
this an ideal framework should possess the following features:

S1 it should be capable of producing collections of possible future paths rather than a
unique path per variable of interest;

S2 the implied joint distribution of various paths should be consistent with the one in
the historical data;

S3 the framework should not incorporate unverified assumptions on the joint evolution
of variables. This concerns, for example, market efficiency, or rationality of certain
agents, or assumptions about policy functions except when absolute certainty exists
regarding their validity;

In addition, it is practically useful for the purposes of supervisory scenarios that

S4 the framework should allow the user to condition projections on exogenously given
series, e.g., global scenario variables already modelled;
S5 the framework should allow the user to impose constraints on the evolution of variables of interest, to produce stylised scenarios, e.g. featuring “double dip” in the GDP growth, or a rapid rise of interest rates, or a specific profile for asset price indices, etc.

S6 whenever extra conditions are imposed on the generated collections of paths the framework should provide information on the unconditional likelihood thereof.

The subsections of this section review the options available for designing, specifically, macroeconomic scenarios.

2.1 Why linear models?

Any quantitative macroeconomic model must be built around a system of parametric equations. This is because the length of historical macroeconomic series does not exceed a few hundred observations (quite often it is considerably shorter). Even low-dimensional non-parametric models require a lot longer series to robustly estimate functional relationships between variables. In principle, parametric models need not be restricted to linear ones and can be specified up to any desired order. In practice, since this increases dramatically the number of parameters to be estimated, even second-order approximations are much less frequently used than linearisations. Non-linearities may be obviated to some extent by using transformed variables, for example, natural logarithms of the raw data to capture the apparent scale-neutrality (one of the so-called Kaldor facts).

If $y_t$ is a vector of observations of macroeconomic variables in a given time $t$, $\varepsilon_t$ a vector of unobserved disturbances or shocks, a generic backward-looking linear model is written as

$$ y_t = A(L)y_t + \varepsilon_t, $$

where $y_t, \varepsilon_t \in \mathbb{R}^n$, $\text{Var}(\varepsilon_t) = \Omega$, and $A(L)$ is a lag operator polynomial:

$$ A(L) y_t \equiv \left( \sum_{k=0}^{+\infty} A_k L^k \right) y_t = \sum_{k=0}^{+\infty} A_k y_{t-k}. $$

2.2 Why vector autoregression approach?

Models (1) include an important subclass of structural models, which is generically characterised by the non-zero matrix of contemporaneous coefficients $A_0$. In order to consistently estimate these models, theoretical identification restrictions are required. However, these equations can always be restated in reduced form (RF), which can be directly estimated from the data with no assumptions on the theoretical structure of the model needed:
2.2 Why vector autoregression approach?

\[ y_t = \sum_{k=1}^{+\infty} B_k y_{t-k} + \varepsilon_t, \]  

(2)

Despite an arsenal of methods employed to specify or estimate linear models, all of them have a canonical RF VAR representation (2), though possibly a restricted one. For example, many quantitative implementations of non-linear DSGE models are based on linear approximations and are estimated subject to constraints implied by the theory either in the traditional linear regression fashion, or in the “calibration” fashion.

The class of forward-looking “rational expectations” models is characterised by the structural form in which

\[ y_t = \sum_{k=0}^{+\infty} A_k y_{t-k} + E_t \sum_{k=1}^{+\infty} A_{-k} y_{t+k} + \varepsilon_t, \]

where \( E_t \) denotes an expectation operator conditional on the information available at time \( t \). In practical implementation, the order of the lag polynomial \( B(L) \) must be finite \( l < +\infty \).

All regular forward-looking models allow the RF representation \( \text{(2)} \) [Sims, 2002]. Thus NiGEM, a popular large forward-looking macro model, can be made equivalent to a backward-looking vector equation \( \text{(2)} \) with a sparse parameter matrix (embodying a huge number of zero restrictions) and deterministic disturbances \( \varepsilon_t \).

Empirical implementations of \( \text{(2)} \), known as vector autoregression (VAR) models, have become widely adopted in macroeconomic econometrics following the work by Sims [1980; 1986]. The advantage of the vector autoregression (VAR) approach lies in explicitly acknowledging multiple sources of shocks to model equations and estimating their statistical properties (e.g. covariance matrix of disturbances \( \Omega \)). Consequently, the approach allows simulating random realisations of future shocks in Monte-Carlo fashion, generating arbitrarily large collections of potential paths. *If the model is properly specified*, then these families of projections are asymptotically accurate approximations of the true underlying distribution of future variable realisations. Conditions \[ \text{S1} \] and \[ \text{S2} \] then follow.

Since linear models with forward looking terms admit reduced form VAR representation, the latter could be used for forecasting and scenario design without loss of generality. They also have a significant advantage over the structural and/or forward-looking models in that they do not rely on any ex ante assumptions and restrictions. Given that credible macroeconomic identification schemes are few and far between, a potentially severe misspecification is avoided, thus delivering condition \[ \text{S3} \].

Another, practical advantage of reduced-form autoregressions over some forms of struct-

---

1 Calibration is a statistical estimation procedure differing from regression with respect to the sample objective function.
2.3 Why augmenting regressions with factors?

Estimation of multivariate models with lags gets progressively harder as the number of endogenous variables increases, jeopardising condition \( S_2 \). With \( l \) lags and \( n \) endogenous variables, there are at least \( n(n+l+nl-1) \) parameters of the coefficient and covariance matrices to be estimated. This means that forecasters are often restricted in the number of economic variables they could consider, despite the existence of many more potentially relevant ones. This is known as the "curse of dimensionality" problem. On the other hand one of the oft heard criticisms of the classical VAR approach to forecasting concerns the small amount of information used by low-dimensional VARs.

In order to overcome this problem, forecasting researchers introduced various factor models. These postulate that a small set of underlying (observable or latent) factors can account for most of the variation in observed series (Sargent and Sims [1977], [Sargent, 1989], [Stock and Watson, 1989], [Stock and Watson, 1991]). Assuming those factors can be estimated accurately, they can summarise the driving forces behind a large set of economic variables and thus significantly improve the performance and efficiency of macroeconomic forecasting. A famous study by Bernanke, Boivin and Eliasz [2005] popularised the factor-based VAR modelling approach to hypothesis testing.

The forecasting strength of dynamic factor models has been confirmed by several
2.4 Why Bayesian estimation?

