



Strathprints Institutional Repository

Koop, Gary and Korobilis, Dimitris (2014) A new index of financial conditions. *European Economic Review*, 71. pp. 101-116. ISSN 0014-2921 , <http://dx.doi.org/10.1016/j.euroecorev.2014.07.002>

This version is available at <http://strathprints.strath.ac.uk/55408/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to Strathprints administrator: strathprints@strath.ac.uk

A New Index of Financial Conditions*

Gary Koop
University of Strathclyde

Dimitris Korobilis[†]
University of Glasgow

July 4, 2014

Abstract

We use factor augmented vector autoregressive models with time-varying coefficients and stochastic volatility to construct a financial conditions index that can accurately track expectations about growth in key US macroeconomic variables. Time-variation in the model's parameters allows for the weights attached to each financial variable in the index to evolve over time. Furthermore, we develop methods for dynamic model averaging or selection which allow the financial variables entering into the financial conditions index to change over time. We discuss why such extensions of the existing literature are important and show them to be so in an empirical application involving a wide range of financial variables.

Keywords: Bayesian model averaging; dynamic factor model; dual Kalman filter; forecasting

JEL Classification: C11, C32, C52, C53, C66

*We want to thank Gianni Amisano, Klemens Hauzenberger, Sylvia Kaufmann, Michael McCracken, Francesco Ravazzolo, and Christian Schumacher for helpful comments. Special acknowledgements to Barbara Rossi for her extensive and insightful suggestions. Finally we want to thank participants at: the 21th Symposium of the Society for Nonlinear Dynamics and Econometrics in Milan; the Conference Forecasting Structure and Time Varying Patterns in Economics and Finance in Rotterdam; the 7th International Conference on Computational and Financial Econometrics in London, and seminar attendees at Birkbeck College, Norges Bank, and Université de Rennes 1. This research was supported by the ESRC under grant RES-062-23-2646.

[†]Corresponding author. Address: University of Glasgow, Adam Smith Business School, Gilbert Scott building, Glasgow, G12 8QQ, United Kingdom, Tel: +44 (0)141 330 2950, e-mail: Dimitris.Korobilis@glasgow.ac.uk

1 Introduction

The recent financial crisis has sparked an interest in the accurate measurement of financial shocks to the real economy. An important lesson of recent events is that financial developments, not necessarily driven by monetary policy actions or fundamentals, may have a strong impact on the economy. The need for policy-makers to closely monitor financial conditions is clear. In response to this need, a recent literature has developed several empirical econometric methods for constructing financial conditions indices (FCIs). FCIs are used for several purposes. For instance, they can be used to identify periods when financial conditions suddenly deteriorate (e.g., Lo Duca and Peltonen, 2013), assess credit constraints or forecast economic developments. An FCI summarizes in one single number information from many financial variables. Many financial institutions (e.g. Goldman Sachs, Deutsche Bank and Bloomberg) and policy-makers (e.g. the Federal Reserve Bank of Kansas City) produce closely-watched FCIs. Estimation of such FCIs ranges from using simple weighted averages of financial variables through more sophisticated econometric techniques. An important recent contribution is Hatzius, Hooper, Mishkin, Schoenholtz and Watson (2010) which surveys and compares a variety of different approaches. The FCI these authors propose is based on simple principal components analysis of a very large number of quarterly financial variables. Other recent notable studies in this literature include English, Tsatsaronis and Zoli (2005), Balakrishnan, Danninger, Elekdag and Tytell (2008), Beaton, Lalonde and Luu (2009), Brave and Butters (2011), Gomez, Murcia and Zamudio (2011) and Matheson (2011).

In this paper our goal is to accurately monitor financial conditions through a single latent FCI. The construction and use of an FCI involves three issues: i) selection of financial variables to enter into the FCI, ii) the weights used to average these financial variables into an index and iii) the relationship between the FCI and the macroeconomy. There is good reason for thinking all of these may be changing over time. Indeed, Hatzius et al (2010) discuss at length why such change might be occurring and document statistical instability in their results. For instance, the role of the sub-prime housing market in the financial crisis provides a clear reason for the increasing importance of variables reflecting the housing market in an FCI. A myriad of other changes may also impact on the way an FCI is constructed, including the change in structure of the financial industry (e.g. the growth of the shadow banking system), changes in the response of financial variables to changes in monetary policy (e.g. monetary policy works differently with interest rates near the zero bound) and the changing impact of financial variables on real activity (e.g. the role of financial variables in the recent recession is commonly considered to have been larger than in other recessions).

Despite such concerns about time-variation, the existing literature does little to statistically model it. Constant coefficient models are used with, at most, rolling

methods to account for time-variation. Furthermore, many FCI's are estimated *ex post*, using the entire data set. So, for instance, at the time of the financial crisis, some FCIs will be based on financial variables which are selected after observing the financial crisis and the econometric model will be estimated using financial crisis data. The major empirical contribution of the present paper is to develop an econometric approach which allows for different financial variables to affect estimation of the FCI, with varying (or zero, when not selected) weight each. In this manner, we develop an econometric tool that explicitly takes into account the fact that each financial crisis has different causes, and is transmitted to the real economy with varying intensity.

Following a common practice in constructing indices, we use factor methods. To be precise, we use extensions of Factor-augmented VARs (FAVARs) which jointly model a large number of financial variables (used to construct the latent FCI) with key macroeconomic variables. Following the recent trend in macroeconomic modelling using VARs and FAVARs (Primiceri, 2005; Korobilis, 2013) we work with time-varying parameter FAVARs (TVP-FAVARs) which allow coefficients and loadings to change in each period. TVP-FAVARs have enjoyed increasing popularity for forecasting macroeconomic variables (see, among others, Eickmeier, Lemke and Marcellino, 2011a and D'Agostino, Gambetti and Giannone, 2013).

Additionally, we work with a large set of (TVP-) FAVARs that differ in which financial variables are included in the estimation of the FCI. Faced with a large model space and the desire to allow for model change, we follow Koop and Korobilis (2012) and use efficient methods for Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA). These methods forecast at each point in time with a single optimal model (DMS), or reduce the expected risk of the final forecast by averaging over all possible model specifications (DMA). We implement model selection or model averaging in a dynamic manner. That is, DMS chooses different financial variables to make up the FCI at different points in time. DMA constructs an FCI by averaging over many individual FCIs constructed using different financial variables. The weights in this average vary over time.

From an econometrician's point of view, there is also growing theoretical evidence in favor of our modelling strategy. Boivin and Ng (2006) show that using all available data to extract factors (the FCI in our case) is not always optimal in factor analysis, thus providing support for implementing DMA/DMS to construct our FCI. Additionally, there is much econometric evidence in favor of structural instabilities in the coefficients or loadings of macroeconomic and financial factor models; see, among others, Banerjee, Marcellino and Masten (2006) and Bates, Plagborg-Møller, Stock and Watson (2013).

Econometric methods for estimating FAVARs and TVP-FAVARs are well-established; see, e.g., Bernanke, Boivin and Elias (2005), and Korobilis (2013). However, the likelihood-based estimation techniques used in the literature (e.g. Bayesian methods using Markov chain Monte Carlo algorithms) rely on simulation algo-

rithms or complex numerical methods, all of which are computationally extremely demanding in high dimensions. With our large model space, and our wish to implement recursive forecasting, it is computationally infeasible to use such methods. Therefore, our major econometric contribution in this paper lies in the development of fast estimation methods which are based on the Kalman filter and smoother and are simulation-free. When dealing with the FAVAR with constant parameters, our algorithm collapses to the two-step estimator for dynamic factor models of Doz, Giannone and Reichlin (2011). In the case of estimating models with time-varying parameters and stochastic volatility (TVP-FAVARs), our algorithm provides an extension of Doz, Giannone and Reichlin (2011).

Our results indicate that financial variables do have predictive power for macroeconomic variables (GDP growth, inflation and unemployment). Additionally, time variation in the parameters is important for providing accurate short-run forecasts. Finally, model averaging and/or selection also result in the improvement of forecast accuracy over using a single model with all the available financial variables. In the remainder of the paper we examine all these issues in depth, and we provide evidence by using different forecast metrics and by conducting several robustness checks.

In particular, in the next section we introduce our modeling framework and sketch the features of our novel estimation algorithm (complete details are provided in the Technical Appendix), plus we describe how we implement DMA or DMS methods in the face of the large number of models we work with. In Section 3 we present our data, estimates of different FCIs, and results of a recursive forecasting exercise which is the main tool for evaluating the performance of our FCI. Section 4 concludes the paper. An empirical appendix provides an extensive sensitivity analysis to various aspects of our specification.

2 Factor Augmented VARs with Structural Instabilities

2.1 The TVP-FAVAR Model and its Variants

Let x_t (for $t = 1, \dots, T$) be an $n \times 1$ vector of financial variables to be used in constructing the FCI. Let y_t be an $s \times 1$ vector of macroeconomic variables of interest. In our empirical work $y_t = (\pi_t, u_t, g_t)'$ where π_t is the GDP deflator inflation rate, u_t is the unemployment rate, g_t is the growth rate of real GDP. The p -lag TVP-FAVAR takes the form

$$x_t = \lambda_t^y y_t + \lambda_t^f f_t + v_t$$

$$\begin{bmatrix} y_t \\ f_t \end{bmatrix} = c_t + B_{t,1} \begin{bmatrix} y_{t-1} \\ f_{t-1} \end{bmatrix} + \dots + B_{t,p} \begin{bmatrix} y_{t-p} \\ f_{t-p} \end{bmatrix} + \varepsilon_t, \quad (1)$$

where λ_t^y are regression coefficients, λ_t^f are factor loadings, f_t is the latent factor which we interpret as the FCI, c_t is a vector of intercepts, $(B_{t,1}, \dots, B_{t,p})$ are VAR coefficients and u_t and ε_t are zero-mean Gaussian disturbances with time-varying covariances V_t and Q_t , respectively. We adopt the common identifying assumption in the likelihood-based factor literature¹ that V_t is diagonal, thus ensuring that v_t is a vector of idiosyncratic shocks and f_t contains information common to the financial variables. This model is very flexible since it allows all parameters to take a different value at each time t . Such an assumption is important since there is good reason to believe that there is time variation in the loadings and covariances of factor models which use both financial and macroeconomic data (see Banerjee, Marcellino and Masten, 2006). For recent discussions about the implication of the presence of structural breaks in factor loadings, the reader is referred to Breitung and Eickmeier (2011) and Bates, Plagborg-Møller, Stock and Watson (2013).

Following the influential work of Bernanke, Boivin and Elias (2005) our factor model in (1) consists of two equations: one equation which allows us to extract the latent financial conditions index (FCI) from financial variables x_t ; and one equation which allows to model the dynamic interactions of the FCI with macroeconomic variables y_t . This econometric specification is important for two reasons. First, unlike Stock and Watson (2002) who extract a factor and then use it in a separate univariate forecasting regression, we use a multivariate system to forecast macroeconomic variables using the FCI. Thus, we jointly model all the variables in the system which should allow us to better characterize their co-movements and interdependence. Second, we are able to purge from the FCI the effect of macroeconomic conditions. Thus, the final estimated FCI reflects information solely associated with the financial sector.

It is worth digressing to expand on the manner in which we include macroeconomic variables in our model. Including y_t on the right-hand side of the first equation of (1) is intended solely to ensure the FCI reflects only financial conditions. This is also done in Hatzius et al. (2010) for the same reason. However, it is worth stressing that, by doing this, we are only purging the FCI of the effect of current macroeconomic conditions. Financial variables can also reflect expectations of future macroeconomic variables and we are not purging the FCI of these future expectations. This is an issue common to all FCIs. As a robustness check, our empirical results (see Appendix C) also include a case where y_t in the first equation of (1) is replaced by professional forecasts of macroeconomic variables (which will reflect expectations of future macroeconomic conditions).

Including y_t on the left-hand side of the second equation of (1) is done so as to provide a metric for evaluating our FCI. That is, in answer to the question: “what makes a good FCI?”, the approach in this paper provides the answer: “it is one which forecasts y_t as well as possible”. There are, of course, other possible answers

¹Some approaches to dynamic factor models which do not use likelihood-based methods allow for weak correlations between the elements of v_t .

to this question which would lead to other FCIs. We are not attempting to build a structural model of the economy (e.g. a structural VAR or a DSGE) through the manner we are including y_t in (1). Hence, issues which arise with structural models (e.g. structural VARs often involve an ordering of variables reflecting an assumed causal structure) do not need to be addressed here.

In order to complete our model, we need to define how the time varying parameters evolve. While the specification of all time-varying covariances is discussed in the following subsection, we define here the vectors of loadings $\lambda_t = \left((\lambda_t^y)', (\lambda_t^f)' \right)'$ and VAR coefficients $\beta_t = (c_t', \text{vec}(B_{t,p})', \dots, \text{vec}(B_{t,p})')'$ to evolve as multivariate random walks of the form

$$\begin{aligned}\lambda_t &= \lambda_{t-1} + v_t, \\ \beta_t &= \beta_{t-1} + \eta_t,\end{aligned}\tag{2}$$

where $v_t \sim N(0, W_t)$ and $\eta_t \sim N(0, R_t)$. Finally, all disturbance terms presented in the equations above are uncorrelated over time and with each other.

We call the full model described in equations (1) and (2) the TVP-FAVAR. We also consider several restrictions on the TVP-FAVAR which result in other popular multivariate models:

1. Factor-augmented VAR (FAVAR): This model is obtained from the TVP-FAVAR under the restriction that both λ_t and β_t are time-invariant ($W_t = R_t = 0$).
2. Time-varying parameter VAR (TVP-VAR): This model can be obtained from the TVP-FAVAR under the restriction that the number of factors is zero (i.e. $f_t = 0$).
3. VAR: This model is obtained when the number of factors is zero and both λ_t and β_t are time-invariant.

In addition, we consider heteroskedastic (when V_t and Q_t are allowed to be time-varying) and homoskedastic ($V_t = V$ and $Q_t = Q$) variants of all models.

2.2 Estimation of a Single TVP-FAVAR

Bayesian estimation of FAVARs (as well as VARs) with time-varying parameters is typically implemented using Markov Chain Monte Carlo (MCMC) methods, which sample from the very complex multivariate joint posterior density of the factor f_t and the remaining model parameters; see, e.g., Primiceri (2005), or Del Negro and Otrok (2008). Such Bayesian simulation methods are computationally expensive

even in the case of estimating a single TVP-FAVAR. When faced with multiple TVP-FAVARs and when doing recursive forecasting (which requires repeatedly doing MCMC on an expanding window of data), the use of MCMC methods is prohibitive.²

In this paper, we use a fast two-step estimation algorithm which vastly reduces the computational burden, and greatly simplifies the estimation of the FCI. Following Koop and Korobilis (2013) we combine the ideas of variance discounting methods with the Kalman filter in order to obtain analytical results for the posteriors of the state variable (f_t) as well as the time-varying parameters $\theta_t = (\lambda_t, \beta_t)$. To motivate our methods, note first that, as long as both the factor, f_t , and the loadings, λ_t , in the measurement equation are unobserved, application of the typical Kalman filter recursions for state-space models is not possible. Therefore, we adapt ideas from Doz, Giannone and Reichlin (2011) and the state-space literature (Nelson and Stear, 1976) and develop a dual, conditionally linear filtering/smoothing algorithm which allows us to estimate the unobserved state f_t and the parameters $\theta_t = (\lambda_t, \beta_t)$ in a fraction of a second.