While VARs are commonly estimated using classical (frequentist) approach, there are several distinct advantages to Bayesian treatment, especially if one is interested in tail outcomes rather than the mean predictions. Bayesian approach postulates a joint parametric distribution (the prior) of unknown model coefficients and then conditions it on the data, essentially improving on the original guess (this updated distribution is called a posterior). From this posterior distribution one can then pick a single salient element, for example, the mean, to serve as a point estimate. Though the procedures are seemingly different, for some choices of prior distribution the means of model coefficients obtained from Bayesian posteriors coincide with the estimates obtained via classical maximum likelihood approach. This mutual consistency underscores the fact that classical econometric analysis, rather than being fundamentally different, should be viewed as a special case of Bayesian estimation. Below I go through several conceptual and practical advantages offered by Bayesian approach.

2.4.1 Parameter uncertainty and sampling uncertainty

The differences between the two approaches arise when using VAR model estimates for specific purposes, such as forecasting. In classical econometrics, one-step ahead projections are performed by way of drawing multiple realisations of disturbances from a hypothesised parametric distribution. So for a data sample ending in period \( T \), period \( T + 1 \) projections are computed according to:

\[
\tilde{y}_{T+1}^j = \sum_{k=1}^{l} B_k y_{T+1-k} + \tilde{\varepsilon}_{T+1}^j,
\]

where \( \tilde{\varepsilon}_{T+1}^j \) is a \( j \)th draw from the (estimated) distribution of disturbances. Denote a single point projection by \( \tilde{y}_{T+1}^j \in \mathbb{R}^n \). If the number of draws of \( \tilde{\varepsilon}_{T+1}^j \) is sufficiently large, the resulting family of projections \( \{\tilde{y}_{T+1}^j\} \) should be representative of all potential
future realisations of $\mathbf{y}$ (conditional on knowing $\mathbf{B}$ with certainty). In practice, since $\mathbf{B}(L)$ polynomial is unknown, a point estimate $\hat{\mathbf{B}}(L)$ is used. This is independent Monte Carlo simulation. For multiple-step ahead projections, one simply iterates on (3) A single path projection $\{\tilde{\mathbf{y}}_{T+\tau}^j\}_{\tau=1}^{H} \in \mathbb{R}^{n \times H}$ may be referred to as a scenario of length $H$:

$$\tilde{\mathbf{y}}_{T+\tau}^j = \sum_{k=1}^{l} \hat{B}_k \tilde{\mathbf{y}}_{T+\tau-k}^j + \tilde{\epsilon}_{T+\tau}^j$$

Bayesian approach explicitly treats coefficients of $\mathbf{B}(L)$ as unknowns. Hence, the Monte Carlo simulation is performed with notationally small but statistically meaningful difference, substituting $\hat{B}_k$ with $\tilde{B}_k^j$, which is sampled concurrently with $\tilde{\epsilon}_{T+1}^j$:

$$\tilde{\mathbf{y}}_{T+\tau}^j = \sum_{k=1}^{l} \tilde{B}_k^j \tilde{\mathbf{y}}_{T+\tau-k}^j + \tilde{\epsilon}_{T+\tau}^j$$

It is easy to see that classical approach, unlike the Bayesian, ignores the parameter uncertainty whenever it is present, artificially shrinking the support of the distribution of outcomes. By pretending that the model parameters, once estimated, are known with certainty, classical approach effectively distorts that distribution. The bigger the parameter uncertainty in the sample, the more severe is the distortion. When the research objective is to model the distribution of future outcomes, Bayesian methods are much more likely to satisfy condition $S_2$ especially if one is interested in extreme realisations as in stress scenario design.

Classical approach is also not robust to deviations from parameter stability either through the sample period or across the different quantiles of the joint underlying distribution. In contrast, deviations from parameter stability can be treated in Bayesian framework either explicitly or implicitly.

### 2.4.2 Bayesian shrinkage

Bayesian analysis offers an attractive solution to the “curse of dimensionality” problem in multidimensional VAR models. Bayesian VAR models have been found to be a useful forecasting tool by many studies, see [Stock and Watson, 2004] for a review. One of the seminal contributions was made by Litterman [1986], who proposed a simple Gaussian prior for the coefficients of the matrix polynomial $\tilde{\mathbf{B}}$, while treating the covariance matrix $\Omega$ as known.

To understand why Bayesian regressions go a step beyond frequentist estimation in

---

2 An alternative to iterated projections is offered by direct $\tau$-step ahead forecasting, however, no systematic differences in forecasting performance have been identified in the literature.

3 See also subsection 2.4.4 below.
producing tight coefficient estimates it is useful to invoke the concept of Bayesian shrinkage. The problems with obtaining precise estimates of the coefficients of interest are due to the effects of sampling variation. There is a fundamental trade-off between potentially introducing the omitted variable bias by including too few regressors and increasing the noise content of the coefficient estimates by including too many. Positing a prior assumption (e.g., mean and variance) about the distribution of the unknown coefficient of interest may help shrinking the posterior variance. In the Monte Carlo simulation context this means limiting the explosive effect of sampling extreme values. Of course, strong priors come with the risk of misspecifying the model.

It seems natural to attempt to combine the benefits in terms of parsimony and shrinkage offered by factor models and Bayesian estimation techniques. Following in the footsteps of Bernanke et al. [2005], others have applied FAVARs in combination with Bayesian estimation techniques for use in forecasting. For instance, Vitner and Iqbal [2009] used six different models to forecast US house prices and found that the BFAVAR method outperforms other models in term of simulated out-of-sample root mean square error criteria. See also [Banbura, Giannone and Reichlin, 2010] for a comprehensive survey of the empirical literature.

It should be noted that objections to mixing the Bayesian approach and factor modelling have been voiced from purists on either side. Bernanke et al. [2005] are primarily concerned with hypothesis testing in the context of a low-dimensional model. They do not find any advantages of Bayesian approach over classical approach, while noting that the former is computationally more burdensome.

On the other hand, Banbura et al. [2010], departing from the large Bayesian VARs standpoint, are skeptical about the need for factor augmentation. They compare forecasting performance of a 20-variable Bayesian non factor-augmented VAR with that of a factor augmented model and find the former generally superior. It should be noted that Banbura et al. [2010] focus specifically on the accuracy of point forecasts.

In section 4 we will revisit the implications of using this prior specification as well as the relative merits of the factor approach.