The idea of using a dual linear Kalman filter is very simple: first update the parameters θ_t given an estimate of f_t , and subsequently update the factor f_t given the estimate of θ_t . Such conditioning allows us to use two distinct linear Kalman filters or smoothers,³ one for θ_t and one for f_t . The main approximation involved is that \tilde{f}_t , the principal components estimate of f_t based on $x_{1:t}$, is used in the estimation of θ_t . Such an approach will work best if the principal component(s) provide a good approximation of the factor(s) coming from a FAVAR with structural instabilities. A theoretical proof that this is the case is not available for our flexible and highly nonlinear specification. However, given the recent findings of Stock and Watson (2009) and Bates, Plagborg-Møller, Stock and Watson (2013), there is strong theoretical and empirical evidence to believe that this is the case. Empirically, Bates, Plagborg-Møller, Stock and Watson (2013) conduct extensive Monte Carlo experiments and show that principal components can support large amount of time variation in the loading coefficients λ_t . Theoretically, they prove that principal component estimates of factors are consistent even if there is a substantial amount of time variation or structural change in the factor loadings. For instance, they find that: “deviations in the factor loadings on the order of $o_p(1)$ [as would occur with random walk variation in the factor loadings] do not break the consistency of the principal components estimator” (page 290). This is the econometric theory

²To provide the reader with an idea of approximate computer time, consider the three variable TVP-VAR of Primiceri (2005). Taking 10,000 MCMC draws (which may not be enough to ensure convergence of the algorithm) takes approximately 1 hour on a good personal computer. Thus, forecasting at 100 points in time takes roughly 100 hours. These numbers hold for a single small TVP-VAR, and would be much infeasible for the hundreds of thousands of larger TVP-FAVARs we estimate in this paper.

³The other alternative being to use a joint nonlinear filter, e.g. the Unscented Kalman Filter (UKF) and the Extended Kalman Filter (EKF). We have found such filters to be very unstable given the dimension of our model, and the relatively few time-series observations.

we draw upon to justify use of principal components estimates at this stage in our algorithm.

Error covariance matrices in the multivariate time series models used with macroeconomic data are usually modeled using multivariate stochastic volatility models (see, e.g., Primiceri, 2005), estimation of which also requires computationally intensive methods. In order to avoid this computational burden, we estimate (V_t, Q_t, W_t, R_t) recursively using simulation-free variance matrix discounting methods (e.g. Quintana and West, 1988). The Technical Appendix provides complete details. For V_t and Q_t we use exponentially weighted moving average (EWMA) estimators. These depend on decay factors κ_1 and κ_2 , respectively. Such recursive estimators are trivial computationally. Additionally, the EWMA is an accurate approximation to an integrated GARCH model. Such a feature is in line with authors such as Primiceri (2005) and Cogley and Sargent (2005) who, in the context of macroeconomic VARs, work with integrated stochastic volatility models. The covariance matrices W_t, R_t are estimated using the forgetting factor methods described in Koop and Korobilis (2012, 2013)⁴ which depend on forgetting factors κ_3 and κ_4 , respectively. Decay and forgetting factors have very similar interpretations. Lower values of the decay/forgetting factors imply that the more recent observation $t-1$, and its squared residual, take higher weight in estimating V_t and Q_t compared to older observations. The choice of the decay/forgetting factors can be made based on the expected amount of time-variation in the parameters.⁵ Note that the choice $\kappa_1 = \kappa_2 = 1$ make V_t and Q_t constant, while $\kappa_3 = \kappa_4 = 1$ imply that $W_t = R_t = 0$ in which case λ_t and β_t are constant.

A sketch of the structure of our estimation algorithm is given in the following table.

Algorithm for estimation of the TVP-FAVAR

1. a) Initialize all parameters, $\lambda_0, \beta_0, f_0, V_0, Q_0$
 - b) Obtain the principal components estimates of the factors, \tilde{f}_t
 2. Estimate the time varying parameters θ_t given \tilde{f}_t
 - a) Estimate V_t, Q_t, R_t , and W_t using VD
 - b) Estimate λ_t and β_t , given (V_t, Q_t, R_t, W_t) , using the KFS
 3. Estimate the factors f_t given θ_t using the KFS
-

In the table VD stands for “Variance Discounting” and KFS stands for “Kalman filter and smoother”. The steps above can also be considered to be a generalization of

⁴An EWMA estimation scheme can also be applied to these matrices, but due to their large dimension we found better numerical stability and precision when using forgetting factors.

⁵Choice of forgetting factors is similar in spirit to choice of prior. Empirical macroeconomists frequently impose subjective priors on the degree of time variation in their parameters; see for instance the very informative priors used in the TVP-VARs of Primiceri (2005) and Cogley and Sargent (2005).

the estimation steps introduced by Doz et al. (2011) for the estimation of constant parameter dynamic factor models. In fact, if we fix all time-varying coefficients and covariances to be constant, our algorithm collapses to the FAVAR equivalent of the two-step estimation algorithm for dynamic factor models of Doz et. al (2011).

Identification in the FAVAR is achieved in a standard fashion by restricting V_t to be a diagonal matrix. This restriction ensures that the factors, f_t , capture movements that are common to the financial variables, x_t , after removing the effect of current macroeconomic conditions through inclusion of the $\lambda_t^y y_t$ term. Further restrictions usually imposed in likelihood-based estimation of factor models, e.g. normalizing the first element of the loadings matrix to be 1 (Bernanke, Boivin and Elias, 2005) are not needed here since the loadings λ_t are identified (up to a sign rotation) from the principal components estimate of the factor.

2.3 Dynamic Model Averaging and Selection with many TVP-FAVARs

In this paper, we work with M_j , $j = 1, \dots, J$, models which differ in the financial variables which enter the FCI. In other words, a specific model is obtained using the restriction that a specific combination of financial variables have zero loading on the factor at time t or, equivalently, that different combinations of columns of x_t are set to zero. Thus, M_j can be written as

$$x_t^{(j)} = \lambda_t^{y(j)} y_t + \lambda_t^{f(j)} f_t^{(j)} + u_t^{(j)}$$

$$\begin{bmatrix} y_t \\ f_t^{(j)} \end{bmatrix} = c_t^{(j)} + B_{t,1}^{(j)} \begin{bmatrix} y_{t-1} \\ f_{t-1}^{(j)} \end{bmatrix} + \dots + B_{t,p}^{(j)} \begin{bmatrix} y_{t-p} \\ f_{t-p}^{(j)} \end{bmatrix} + \varepsilon_t^{(j)}, \quad (3)$$

where $x_t^{(j)}$ is a subset of x_t , and $f_t^{(j)}$ is the FCI implied by model M_j . Since x_t is of length n , there is a maximum of $2^n - 1$ combinations⁶ of financial variables that can be used to extract the FCI.

When faced with multiple models, it is common to use model selection or model averaging techniques. However, in the present context we wish such techniques to be dynamic. That is, in a model selection exercise, we want to allow for the selected model to change over time, thus doing DMS. In a model averaging exercise, we want to allow for the weights used in the averaging process to change over time, thus leading to DMA. In this paper, we do DMA and DMS using an approach developed in Raftery et al (2010) in an application involving many TVP regression models. The reader is referred to Raftery et al (2010) for a complete derivation and motivation of DMA. Here we provide a general description of what it does.

The goal is to calculate $\pi_{t|t-1,j}$ which is the probability that model j applies at time t , given information through time $t - 1$. Once $\pi_{t|t-1,j}$ for $j = 1, \dots, J$ are obtained they can either be used to do model averaging or model selection. DMS

⁶We remove from the model set the model with zero financial variables, i.e. with no FCI extracted.

arises if, at each point in time, the model with the highest value for $\pi_{t|t-1,j}$ is used. Note that $\pi_{t|t-1,j}$ will vary over time and, hence, the selected model can switch over time. DMA arises if model averaging is done in period t using $\pi_{t|t-1,j}$ for $j = 1, \dots, J$ as weights. The contribution of Raftery et al (2010) is to develop a fast recursive algorithm for calculating $\pi_{t|t-1,j}$.

Given an initial condition, $\pi_{0|0,j}$ for $j = 1, \dots, J$, Raftery et al (2010) derive a model prediction equation using a forgetting factor α :

$$\pi_{t|t-1,j} = \frac{\pi_{t-1|t-1,j}^\alpha}{\sum_{l=1}^J \pi_{t-1|t-1,l}^\alpha}, \quad (4)$$

and a model updating equation of:

$$\pi_{t|t,j} = \frac{\pi_{t|t-1,j} f_j (Data_t | Data_{1:t-1})}{\sum_{l=1}^J \pi_{t|t-1,l} f_l (Data_t | Data_{1:t-1})}, \quad (5)$$

where $f_j (Data_t | Data_{1:t-1})$ is a measure of fit for model j .⁷ Many possible measures of fit can be used. Inspired by is a large literature (e.g., among many others, Forni, Hallin, Lippi and Reichlin, 2003) which investigate the ability of financial variables to forecast macroeconomic ones, we focus on the ability of the FCI to forecast y_t . Accordingly, we set as a measure of fit the predictive likelihood for the macroeconomic variables, $p_j (y_t | Data_{1:t-1})$. α is a forgetting factor with $0 < \alpha \leq 1$ which, similar to the decay/forgetting factors ($\kappa_1, \kappa_2, \kappa_3, \kappa_4$) used for estimating the error covariance matrices, tunes how rapidly switches between models should occur. Lower values of α allow for an increasing amount of switching between the number of variables that enter the FCI each time period. If $\alpha = 0.99$, forecast performance five years ago receives 80% as much weight as forecast performance last period. The case $\alpha = 1$ leads to conventional Bayesian model averaging implemented on an expanding window of data.

3 Empirical Results

3.1 Data and Model Settings

We use 18 financial variables⁸ which cover a wide variety of financial considerations (e.g. asset prices, volatilities, credit, liquidity, etc.). These are gathered from several

⁷Throughout this paper past data up to time t will be denoted by $1 : t$ subscripts, e.g., $Data_{1:t} = (Data_1, \dots, Data_t)$.

⁸We have gathered a range of widely-used financial variables, following the recommendations of Hatzius et al. (2010). The final number of variables, though, was restricted by computational constraints. As we explain in the Technical Appendix, availability of computer clusters (such as the ones maintained by central banks, e.g. the ECB) could allow for the implementation of model averaging with many more variables.

sources. All of the variables (i.e. both macroeconomic and financial variables) are transformed to stationarity following Hatzius et al (2010) and many others. The Data Appendix provides precise definitions, acronyms, data sources, sample spans and details about the transformations. Our data sample runs from 1970Q1 to 2013Q3. Notice that all of our models use four lags and, hence, the effective estimation sample begins in 1971Q1. The three macroeconomic variables that complete our model are the GDP deflator, the unemployment rate, and the real GDP. We use real time data such that forecasts are at time t are always made using the vintage of data available at time t . All of these series are observed in real-time from 1970Q1, are seasonally adjusted, and can be found in the Real-Time Data Set for Macroeconomists provided by the Philadelphia Fed. Macroeconomic variables which are not already in rates, that is the GDP deflator and real GDP, are converted to growth rates by taking first log-differences, which we will refer to as inflation and output growth, respectively.

Some of the financial variables have missing values in that they do not begin until much after 1970Q1. In terms of estimation with a single TVP-FAVAR model, such missing values cause no problem since they can easily be handled by the Kalman filter (see the Technical Appendix for more details). However, when we are using multiple models, there is a danger that in a specific model the value of the FCI in a period (say 1970Q1-1982Q1) has to be extracted using financial variables which all have missing values for that period. In such a case, the value of extracted FCI will be nil for the specified period, and the FCI will be estimated only after at least one variable becomes observed. We introduce a simple restriction to prevent such estimation issues. In each model, we always include the S&P500 in the list of financial variables, a variable which is observed since 1970Q1. This means that, at a minimum, the FCI will be extracted based on this financial variable. This restriction implies that the S&P500 is not subject to model averaging/selection and we instead perform DMA/DMS using the remaining 17 financial variables. Therefore, we have a model space of $2^{17} = 131,072$ TVP-FAVARs. We remind the reader that a list of the different specifications estimated (and their acronyms) is given at the end of Section 2.1.

To summarize, our models which produce an FCI are the TVP-FAVARs and the FAVARs. In our forecasting exercise, for the purpose of comparison, we also include some forecasting models which do not produce an FCI. These are the VARs and TVP-VARs. With these model spaces, we investigate the use of DMS, DMA and a strategy of simply using the single model which includes all 18 of the financial variables.

Some authors (e.g. Eickmeier, Lemke and Marcellino, 2011b) use existing FCIs (i.e. estimated by others) in the context of a VAR model. In this spirit, we also present results for VARs and TVP-VARs where the factors are not estimated from the factor model equation in (1), rather they are replaced with an existing estimate. To be precise, we use $\left(y'_t, \hat{f}_t\right)'$ as dependent variables for different choices of \hat{f}_t . Table 1 lists these choices from a set of financial conditions and financial stress indices

maintained by Federal Reserve Banks. Financial Stress indices (FSIs) are similar to FCIs, but have opposing signs: a decrease in financial conditions means increased financial stress, and vice-versa. However, FSIs tend to focus on different issues than FCIs. The latter tend to focus on broad measures of financial conditions, whereas FSIs narrowly focus on measuring instability in the financial system (i.e. the current level of frictions, stresses and strains). For this reason, we would argue that the Chicago Fed National FCI is the most comparable to the FCI we are producing in our empirical results. Again, these models are a restricted special case of our TVP-FAVAR and estimation proceeds accordingly. The error covariance matrix is modelled in the same manner as the TVP-FAVAR. We use an acronym for these TVP-VARs such that, e.g., TVP-VAR + FCI 4, is the VAR involving the three macroeconomic variables and the Chicago Fed National FCI.

Table 1. Financial Conditions and Stress indices

Name	Acronym	Source	Sample
St. Louis Financial Stress Index	FCI 1	St Louis Fed	1993Q4 - 2012Q1
Kansas City Fed Financial Stress Index	FCI 2	Kansas Fed	1990Q1 - 2012Q1
Cleveland Fed Financial Stress Index	FCI 3	Cleveland Fed	1991Q3 - 2012Q1
Chicago Fed National FCI	FCI 4	Chicago Fed	1973Q1 - 2012Q1

3.2 Choice of hyperparameters and initial conditions

In this section we outline the setting of various hyperparameters and initial conditions. All our benchmark choices that apply in the next two subsections are fairly non-informative. In Appendix C we implement a sensitivity analysis using priors based on a training data sample, thus extending the recommendations of Primiceri (2005) to our TVP-FAVARs.

The first step is to set the initial conditions for the factor f_t (FCI), the time-varying parameters λ_t, β_t , the time-varying covariances V_t, Q_t , and, for doing DMA and DMS, we must specify $\pi_{0|0,j}$, $j = 1, \dots, J$. These initial conditions are set to the following (relatively non-informative) values

$$\begin{aligned}
 f_0 &\sim N(0, 4), \\
 \lambda_0 &\sim N(0, 4 \times I_{n(s+1)}), \\
 \beta_0 &\sim N(0, V_{MIN}) \\
 V_0 &\equiv 1 \times I_n, \\
 Q_0 &\equiv 1 \times I_{s+1}, \\
 \pi_{0|0,j} &= \frac{1}{J}
 \end{aligned}$$

where V_{MIN} is a diagonal covariance matrix which, following the Minnesota prior tradition, penalizes more distant lags and is of the form

$$V_{MIN} = \begin{cases} 4, & \text{for intercepts} \\ 0.1/r^2, & \text{for coefficient on lag } r \end{cases}, \quad (6)$$

where $r = 1, \dots, p$ denotes the lag number. Note that estimates of W_t and R_t are proportional to the respective state covariance matrices obtained from the Kalman filter, therefore there is no need to initialize these matrices; see the Technical Appendix for more details.

Regarding the decay and forgetting factors we have introduced in our model it is worth noting that we can estimate these from the data. However, computation increases substantially (we need to evaluate or maximize the predictive likelihood for each combination of the various factors) and, as shown in Koop and Korobilis (2013), the existence of value added in forecasting performance from such a procedure is questionable. Given these considerations, we choose to fix the values of the decay/forgetting factors, but investigate sensitivity to their choice in Appendix C.

For the decay factors κ_1, κ_2 which control the variation in the covariance matrices, we fix these to the value 0.96. Such values provide volatility estimates which are quite close to the ones expected by integrated stochastic volatility models that have been used extensively in the Bayesian VAR and FAVAR literature (Primiceri, 2005; Korobilis, 2013). For the forgetting factors κ_3, κ_4 , we follow the “business as usual prior” approach of Cogley and Sargent (2005) and assume that changes each period are relatively slow and stable under the random walk specification in equation (2). In order to achieve this slow time variation in the coefficients, we set $\kappa_3 = \kappa_4 = 0.99$, a setting we use in all TVP-FAVAR and TVP-VAR specifications. As described in the Technical Appendix, restricted versions of our general model can be obtained by setting the forgetting factors to one. For instance when $\kappa_3 = \kappa_4 = 1$, we obtain the VAR model.