2.4.3 Model uncertainty

In higher-dimensional models neither the factor approach nor Bayesian sampling may be enough to resolve the dimensionality issue. However, it is often reasonable to assume that many of the coefficients of the \( B \) and \( \Omega \) matrices are zero. The problem is that in the reduced form context it is near impossible to postulate credibly which coefficients should be excluded. Traditionally, econometricians relied on comparisons of all the possible submodels (incorporating all possible combinations of zero restrictions). With the exponential
cost of enumerating all the submodels, the computational requirements for this procedure are prohibitive. That is why the stochastic search variable selection procedure proposed by George, Sun and Ni [2008] presents an attractive solution to dealing with the model uncertainty. George et al. [2008] embed the regression setup in a hierarchical Bayesian normal mixture model, where latent variables are used to identify model subset choices. The idea goes back to the results of Aitchison [1975] and Madigan and Raftery [1994], who saw the advantage of the Bayesian approach in producing finite sample inferences on the parameters of interest simultaneously allowing for both model and parameter uncertainty. Madigan and Raftery [1994] argue that under the prior incorporating model uncertainty it is better to average the predictive density over the model space than to condition on any single model [4]. This makes Bayesian analysis with a join SSVS prior a prime candidate for satisfying S2 in the real world, where model and parameter uncertainty are pervasive. The details of the SSVS prior are supplied in appendix A.

2.4.4 Imposing constraints on scenario families

The frequentist approach may be inconvenient for imposing constraints on the scenario to be chosen. Suppose one is interested in projecting the possible evolution of $y$ subject to a constraint $\{\tilde{y}_j^T + \tau\} \in \mathbb{Y} \subseteq \mathbb{R}^{n \times H}$. Since the coefficients of $B$ polynomial are rigidly set, extreme demands may be placed on the realisations of $\tilde{\varepsilon}_j^T$ to obtain draws that satisfy the constraint [5].

In contrast, off-the-shelf Bayesian techniques exist for the constructing predictive densities conditional on constraints imposed on projected variable paths, see [Waggoner and Zha, 1999]. This makes a Bayesian model much more likely to generate ranges of scenarios consistent with a set of constraints imposed through scenario selection, satisfying S5.

3 Model evaluation criteria

There are two broad approaches to evaluating predictive power of competing models. The forecasting literature has traditionally focused on measures of out-of-sample forecast accuracy based on loss functions. These are applied to point forecasts, corresponding either to point estimates of the model parameters, or in case of Bayesian models, to posterior means of the values of interest. A popular criterion uses the quadratic loss function to compute the root mean square error (RMSE), defined for $y_i$, an element of $Y$.

---

4 As George et al. [2008] argue, what SSVS-driven integration achieves is the minimisation of the Bayes risk under Kullback-Leibler loss.

5 Technically, this means that the set of all $\{\tilde{y}_j^T + \tau\}_{\tau=1}^H$ such that $\{\tilde{y}_j^T + \tau\}_{\tau=1}^H \in \mathbb{Y}$ has a low measure or is even empty.
vector $\mathbf{y}$, as
\[
\left(\frac{1}{H} \sum_{\tau=1}^{H} (\bar{y}_{i,T+\tau} - y_{i,T+\tau}^O)^2\right)^{\frac{1}{2}},
\]
where $\bar{y}_{i}$ are the forecast values, $y_{i}^O$ observed values, $i$ is the index of the variable in the vector, and $H$ is the forecast horizon. Note that because units of RMSE may not be comparable across variables, these measures do not admit a straightforward aggregation across forecast dimensions.

If the forecasting objective is to predict the whole distribution of future outcomes, as in the scenario analysis, the RMSE criterion may fail to distinguish meaningfully between competing models. Consider two standard linear structures as in (3) that have identical coefficient matrices $\mathbf{B}_k$ but different covariance matrices of shocks $\varepsilon_{T+1}$. Their posterior mean forecasts will be identical, however, the posterior distributions could have arbitrarily large discrepancies. It is this inability of the RMSE criterion to account for substantial differences in the competing models that renders it inadequate for the scenario design.

An alternative to criteria based on point forecast comparison is offered by the model uncertainty literature, which has relied on the Bayes factors (ratios of marginal likelihoods) to evaluate the evidence in favour of hypotheses. It is useful to introduce formalism here, denoting the historical observations up to time $t$ by $y_{t}^O$, parametric models by $M$, and parameter vectors by $\theta$. Bayesian inference posits a prior density $p(\theta| M)$ and analyzes the model posterior $p(\theta| y_{t}^O, M) \propto p(y_{t}^O | \theta, M) p(\theta| M)$, where $p(y_{t}^O | \theta, M)$ is the likelihood function. The Bayes factor in favour of model $M_1$ as opposed to an alternative $M_2$ is defined as
\[
\frac{p(y_{t}^O | M_1)}{p(y_{t}^O | M_2)},
\]
where
\[
p(y_{t}^O | M_i) = \int p(y_{t}^O | \theta, M) p(\theta | M) d\theta \tag{4}
\]
is the likelihood marginalised with respect to the parameter vector $\theta$. While this is superficially similar to the frequentist likelihood ratio test statistic, it can be used to compare broader classes of models. For example, the likelihood ratio test, to be meaningful, requires models to be nested. In contrast, Bayes factors are meaningful even if $M_1$ and $M_2$ are non-nested. Kass and Raftery [1995] describe various approaches to calculating Bayes factors in applications. One convenient way of doing this, which also addresses the issue of forward projection, is posterior simulation. However, one substantial limitation to model comparison is that Bayes factors may not be well-defined for non-Bayesian or hybrid models (those mixing frequentist and Bayesian features, etc. the Litterman’s Minnesota prior to be discussed below). For an MLE $\hat{\theta}$, in absence of the clear framework to
account for sampling and parameter uncertainty, looks like

\[ p \left( y_t^O \mid M_i \right) = p \left( y_t^O \mid \hat{\theta}, M \right) , \]

which is already maximised with respect to \( \hat{\theta} \), and so by definition will beat any Bayesian model that is averaging over the \textit{a priori} likely values of \( \theta \).