Finally, we need to choose our prior beliefs about model change. The value of the forgetting factor α determines how fast model switches occur, and thus we use two values: $\alpha = 1$ which implies that we are implementing Bayesian model averaging (BMA) given data up to time t ; and $\alpha = 0.99$ which implies that we implement dynamic model averaging (DMA) with relatively slowly varying model probabilities.

3.3 Estimates of the Financial Conditions Index

Before we proceed to the forecasting exercise, it is important to understand how both our estimation algorithm and model averaging work in the context of estimating an FCI. The results in this section are recursive. That is the estimate of the FCI at time t is made using data up to time t .

Figure 1 shows the FCI estimated in various ways using all 18 financial variables without any model selection or model averaging being done. The shaded regions in this figure (and subsequent figures) are the NBER recession dates. The estimated FCIs are all similar to each other. In particular, the TVP-FAVAR and principle components are producing very similar FCIs. The FCI produced by the FAVAR does differ from the others at some points, particularly in the first half of the sample. This indicates the potential importance of allowing for time variation in parameters. It is interesting to note that the FCIs start declining before the beginning of the recent recession with all of them bottoming out in early 2009. However, after 2009 some discrepancies appear between the FCIs.

Figure 2 shows the impact of model averaging and selection on the estimate of the FCI, focussing on the TVP-FAVARs. Although the broad patterns in the FCIs plotted in Figure 2 are similar, there are some differences. In general, the FCIs in Figure 2 are less smooth than in Figure 1 indicating that model switching/averaging is reacting more quickly than methods without such a feature. And there are some interesting small divergences between the two figures. For instance, in Figure 2 there is a slight improvement in the FCIs early on in the recent recession which is missing from Figure 1. DMA and DMS are producing factor estimates which are very similar to one another.

In Figure 3 we perform a comparison of the FCI constructed using DMA on the TVP-FAVAR models with the four FCIs or FSIs maintained by four Federal Reserve Banks (see Table 1). As discussed previously, FCIs and FSIs are somewhat different since FSIs are measuring financial stress (and, hence, it is the comparison of our approach with the Chicago Fed's National Financial Conditions Index which is the most relevant). We have multiplied the FSIs by minus one to maintain comparability with the FCIs. Additionally, we standardize all the indices to have mean zero and variance one. Our FCI does indeed match up most closely with the Chicago Fed's FCI. Our FCI started dropping earlier than the Chicago Fed's in 2008 and dropped to a lower value at the depth of the current recession. In contrast, after the 2001 recession ended, our FCI grew faster and peaked at a higher level than the Chicago Fed's. However, there are substantive differences with the FSIs. It is interesting to note that, after this 2001 recession ended, the FSIs continued to signal deteriorating financial conditions for much longer than our FCI. In general, the FSIs (in particular, the Cleveland Fed's FCI) exhibit substantively different behavior from our FCI.

Figures 1 through 3 compare a range of different FCI estimates. At this stage, we express no view on whether any FCI is better or worse than any other. The key finding we stress is that, although they are similar to one another in many respects, differences can occur. These differences are most notable when we compare our TVP-FAVAR based estimates to conventional estimates produced by Federal Reserve Banks.

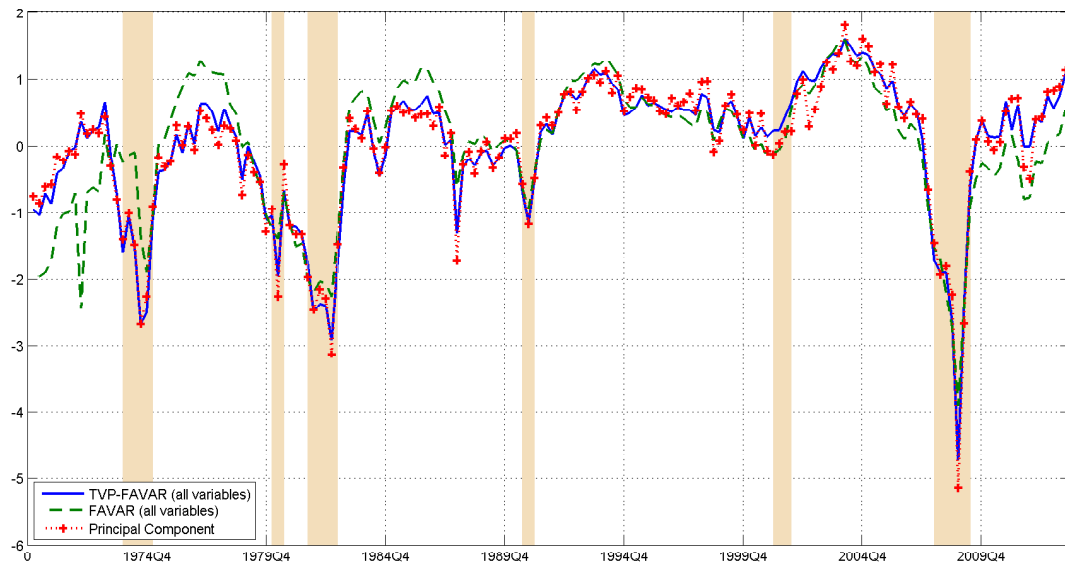


Figure 1. FCIs constructed from several versions of the heteroskedastic factor-augmented VAR model with all 18 financial variables used (no model averaging/selection). For comparison, the principal component of the 18 financial variables is also plotted.

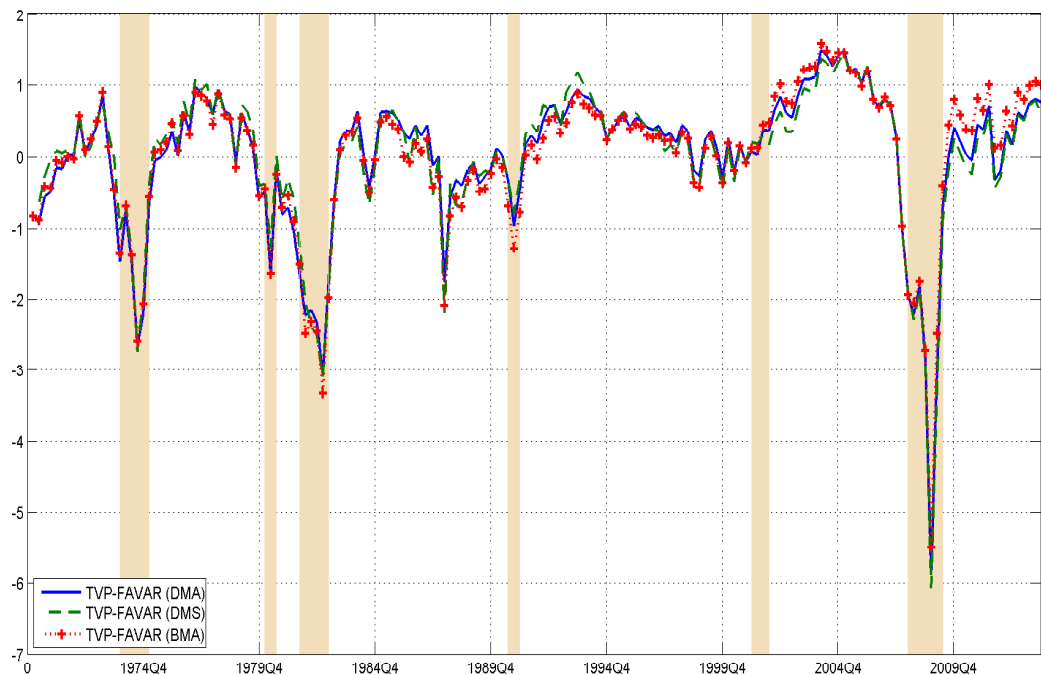


Figure 2. FCIs implied by BMA, DMA and DMS on the TVP-FAVARs (with DMA results for FAVARs provided for comparison)