It is nevertheless possible to use the concept closely related to the marginal likelihood to evaluate the predictive performance of the models regardless of their structure. Define predictive density for one period ahead projection as

\[ p \left( y_{t+1} \mid y_t^O, M \right) = \int p \left( y_{t+1} \mid y_t^O, \theta, M \right) p \left( \theta \mid y_t^O, M \right) d\theta. \] (5)

The corresponding distribution combines the uncertainty about the unknown parameter vector \( \theta \) and the uncertainty about the future values of \( y_{t+1} \) (for any given \( \theta \)), conditional on the observed history \( y_t^O \) and the assumptions of the model \( M \). Substituting the observed values of \( y_t^O \) results in the predictive likelihood \( p \left( y_{t+1} \mid y_t^O, M \right) \) of model \( M \) conditional on the actual realisations. Furthermore, since the marginal likelihood of model \( M \) can be written as

\[ \log p \left( y_T^O \mid M \right) = \sum_{t=1}^{T} \log PL_M (t), \]

an additive decomposition is possible

\[ \log p \left( y_T^O \mid M \right) = \sum_{t=1}^{S} \log PL_M (t) + \sum_{t=S+1}^{T} \log PL_M (t) = \log p \left( y_S^O \mid M \right) + \log p \left( y_T^O \mid y_S^O, M \right) \]

Comparing the second terms for different models allows using standard conventions of the Bayesian hypothesis testing. The difference \( \log \frac{p \left( y_T^O \mid y_S^O, M_1 \right)}{p \left( y_T^O \mid y_S^O, M_2 \right)} \) is conceptually similar to Bayes factor and therefore can be used for the meaningful comparison against the standard decision thresholds. Traditionally, \( \log_{10} \) values of the Bayes factor in excess of 0.5 signify that the evidence in favour of \( M_1 \) is substantial; for values between 1 and 2 the evidence is strong, and above 2, decisive [Kass and Raftery, 1995]. Note that the predictive
likelihood can be computed for non-Bayesian and hybrid models and is comparable to that of Bayesian models, since the conditionality on the training sample \( \{1, \ldots, S\} \) implies also conditionality on the point estimators that are functions of the data in that sample. The model evaluation then amounts to computing the log score for the forecasting model employing point estimates. Further details are supplied in Geweke [2005] and Geweke and Amisano [2010].

Note that the predictive likelihood may be evaluated on a strict subset of the variables predicted within a given model: \( p\left(x_t^O \mid y_{t-1}^O, M\right) \) is the induced density where \( x_t^O \) is a section of the full vector \( y_t^O \). This allows comparing the predictive performance of the models of different size with respect to that particular density on the subset.

4 Predictive density evaluation and extreme scenarios: the Great recession

4.1 Bayesian projection and scenario selection

This section implements forward projection, based on the SSVS BFAVAR and a few competing models, in the context of historical UK macroeconomic series. Gauging the ability of the framework to accurately project scenario densities is best accomplished in those periods where most posterior mean forecasts have failed, i.e. the extreme historical episodes. A good candidate for such an episode is the period starting in Q4 2007, when the growth of UK GDP has sharply decelerated. Since the recession that followed was far from universally predicted, it is reasonable to consider the subsequent GDP realisations a tail outcome.

The dataset, used in construction of the latent factor variables consists of a collection of publicly available UK time series\(^6\) with the back history dating from at least Q1 1972 (this was chosen as the starting point because, for example, UK unemployment data are only available at quarterly frequency from that time). Only data up to Q3 2007 (as would be available to an econometrician at the time) were employed. The full list of variables is in the Appendix \(^7\). All variables were seasonally adjusted and transformed to stationary.

In addition all interest rates \( i \) were transformed by \( \log (1 + i) \) to preserve the log-linearity of interest parity conditions, such as Fischer’s formula\(^7\).

A Monte-Carlo simulation is used to project the evolution of the factor variables.

---

\(^6\)All series with the exception of UK corporate bond yield are taken from Thomson-Reuters Datas‐stream. UK corporate bond is INGBRW composite index of corporate bond yields, maintained by The Economist magazine until 15/10/2011.

\(^7\)For example, for an inflation linked yield \( r \), a nominal rate \( i \) (both riskless), and a relevant price index \( p \) the following would hold: \( \log (1 + i_t) - \log (1 + r_t) = \pi_t = \log (p_t - p_{t-1}).\)
4.2 Prior specifications and the sample scenarios

When it is impossible to sample directly from the joint posterior distribution of model parameters, the appropriate Gibbs sampler is used. At the $n$th iteration of the (MC)MC algorithm we obtain a valid draw from the joint distribution of $\Phi^n$, $\Sigma^n$, and $\{\varepsilon^n_{T+\tau}\}_{\tau=1}^H$, which is used to recursively project the future evolution of the factors for $H$ periods as follows:

$$
\begin{align*}
\hat{F}_{T+\tau}^n &= \Phi^n \hat{F}_{T+\tau-1}^n + \varepsilon_{T+\tau}^n, \\
\varepsilon_{T+\tau}^n &\sim N(0, \Sigma^n), \\
\tau &= 1, \ldots, H.
\end{align*}
$$

Here $\{\hat{F}_T^n\}_{\tau=1}^H$ is the $n$th raw scenario. For the purposes of this paper, the raw scenario family includes 50,000 alternative joint histories for all factor variables. A stylised scenario can then be constructed by grouping raw scenarios into equivalence classes with respect to a criterion of interest. For example, to produce a stylised scenario for GDP, the raw scenarios can be ordered according to a sorting criterion such as the average GDP level during the scenario period or any sub-period thereof, start to trough growth, start to a fixed point growth, etc.

Once 50,000 raw scenarios are sorted according to the criterion above, resulting in the ordering $\{n_k\}$, the percentiles are naturally defined as integer indices $n_{100\times50,000}$. To create a robust and representative stylised scenario, a median path is extracted from a strip of adjacent raw scenario paths. Thus the raw scenario paths that fall within the same strip according to the sorting criterion are deemed to be in the same equivalence class. To construct the $j$th percentile, sorted raw scenarios with indices $n_{j-\delta100\times50,000}$ to $n_{j+\delta100\times50,000}$ are used, where $\delta$ is the strip width. In what follows $\delta$ is set to equal 0.005, half of one percentage band. Figure presents some examples of how dissimilar stylised scenario sets look like when different sorting criteria are applied to UK GDP paths from the same underlying family of 50,000 raw scenarios.

4.2 Prior specifications and the sample scenarios

Initially eight different specifications were run as summarised in Table I:

1. **SSVS** is a factor-augmented VAR with the joint SSVS prior for $\Phi$ and $\Sigma$.

2. **MN** is a factor-augmented VAR with the standard Minnesota prior for $\Phi$. This specification assumes that $\Sigma$ is diagonal and known with certainty.