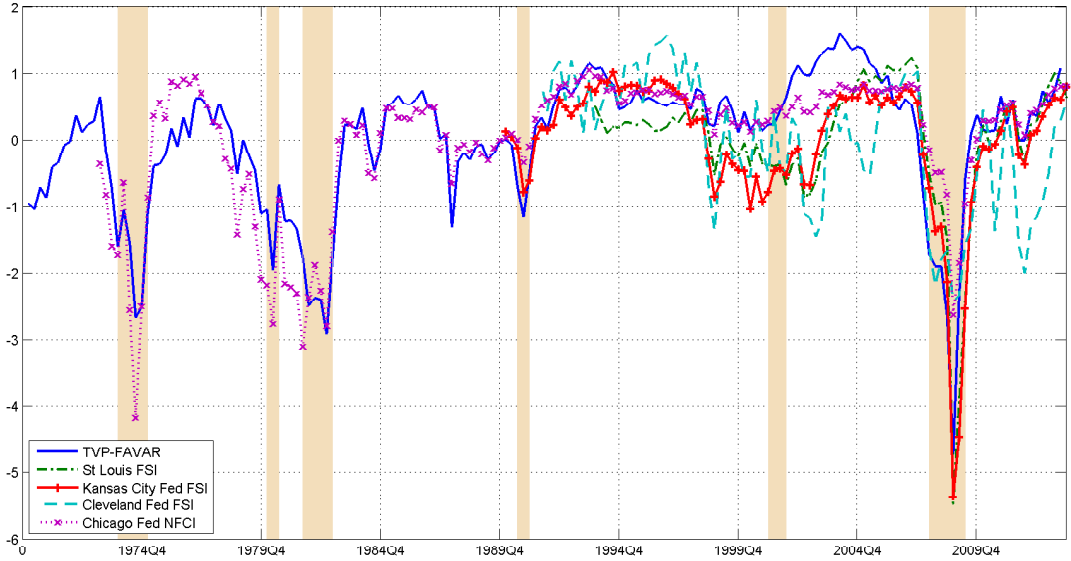


Figure 3. The FCI from the TVP-FAVAR with DMA compared to existing financial indexes maintained by four regional US central banks.

To provide some additional insight on what DMA is doing, we present Figures 4 and 5 which shed light on the number of variables selected when we do DMA or DMS on the TVP-FAVARs. In particular, Figure 4 calculates the expected number of variables used to extract the FCI at each point in time. If we denote by n_j the number of variables which load on the FCI under model M_j , then we calculate each time period the following expectation⁹

$$E(n_t^{DMA}) = \left(\sum_{j=1}^J \pi_{t|t,j} \times n_j \right) - 1.$$

Figure 4 shows DMA or DMS is achieving a strong degree of parsimony. Given 17 variables to choose between, it is tending to choose between 5 and 8. There is a slow decrease in the number of variables chosen until the late 90s, then there was an abrupt increase until the current recession. Interestingly, in the middle of the recent recession, the number of variables selected started dropping again before stabilizing at the end of the recession.

⁹We subtract one since the S&P500 variable is always included in all models.

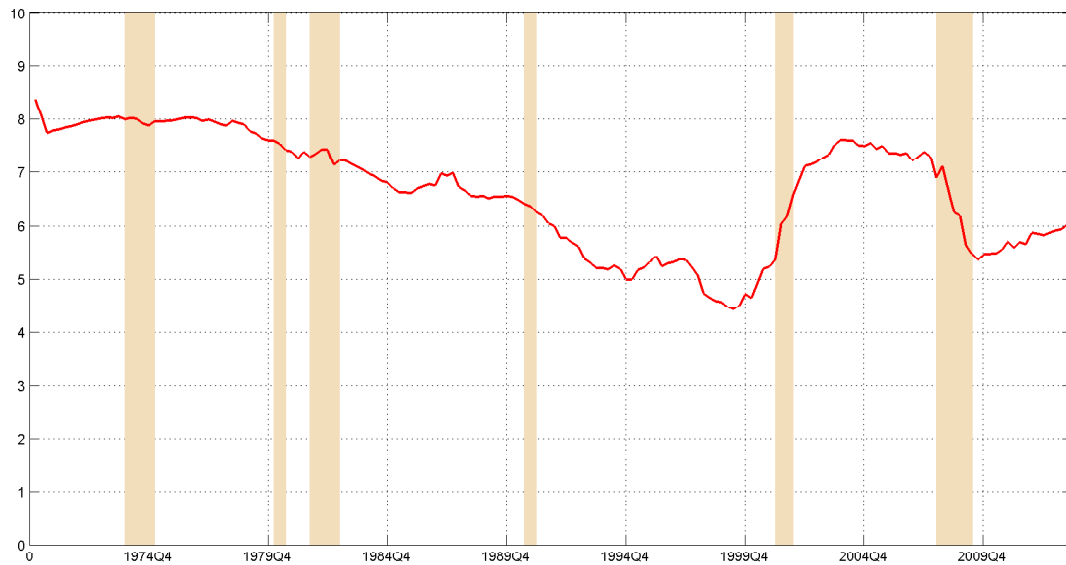


Figure 4. Average number of variables used to extract the FCI at each point in time as implied by DMA applied in the full TVP-FAVAR specification.

Figure 5 provides evidence on which variables receive most weight in the DMA procedure (or are selected by DMS). The numbers in each panel of this figure are the total probability DMA attaches to models which contain the variable named in the title on the panel. It is worth noting that there is substantial variable switching. That is, there are a few variables which enter then leave (or vice versa) the FCI.

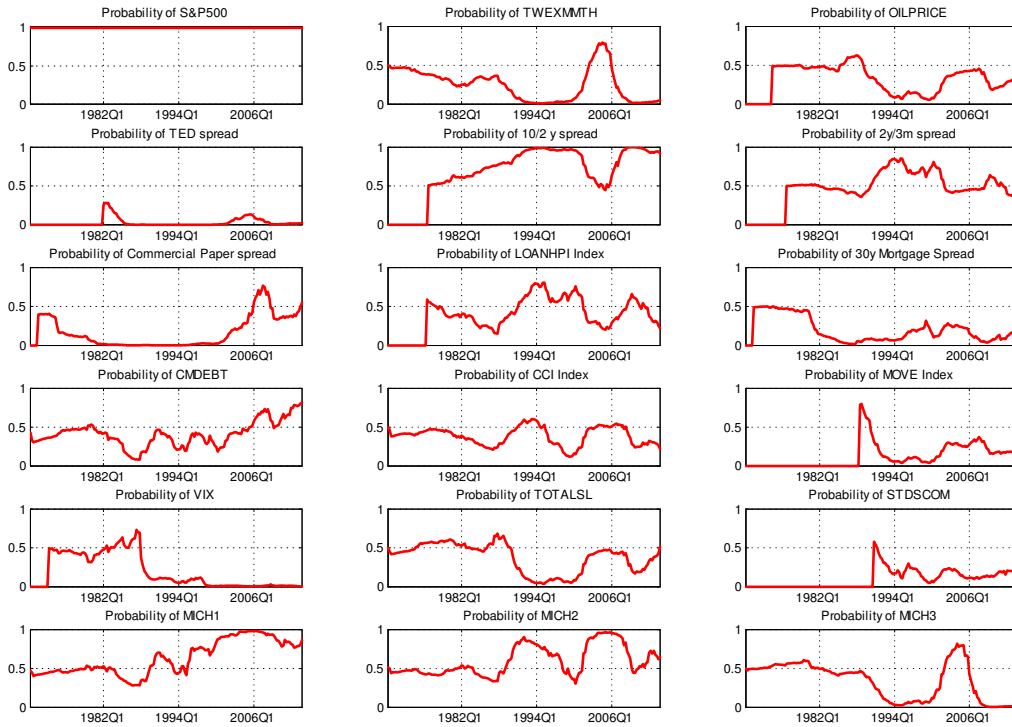


Figure 5. Time-varying probabilities of inclusion to the final FCI for each of the 18 financial variables (S&P500 is always included; see Section 3.1). Zero probabilities at the beginning of the sample for some of the variables correspond to periods of missing observations.