3. **MLE** is a factor-augmented VAR that uses point estimates of $\hat{\Phi}$ and $\hat{\Sigma}$ in all simulations.

For some of the models the parameter matrices are invariant across the simulations, etc. for the MLE model $\Phi^n = \hat{\Phi}$, $\Sigma^n = \hat{\Sigma}$.
4.2 Prior specifications and the sample scenarios

Figure 1: Examples of stylised scenario sets resulting from the same raw scenario family
4. SSVS BVAR is a BVAR of the same dimensionality as 1-3, but specified in terms of observable series, not factors. The list of variables entering the VAR is provided in appendix C. The joint SSVS prior is assumed for $\Phi$ and $\Sigma$.

5. ALLINC is a factor-augmented VAR with the joint prior for $\Phi$ and $\Sigma$, differing from SSVS in that distributions of all coefficients in both matrices are assumed to be non-degenerate (equivalent to SSVS with all $\gamma_i = 1$ and $\omega_{ij} = 1$).

6. FA INIW is a factor-augmented VAR with the Independent Normal-Inverse Wishart prior of Karlsson [2013]. The prior distribution $\Sigma$ is assumed to be inverse Wishart with no additional restrictions.

7. FA BGR is a factor-augmented VAR with the modification of the Extended Natural Conjugate prior proposed by Banbura et al. [2010]. It imposes Minnesota-type restrictions on the multivariate normal distribution of $\Phi$ and infers hyperparameter values for the inverse Wishart distribution of $\Sigma$ based on the sample residual covariance matrix\(^9\). The MCMC is implemented via a VAR on the sample augmented with dummy observations with a Normal-Diffuse prior as in Banbura et al. [2010].

8. BGR BVAR is a VAR of the same dimensionality as above, specified in terms of the same set of observable series as SSVS BVAR. The prior assumption and the MCMC implementation is the same as in FA BGR.

All VARs included 8 quarterly lags. All factor-augmented VAR specifications used the same point estimates of the factor coefficients (obtained via principal components as the least CPU-intensive option). The number of the latent variables to be used was chosen by the Kaiser criterion, while UK GDP was used an an explicit factor variable. The predictive likelihood for these models is conditional on the coefficients of matrix factor extraction procedure. Notice that the non-factor augmented model do not predict all the information variables necessary to construct values of the factor variables in the evaluation sample. Thus the minimum common subset of predicted variables across all the models is just the GDP. The partial predictive likelihood can be evaluated for its induced univariate distribution, allowing comparison across all the models. The implementation of the full predictive density for all the variables in the information dataset, although conceptually straightforward, is computationally and memory intensive and therefore is left to future research.

Figure 2 presents the evolution of the partial predictive likelihood over the evaluation sample (Q4 2008 – Q3 2011). It shows that the evidence in favour of specification SSVS

---

\(^9\)As in [Banbura et al., 2010], it is implemented as a VAR with dummy observations under a Normal-Diffuse prior.
4.2 Prior specifications and the sample scenarios

Prior/estimation scheme:

- Point estimates $\hat{\Phi}$, $\hat{\Sigma}$
- Minnesota prior for $\Phi$
- Conjugate Wishart prior for $\Sigma^{-1}$
- Independent Wishart prior for $\Sigma^{-1}$
- Element-wise prior for $\Sigma$
- SSVS hierarchical hyperprior

Factor augmented

<table>
<thead>
<tr>
<th>Cumulative partial predictive</th>
<th>SSVS</th>
<th>ALLINC</th>
<th>MN</th>
<th>MLE</th>
<th>INIW</th>
<th>BGR</th>
<th>BVAR</th>
<th>FA BGR</th>
<th>SSVS BVAR</th>
<th>Actuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayes factor vs. SSVS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start to trough drop within 2 years, %</td>
<td>-6.3</td>
<td>-0.9</td>
<td>-0.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-25.20</td>
<td>-26.60</td>
<td>-6.0</td>
</tr>
<tr>
<td>Start to trough drop during the scenario period, %</td>
<td>-8.1</td>
<td>-1.2</td>
<td>-0.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-6.0</td>
</tr>
</tbody>
</table>

Table 1: Summary of models estimated on UK data
4.2 Prior specifications and the sample scenarios

Figure 2: Predictive Bayes factor vs. SSVS (GDP)

![Graph showing predictive Bayes factor vs. SSVS (GDP)]

The failure of the alternatives to SSVS can be further illustrated by considering the stylised scenario families generated by those models. Figure 3 displays the 3rd percentile paths from the families generated using MLE, MN, and the SSVS prior specifications. In all simulations, projected raw scenario families were sorted by the GDP level as of Q4 2008 (referred to as Midterm level in the figures). The 3rd percentile itself was chosen because it was the closest integer percentile at which the stylised scenario path simulated with the SSVS specification traced the actual GDP evolution. Table 1 reports summary statistics for these paths.

It is clear that the two former specifications do not generate the range of outcomes that is wide enough to encompass the actual negative realisations. Moreover, the growth rebounds quite strongly in both of them. The relative ranking of the severity of the 3rd percentile paths is intuitive: the MLE does not allow for any parameter uncertainty and hence artificially restricts the set of possible parametric configurations to point estimates. Any variance in the resulting raw scenario families is due to the sampling variance of the residuals. MN model provides slightly more flexibility as the Monte-Carlo simulation integrates over the space of possible $\Phi$, but the rudimentary treatment of the $\Sigma$ matrix does not allow it to truly account for all the sources of uncertainty present in the data. In
4.2 Prior specifications and the sample scenarios

contrast, the SSVS specification appears to do a reasonable job (that is, scenarios similar to the actual outcome appear in the projected family at a conservative but not extreme level of likelihood).

Figure 3 displays assorted paths from the families generated using SSVS BVAR, joint non-SSVS BFAVAR (ALLINC), FA INIW and FA BGR prior specifications. No BGR BVAR scenario paths are plotted because they are completely off the scale for even small deviations from the median. The plots demonstrate that these specifications result in excessively noisy projections, and it’s not hard to see why.

If infinitely long samples were available, ALLINC joint prior specification would obviate the need to draw restricted models because it nests all of them. However, with finite samples this specification piles up too much variance from every single coefficient (note that in a 11-dimensional VAR with 8 lags there are 979 elements in matrix $\Phi$ and 121 in matrix $\Sigma$) without distinguishing whether this variance is capturing the genuine parameter/model uncertainty or is just noise.