3.4 Forecasting

In this section, we investigate the performance of a wide range of models and methods for forecasting inflation, output growth and the unemployment rate. Our forecast evaluation period is 1990Q1 through 2013Q3- h for $h = 0, 1, 2, 3, 4$ quarters ahead. Note that, since we are using real-time data, it is also informative to present nowcasts and these are labelled $h = 0$ in the tables.¹⁰ Evaluation of forecast accuracy is based on the mean squared forecast error (MSFE) and average of predictive likelihoods (APL). The former evaluates the quality of point forecasts whereas the latter evaluates the quality of the entire predictive distribution. Results are normalized by dividing an MSFE or APL by the corresponding value produced by a benchmark model. We use a homoskedastic VAR (with no FCI) as our

¹⁰To be precise, at time t (given delays in release of macroeconomic variables) we have macroeconomic data available through period $t - 1$. This is used to provide nowcasts ($h = 0$) and forecasts ($h = 1, \dots, 4$) using iterative methods.

benchmark.¹¹ For this benchmark model, the table presents the actual value of MSFE or APL. Details of our choices of forgetting and decay factors are given in Section 2 and we investigate the sensitivity of results to their choice in Appendix C. The notation in the table (and the tables in Appendix C) extends our previous notation. Thus, TVP-FAVAR indicates the TVP-FAVAR with all 18 financial variables whereas TVP-FAVAR-DMA uses DMA (with $\alpha = 0.99$) over the 2^{17} TVP-FAVARs. Notation such as TVP-FAVAR (c_1, c_2) means that we set $\kappa_1 = \kappa_2 = c_1$ and $\kappa_3 = \kappa_4 = c_2$. We also present results for a VAR containing the macroeconomic variables plus a factor estimated using principle components (labelled VAR+PC). As another simple comparator, we also produce OLS (diffuse prior Bayesian) forecasts from AR models augmented with a factor (estimated by principal components). At each point in time, we choose the lag length which has produced the lowest MSFE over the last 40 quarters. From this model, we present recursive (labelled AR+PC(rec)) and rolling (labelled AR+PC(rol)) forecasts. For the rolling forecasts, we try windows of 25, 30, 35 and 40 periods and present results for the one which produces the lowest MSFE over the last 40 quarters.

For the MSFE's we also present the Bayesian variant of a test of forecast accuracy developed in Diebold and Mariano (1995). This test is described in Appendix B of Garratt et al (2009). If an MSFE in the table has a *, it means the approach forecasts significantly differently from the benchmark VAR.

Table 2 is organized so that each panel begins with a standard benchmark (e.g. the homoskedastic VAR), then adds heteroskedasticity (e.g. the TVP-VAR(0.96, 1) which selects the forgetting factor so as to make the VAR coefficients constant over time but allows for heteroskedasticity), then adds time variation in coefficients (e.g. the TVP-VAR(0.96, 0.99)). Then the next panel in the table repeats the process with models which include FCIs beginning with the FAVAR. The subsequent panel investigates the usefulness of DMA or DMS.

In general, Table 2 shows a pattern where forecasts improve as we add in extensions. Adding heteroskedasticity tends to improve forecasts substantially, then adding time-variation in parameters tends to improve them a bit more. Moving from TVP-VARs to TVP-FAVARs leads to further improvements in forecast performance. Using DMA or DMS with the 2^{17} TVP-FAVARs leads to yet further improvements. Of course, there are some exceptions to this pattern. Table 2 investigates forecasting performance for three variables at five different horizons using two different forecast evaluation metrics. With so many possible forecast evaluations it is not surprising there are some cases where simpler approaches beat our more complicated TVP-FAVAR-DMS or TVP-FAVAR-DMA approaches. But overall the latter are tending to produce the best forecasts.

Another important pattern in Table 2 is that our approach (again with some

¹¹This VAR is obtained as the restricted special case of the TVP-FAVAR (i.e. dropping the equation for the factors and turning off the variation in coefficients). Thus, its prior is equivalent to the prior for the initial conditions of the TVP-VAR described in Section 3.2.

exceptions) tends to forecast much better than standard approaches, including the VAR, the FAVAR and simple univariate methods. The simple univariate methods occasionally forecast well (e.g. when forecasting unemployment a year ahead the recursive AR model forecasts well when MSFE is used as a forecasting metric and AR+PC(rol) often forecasts inflation well), but often forecast very badly. In contrast, the approaches involving TVP-FAVARs and DMA or DMS typically forecast best and never forecast badly. They also virtually always beat FAVARs, indicating the importance of allowing for time-variation in parameters.

Table 2: Performance of our FCI compared to other forecasting models, 1990Q1 - 2013Q3

Forecast Metric	APL					MSFE				
	INFLATION									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
VAR (no FCI)	0.8587	0.7470	0.6959	0.6319	0.5848	0.0440	0.0540	0.0570	0.0610	0.0730
TVP-VAR (0.96,1)	1.36	1.31	1.32	1.32	1.29	1.05*	0.92*	0.92*	0.95*	0.97*
TVP-VAR (0.96,0.99)	1.35	1.32	1.33	1.32	1.28	0.93*	0.84*	0.87*	0.93*	0.96*
FAVAR	1.00	1.00	0.99	1.02	1.01	0.96	0.90	0.86*	0.89*	0.86
TVP-FAVAR (0.96,1)	1.36	1.32	1.33	1.32	1.32	1.08*	0.94*	0.92*	0.92*	0.93*
TVP-FAVAR (0.96,0.99)	1.34	1.34	1.35	1.34	1.31	0.91*	0.78*	0.78*	0.87*	0.92*
TVP-FAVAR-DMA (0.96,0.99)	1.47	1.46	1.48	1.50	1.47	0.94*	0.78*	0.82*	0.90*	0.91*
TVP-FAVAR-DMS (0.96,0.99)	1.50	1.48	1.50	1.53	1.49	1.02*	0.86*	0.94*	1.07*	1.08*
TVP-FAVAR-BMA (0.96,0.99)	1.35	1.36	1.38	1.38	1.32	0.94*	0.79*	0.83*	0.91*	0.92*
TVP-FAVAR-BMS (0.96,0.99)	1.42	1.42	1.44	1.45	1.42	1.01*	0.84*	0.93*	1.06*	1.06*
VAR+PC	0.99	1.00	1.00	1.02	1.03	0.96	0.91	0.92	0.94*	0.90*
AR + PC (rec)	1.02	1.09	1.06	1.05	1.04	1.41*	1.39*	1.61*	2.09*	2.28*
AR + PC (rol)	1.36	1.58	1.62	1.58	1.59	1.17*	0.90*	0.93*	1.12*	1.21*
	UNEMPLOYMENT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
VAR (no FCI)	0.6508	0.4119	0.3067	0.2468	0.2073	0.1210	0.3890	0.8230	1.4020	2.1100
TVP-VAR (0.96,1)	1.47	1.52	1.50	1.48	1.46	0.77*	0.72*	0.73*	0.72*	0.71*
TVP-VAR (0.96,0.99)	1.48	1.49	1.44	1.41	1.37	0.68*	0.62*	0.66*	0.66*	0.70*
FAVAR	0.91	0.87	0.83	0.82	0.81	0.99	1.14	1.25	1.29	1.34
TVP-FAVAR (0.96,1)	1.50	1.53	1.52	1.53	1.52	0.68*	0.62*	0.63*	0.65*	0.66*
TVP-FAVAR (0.96,0.99)	1.46	1.50	1.46	1.43	1.42	0.69*	0.65*	0.67*	0.67*	0.69*
TVP-FAVAR-DMA (0.96,0.99)	1.61	1.65	1.60	1.57	1.51	0.63*	0.55*	0.55*	0.56*	0.58*
TVP-FAVAR-DMS (0.96,0.99)	1.57	1.63	1.59	1.59	1.56	0.68*	0.56*	0.53*	0.54*	0.56*
TVP-FAVAR-BMA (0.96,0.99)	1.47	1.52	1.42	1.37	1.27	0.63*	0.55*	0.55*	0.56*	0.58*
TVP-FAVAR-BMS (0.96,0.99)	1.56	1.58	1.52	1.48	1.42	0.67*	0.57*	0.55*	0.56*	0.58*
VAR+PC	0.94	0.92	0.91	0.90	0.88	0.93	0.99	1.08	1.13	1.17
AR + PC (rec)	0.97	1.14	1.22	1.28	1.38	1.27	0.79*	0.64*	0.54*	0.47*
AR + PC (rol)	1.12	1.30	1.34	1.34	1.41	1.11	0.76*	0.74*	0.70*	0.55*
	OUTPUT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
VAR (no FCI)	0.3380	0.3170	0.3118	0.3104	0.3099	0.4310	0.5280	0.5470	0.5250	0.5050
TVP-VAR (0.96,1)	1.38	1.41	1.38	1.38	1.38	0.90*	0.83*	0.83*	0.83*	0.85*
TVP-VAR (0.96,0.99)	1.35	1.38	1.36	1.36	1.35	0.89*	0.81*	0.80*	0.80*	0.82*
FAVAR	0.97	0.97	0.97	0.97	0.98	0.92	1.00	0.95	0.87	0.83
TVP-FAVAR (0.96,1)	1.49	1.51	1.47	1.48	1.48	0.76*	0.76*	0.77*	0.79*	0.82*
TVP-FAVAR (0.96,0.99)	1.46	1.46	1.44	1.44	1.43	0.77*	0.72*	0.72*	0.71*	0.72*
TVP-FAVAR-DMA (0.96,0.99)	1.51	1.52	1.49	1.48	1.48	0.74*	0.68*	0.67*	0.69*	0.71*
TVP-FAVAR-DMS (0.96,0.99)	1.54	1.56	1.54	1.55	1.56	0.76*	0.68*	0.68*	0.71*	0.72*
TVP-FAVAR-BMA (0.96,0.99)	1.39	1.38	1.34	1.31	1.30	0.74*	0.68*	0.67*	0.69*	0.71*
TVP-FAVAR-BMS (0.96,0.99)	1.47	1.46	1.43	1.42	1.42	0.78*	0.71*	0.70*	0.72*	0.72*
VAR+PC	1.03	1.01	1.00	0.99	0.99	0.80*	0.93*	0.96*	0.94*	0.94*
AR + PC (rec)	1.10	1.14	1.16	1.17	1.17	0.97*	0.78*	0.83*	0.79*	0.85*
AR + PC (rol)	1.29	1.27	1.26	1.29	1.35	1.04*	0.90*	0.98*	1.00*	0.93*