SSVS BVAR specification is clearly misspecified due to omitting variables (observable or latent) that contribute to explaining variation in the data. Omission of relevant factors predictably leads to loosely estimated equations, which shows up in overblown total variance in projected paths.
4.2 Prior specifications and the sample scenarios

Both FAVAR specifications using the inverse Wishart prior for $\Sigma$ produce explosive collections of scenarios. This is apparently due to the undesirable feature of the prior that makes high variance and extreme correlation values highly likely for sampled covariance matrices.

Finally, a natural reaction to a misspecified mid–size VAR is to increase the dimensionality of the system. A fixed covariance Minnesota type BVAR and a BGR BVAR specifications were run on, respectively, 39 and 30 variable datasets and failed to produce meaningful results. While a medium-to-large BVARs may achieve sufficient shrinkage to produce mean point forecasts that look reasonable (because neither parameter posterior variances nor shock posterior covariance matrix interact with the posterior coefficient means), they are not fit for the purpose of projecting density, because large residual variances and coefficient variances snowball at the simulation stage.

In contrast to the specifications discussed above, the joint SSVS FAVAR specification appears to achieve significant Bayesian shrinkage by keeping both parameter and residual variances under control. It is fair to interpret this as a success of the SSVS technique, which a) apparently samples significantly restricted models, and b) visits (draws from) only those models that are plausible given the historical series.
5 Sensitivity of the results to the information dataset and “narrative” priors

This section considers the question of how to implement prior hypotheses about possible sources of stress. In the context of our 2008 recession example, it is reasonable that the econometrician would be aware about the mounting problems in the US subprime mortgage market. Notice that the simulations of the previous section did not employ US data. This would result in misspecification if there were both relevant sources of US shocks and transmission channels through which US outcomes could affected UK macroeconomic variables. While setting up a structural model spelling out those shocks and channels would be a daunting task, there are straightforward ways of incorporating US variation into the UK reduced-form VAR. There are still non-trivial issues:

- which US variables should be included into a UK VAR system?
- should they be treated as exogenous or jointly endogenous with the UK variables?
- should they receive any extra probabilistic weight in those specification that stochastically search for non-zero coefficients?

To investigate the above issues, in our pre-Q4 2007 universe we hypothesise that concurrent developments in the US housing market may contribute to UK risks. If one were to choose a single US variable to be added to the UK information dataset, then housing starts\textsuperscript{10} series is a good candidate. One reason is that unlike house price indices this variable has reached its peak in Q2 2005\textsuperscript{11} and by Q1 2008 has dropped by roughly 50%. Another tentative reason is suggested by the findings of Chahrour [2010] that hstarts captures a substantial share of the informational content of the estimated latent factors. If a plausible-sounding narrative for the spill-over effect were needed, it would be easy to produce in hindsight. The presence of a large house price bubble in the US makes the situation inherently precarious. While it’s difficult to predict the timing of the bubble burst on the basis of the asset prices alone, the new housing starts may serve as a leading indicator of the market cooling down. Depressed US housing starts suggest a higher probability of house price collapse, which would impact on the solvency and liquidity of the US financial sector and from there spill over to the UK due to market externalities and contagion.

Four different specifications were run as summarised in Table 2:

\textsuperscript{10}referred to as hstarts in what follows; its Thomson Reuters Datastream name is US HOUSING STARTED VOLN.

\textsuperscript{11}if using raw data; if one uses non-seasonally adjusted data, the peak date is Q1 2006.
1. **SSVS exog** is an SSVS VAR formulated in terms of UK GDP, hstarts, and nine latent factors generated from the UK dataset (no US variables other than hstarts are used). Matrix $\Phi$ is restricted to be block-recursive so that the UK variables may be affected by hstarts but not the other way around.\(^{12}\)

2. **SSVS hstarts** is similar to **SSVS exog** but no exogeneity restriction is imposed so that all coefficients may be non-zero. The prior probability that hstarts should have non-zero coefficients is doubled relative to all other variables, so, for example, lags 1-4 of hstarts are included in the model with probability 1 in all equations (the corresponding $\tau_i = 1$).

3. **SSVS impartial** is similar to **SSVS hstarts** but it does not treat hstarts any differently from the UK factor variables (the prior probabilities of inclusion of a coefficient into the model, $p_i$, only depend on the lag order as in section A.2).

4. **SSVS agnostic** adds hstarts to the UK information dataset used to construct the latent factors, but does not explicitly include it in the VAR system.

Table 2 reports values of both partial and full predictive likelihood (when applicable) evaluated on the future values of the factors computed as linear combinations of the ex-post observed information variables. All four specifications are directly comparable to the SSVS specification of Section 4 in terms of the partial predictive likelihood. The partial predictive Bayes factor fail to furnish strong evidence in favour of either of the specifications, although it is symptomatic that SSVS and SSVS agnostic are the best performers. Also, the specifications that include hstarts as an observable factor variable (SSVS exog, SSVS hstarts, and SSVS impartial) produce predictive densities for the identical set of variables (UK GDP, US hstarts, and the nine principal components from the main UK dataset) and therefore are comparable with respect to the full predictive likelihood. **SSVS exog** beats the competing models decisively.

---

\(^{12}\)This is to preserve the dimensionality of the VAR and make it comparable to specifications using only UK data.

\(^{13}\)This is equivalent to sampling outcomes from a univariate regression of hstarts on its own lags and using them as exogenous predictors.
Use of US hstarts:
- US hstarts is on the top of the block-recursive UK FAVAR model
- US hstarts is factor variable in FAVAR with doubled prior inclusion probability
- US hstarts is treated symmetrically with the other factor variables in the FAVAR
- US hstarts is merely added to UK information set

<table>
<thead>
<tr>
<th></th>
<th>SSVS agnostic</th>
<th>SSVS exog</th>
<th>SSVS impartial</th>
<th>SSVS hstarts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative partial predictive Bayes factor vs. UK SSVS</td>
<td>0.12</td>
<td>-0.56</td>
<td>-0.75</td>
<td>-1.17</td>
</tr>
<tr>
<td>Cumulative predictive Bayes factor vs. SSVS exog</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>for UK GDP, US hstarts, and 9 factor variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start to trough drop within the 2 years, %</td>
<td>-6.0</td>
<td>-6.7</td>
<td>-7.0</td>
<td>-7.6</td>
</tr>
<tr>
<td>Start to trough drop during the scenario period, %</td>
<td>-9.5</td>
<td>-12.6</td>
<td>-13.6</td>
<td>-16.0</td>
</tr>
</tbody>
</table>

Table 2: Summary of models estimated on UK and US data
The table also reports summary statistics for the 3rd percentile paths, while figure 5 displays the corresponding trajectories for all SSVS specifications using hstarts. To a superficial observation, all 4 paths look reasonable. Start-to-trough magnitudes reported in table 2 reveal, however, that the three specifications that push hstarts to “cause” UK variables most aggressively (SSVS exog, SSVS hstarts, and SSVS impartial), feature markedly more severe outcomes at the 3rd percentile, especially over the entire scenario period. This may be be interpreted in two different ways: either that these specifications capture genuine downside risks manifested in hstarts, or that introducing a loosely related variable into the VAR interferes with the efficiency of the model selection performed by the SSVS algorithm. If the latter interpretation is true, we would expect the range of outcomes to be wider in general than for the baseline UK SSVS specification.

Figure 6 plots the 5 to 95 percentile range paths for SSVS agnostic and SSVS hstarts specifications. It suggests that the range of outcomes generated by the latter are just wider than that of SSVS and SSVS agnostic. This is consistent with the former specifications failing to achieve Bayesian shrinkage due to introducing possibly irrelevant prior information. A tentative conclusion from the above comparison is that an objective framework should be better trusted to properly utilise the informational content of a potentially
relevant addition to the dataset than a human storyteller, who may introduce more noise and impede Bayesian shrinkage.

6 Conclusions and directions for future research

SSVS BFAVAR framework appears to produce better simulated predictive densities compared to all other alternatives. This is thanks to the combination of several powerful techniques coping with the “curse of dimensionality”. There are several ways in which the framework may potentially be improved, although whether these improvements might do more harm than good in a macroeconomic dataset context is an open question.

While this paper argued strongly for the integrated Bayesian treatment of all parameters, it treated the extracted latent factor series as known. One possible extension is the probabilistic inference on those. Prior assumptions may also be generalised to allow for a varying number of relevant factor variables. A potentially restrictive consequence of using the principal components approach to factor variables was that the underlying series had to be the same length and same frequency. This prevented variables with shorter histories from being in the information dataset, unless one was willing to reduce the size of the sample used for the FAVAR. Bayesian methods are capable of mixing series of different length.
The same applies to variables that are observed at either higher or lower frequency than quarterly. Information is inevitably lost when either excess observations are discarded (as in the former case) or the whole series (as in the latter case). Yet another shortcoming of the principal components approach is the need to make the underlying data stationary. This may theoretically discard useful information contained in the levels of the variables in case long term relationships between non-stationary variables are present in the data.

Finally, a conclusive assessment of the relative accuracy of various statistical specification considered in this paper should ideally be done by comparing full predictive likelihoods evaluated over a large set of observable UK variables of interest. This straightforward but extremely time-consuming task is left to be implemented in the future.
A Overview of the SSVS Bayesian Factor-Augmented VAR (BFAVAR) methodology

A.1 Factor representation of multivariate dynamics

Denote by $X_t$ the set of the observed variables which we refer to as “information” variables. The factor model postulates that the dynamics of these variables can be summarised by the evolution of a handful of factor variables $F_t$, most or all of which may be unobserved. Note that there may be a subset of the factor variables that are observed and are therefore part of the set $X_t$ as well. The observation equation then relates the observation variables to the contemporaneous values of the factor variables:

$$X_t = \Lambda F_t + e_t,$$  

where elements of $e_t$ are independent across members of $X_t$ ($\text{Var}(e_t)$ is diagonal). Meanwhile the state equation describes the joint evolution of the factor variables in terms of a vector autoregression ($\Phi(L)$ is a lag polynomial of a finite order $l$):

$$F_t = \Phi(L) F_{t-1} + \varepsilon_t.$$  

An estimation procedure is needed to come up with estimates of the factor variables themselves, and the parameter matrices $\hat{\Phi}(L)$ and $\hat{\Lambda}$, as well as the parameters of the covariance matrices $\text{Var}(\varepsilon_t)$ and $\text{Var}(e_t)$. If estimates are accurate, one would be able to generate a representative family of potential future paths.

Several approaches exist to recovering the latent factors. The easiest to implement is the principal components (PC) representation, based on the decomposition of the covariance matrix of $X_t$. Other options include spectral methods, state-space models, dynamic regressions, and a fully integrated Bayesian treatment, in which sets of factor variables $\tilde{F}_t$ are sampled jointly with the rest of the parameters. In what follows I describe the priors for estimating $\Phi(L)$ and $\text{Var}(\varepsilon_t)$. The same approach can be applied verbatim to $\Lambda$ and $\text{Var}(e_t)$.

A.2 Bayesian prior specification and posterior sampling

Denote by $F_t = \begin{bmatrix} F_{t-1}' & F_{t-2}' & \ldots & F_{t-l}' \end{bmatrix}'$ and $\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 & \ldots & \Phi_l \end{bmatrix}$.  

---

14 The full list of variables used in trial runs is available in the appendix B.
15 The existence of this matrix relies on the assumption that all variables are stationary.
16 A vector of constant values is also used as a R.H.S. regressor, together with its coefficient, but is omitted to simplify exposition below.
Now define $\phi$ as the vectorised matrix $\Phi$, and specify the SSVS prior, that is, assume that elements of $\phi$ are distributed according to a mixture of two Normal distributions:

$$
\phi_i | \gamma_i \sim (1 - \gamma_i) N(0, \tau_0^2) + \gamma_i N(\alpha_i, \tau_1^2),
$$

where $\gamma_i \sim \text{i.Bernoulli}(p_i)$

Here the hyperparameter $\gamma_j$ is itself an (independent) Bernoulli random variable, treated in the same fashion as other parameters of the model and estimated from the data. Following [George et al., 2008], the “low” standard deviation $\tau_0$ is set to a small value (0.1 times $\hat{\sigma}_{\phi_i}$, the standard error associated with the unconstrained MLE estimate of $\phi_i$), while $\tau_1 = 10 \hat{\sigma}_{\phi_i}$. $\alpha_i$ is the prior mean, conditional on the corresponding regressor being included in the model. It can be set to a point estimate, or zero, or another value, according to the desired shrinkage. The natural prior value for all the coefficients $\phi_i$, $p_i = \frac{1}{2}$, To further limit the number of non-zero coefficients in the model and to achieve effective shrinkage, I set the prior value of $p_i = \frac{1}{2}$ for coefficients on lagged variables with lags $l \leq 4$, $\frac{1}{4}$ for $4 < l \leq 8$, $\frac{1}{8}$ for $8 < l \leq 12$ and so forth.

The residuals $\varepsilon_t$ are assumed to be Gaussian and independent across time:

$$
\varepsilon \sim N(0, \Sigma \otimes I_T)
$$

The intratemporal covariance matrix $\Sigma$ is constructed as a product:

$$
\Sigma^{-1} = \Psi \Psi',
$$

where $\Psi$ is an upper-triangular matrix. This ensures positive definiteness of matrix $\Sigma$.

The diagonal elements $\psi_{ii}$ are drawn from the gamma distribution, and the above diagonal elements are drawn from the SSVS mixture of normal distributions similar to $\phi_j$ above:

$$
\begin{pmatrix}
\psi_{11} & \psi_{12} & \ldots & \psi_{1K} \\
\psi_{22} & \ldots & \\
\vdots & \\
\psi_{KK} & 
\end{pmatrix}
$$

$$
\psi_{ii}^2 \sim \Gamma(a_i, b_i),
$$

$$
\psi_{ij} | \omega_{ij} \sim (1 - \omega_{ij}) N(0, \kappa_{0ij}^2) + \omega_{ij} N(\alpha_j, \kappa_{0ij}^2), \text{ for } i \neq j
$$
A.2 Bayesian prior specification and posterior sampling

\[ \omega_{ij} \sim \text{i.Bernoulli} \left( q_{ij} \right) \]

where \( \omega_{ij} \) is a vector of zeroes and ones of length \( j - 1 \), drawn from a Bernoulli distribution, and \( \kappa^2_{0ij}, \kappa^2_{1ij}, a_i, b_i, q_{ij} \) are the prior hyperparameters (i.e., parameters of the distribution of model parameters). George et al. [2008] provide guidance on how to select their values optimally.

The resulting posterior is too complex to yield analytical results. It is also impossible to sample from the joint posterior, which is why a Gibbs sampler is used to draw sequentially values of the parameters \( \{\phi_i\}, \{\gamma_j\}, \{\psi_{ii}\}, \{\psi_{ij}\}_{i \neq j}, \) and \( \{\omega_{ij}\} \). The details of the conditional posteriors can be found in [George et al., 2008].
B List of UK variables in the FAVAR information dataset

UK GDP AT MARKET PRICES (CVM) CONA
UK FINAL CONSUMPTION EXPENDITURE - HOUSEHOLDS (CVM) CONA
UK GENERAL GOVERNMENT: FINAL CONSUMPTION EXPENDITURE (CVM) CONA
UK OUTPUT INDEX - CONSTRUCTION (CVM) (DISC.) VOLA
UK VOLUME OF EXPORTS SADJ
UK VOLUME OF IMPORTS SADJ
UK EXPORTS OF GOODS AND SERVICES (CVM) CONA
UK IMPORTS OF GOODS AND SERVICES (CVM) CONA
UK EXPORTS - BALANCE OF PAYMENTS BASIS CURA
UK IMPORTS - BALANCE OF PAYMENTS BASIS CURA
UK GROSS FIXED CAPITAL FORMATION CONA
UK GROSS NATIONAL INCOME CURA
UK LFS: UNEMPLOYMENT RATE, ALL, AGED 16 & OVER SADJ
UK INDUSTRIAL PRODUCTION INDEX - MANUFACTURING VOLA
UK INDEX OF PRODUCTION - ALL PRODUCTION INDUSTRIES VOLA
UK INDUSTRIAL PRODUCTION: MINING & QUARRYING (DISC.) VOLN
UK INDUSTRIAL PRODN: ELECTRICITY, GAS & WATER SUPPLY (DISC.) VOLA
UK INDUSTRIAL PRODUCTION: FOOD, DRINK & TOBACCO (DISC.) VOLN
UK INDUSTRIAL PRODN.: COKE, PETROLEUM PROD, NUCLEAR FUEL (DISC VOLN
UK INDUSTRIAL PRODN.: CHEMICALS & MAN-MADE FIBRES (DISC.) VOLN
UK INDUSTRIAL PRODUCTION - ALL PRODUCTION INDUSTRIES VOLA
UK GDP DEFLATOR SADJ
UK HN: RESOURCES: WAGES & SALARIES CURA
UK MONETARY AGGREGATE M4 CURA
UK HN: RESOURCES: DISPOSABLE INCOME GROSS CURA
UK M4 SECT.ANALYSIS: M4 LEND INDIV.: DWELLS.SECURED LOANS: LEVEL
UK HN: REAL HOUSEHOLDS DISPOSABLE INCOME CONA
GBP TO USD - EXCHANGE RATE
SOUTH KOREAN WON TO USD (KO) - EXCHANGE RATE
UK BANK OF ENGLAND BASE RATE (EP)
UK 3 MONTHS TREASURY BILLS YIELD (EP)
UK GROSS REDEMPTION YIELD ON 10 YEAR GILT EDGED STOCKS (AVE)
UK AVERAGE GROSS REDEMPTION YIELD ON 5 YR GOV.SECS (PERIOD AVE)
UK GROSS REDEMPTION YIELD ON 20 YEAR GILTS (PERIOD AVERAGE) NADJ
FTSE ALL SHARE - PRICE INDEX
UK DCLG HOUSE PRICE INDEX (MIX ADJ.) NADJ
UK corporate bond yield

(Source: Thomson-Reuters Datastream)
C   List of variables in the non-factor-augmented BVAR

UK GDP AT MARKET PRICES (CVM) CONA
UK LFS: UNEMPLOYMENT RATE, ALL, AGED 16 & OVER SADJ
UK DCLG HOUSE PRICE INDEX (MIX ADJ.) NADJ
FTSE ALL SHARE - PRICE INDEX
UK BANK OF ENGLAND BASE RATE (EP)
UK 3 MONTHS TREASURY BILLS YIELD (EP)
UK GROSS REDEMPTION YIELD ON 10 YEAR GILT EDGED STOCKS(AVE)
UK GROSS REDEMPTION YIELD ON 20 YEAR GILTS (PERIOD AVERAGE) NADJ
UK M4 SECT.ANALYSIS: M4 LEND INDIV.: DWELLS,SECURED LOANS: LEVEL
UK HN: REAL HOUSEHOLDS DISPOSABLE INCOME CONA
UK MONETARY AGGREGATE M4 CURA
References