Notes: APL is the average predictive likelihood (not in logarithms), and MSFE is the mean squared forecast error. Model's forecast performance is better when APL (MSFE) is higher (lower). For each variable (inflation, unemployment, output) the first line shows the APL and MSFE of the benchmark model for each forecast horizon h . All other models' APL and MSFE are relative to that of the benchmark model. Values of APL (MSFE) higher (lower) than 1 signify better performance than the benchmark.

Also of interest is the performance of our approach relative to VAR forecasts augmented with an existing FCI. Before doing so, we note that such comparisons are extremely difficult since different indices are based on different assumptions, data transformations, frequencies and sample sizes. The earliest common starting date for the FCIs is 1994Q1 and, accordingly, we re-estimate our models using data from this point and use 2000Q1 - 2013Q3- h as our forecast evaluation period. Table 1 describes the existing FCIs and FSIs and defines the acronyms we use in Table 3. We use the same naming convention as in Table 2 so that, for instance, TVP-VAR+FCI4 is the TVP-VAR which includes the macroeconomic variables and the Chicago Fed National FCI. All modelling choices (e.g. priors and discount factors) are the same benchmark choices as those used previously (see Section 3.2).

In relation to the approach developed in this paper, the story is similar to that told by Table 2. TVP-FAVARs with DMA or DMS still are either the best or among the best forecasting models for all the macroeconomic variables at all forecasting horizons. The table also shows that substantial benefits can be achieved by allowing for time-variation in parameters and doing DMA or DMS. Among the existing FCIs, use of the Chicago Fed's National FCI tends to lead to the best forecasting performance. Given the fact that it is this FCI which is most similar to our own (see Figure 3), it is not surprising that the TVP-VAR which includes FCI4 is given forecasting results similar to our TVP-FAVAR which includes all of the financial variables. Nevertheless, it is worth noting that DMA or DMS do add additional improvements in forecast performance so that TVP-FAVAR-DMA and TVP-FAVAR-DMS almost always forecast better than the TVP-VAR-FCI4.

Including any of the FSIs (i.e. FCI1, FCI2 or FCI3) in a TVP-VAR usually leads to fairly good forecast performance, but rarely as good as FCI4. And there are some cases where these other existing indices forecast poorly. For instance, the Kansas City Fed's FSI forecasts of unemployment are appreciably worse than those of other indices.

Table 3a: Performance of our FCI compared to other FCIs: relative APLs, 2000Q1 - 2013Q3

	INFLATION				UNEMPLOYMENT				OUTPUT						
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
VAR (no FCI)	0.8713	0.7572	0.7175	0.6632	0.6184	0.6532	0.4066	0.3004	0.2425	0.2018	0.3416	0.3191	0.3153	0.3121	0.3125
VAR+PC	0.98	1.00	1.00	1.02	1.03	0.96	0.94	0.94	0.92	0.91	1.05	1.03	1.01	1.00	0.99
VAR+FCI1	0.91	0.93	0.90	0.90	0.89	0.98	0.95	0.94	0.90	0.90	0.97	0.95	0.94	0.96	0.96
VAR+FCI2	0.97	0.99	0.98	0.98	0.98	0.71	0.65	0.62	0.61	0.59	1.00	1.02	1.00	1.00	1.00
VAR+FCI3	0.96	0.97	0.99	0.97	0.97	0.87	0.85	0.85	0.82	0.81	1.00	1.01	1.02	1.03	1.04
VAR+FCI4	0.99	0.96	0.97	0.97	0.98	0.98	0.97	0.94	0.91	0.90	1.05	1.04	1.01	1.00	1.00
TVP-VAR (no FCI)	1.20	1.22	1.24	1.23	1.18	1.36	1.34	1.28	1.20	1.13	1.23	1.29	1.28	1.28	1.27
TVP-VAR+PC	1.22	1.26	1.29	1.27	1.23	1.39	1.42	1.34	1.25	1.18	1.34	1.37	1.35	1.32	1.34
TVP-VAR+FCI1	1.16	1.16	1.18	1.15	1.10	1.45	1.40	1.34	1.20	1.12	1.28	1.29	1.25	1.24	1.26
TVP-VAR+FCI2	1.22	1.27	1.28	1.23	1.19	1.13	1.05	0.95	0.87	0.81	1.38	1.40	1.34	1.33	1.32
TVP-VAR+FCI3	1.30	1.28	1.28	1.26	1.18	1.28	1.26	1.19	1.12	1.10	1.29	1.36	1.32	1.31	1.31
TVP-VAR+FCI4	1.29	1.32	1.35	1.35	1.30	1.46	1.42	1.31	1.22	1.14	1.39	1.42	1.35	1.35	1.34
FAVAR	0.99	1.00	1.00	1.01	1.01	0.92	0.89	0.88	0.87	0.86	0.96	0.97	0.97	0.96	0.98
TVP-FAVAR	1.27	1.32	1.34	1.33	1.29	1.46	1.47	1.38	1.26	1.20	1.41	1.43	1.39	1.39	1.40
TVP-FAVAR-DMA	1.35	1.38	1.45	1.37	1.27	1.56	1.66	1.39	1.38	1.31	1.50	1.48	1.54	1.36	1.39
TVP-FAVAR-DMS	1.36	1.39	1.44	1.38	1.30	1.49	1.48	1.36	1.35	1.50	1.59	1.44	1.47	1.56	1.46
TVP-FAVAR-BMA	1.37	1.34	1.45	1.37	1.22	1.62	1.74	1.41	1.49	1.19	1.51	1.42	1.45	1.41	1.47
TVP-FAVAR-BMS	1.37	1.38	1.43	1.36	1.31	1.45	1.47	1.37	1.35	1.29	1.57	1.43	1.45	1.41	1.41
AR+PC (rec)	1.02	1.10	1.09	1.07	1.08	0.98	1.17	1.22	1.29	1.40	1.09	1.15	1.15	1.18	1.18
AR+PC (rol)	1.35	1.68	1.78	1.65	1.80	1.13	1.31	1.33	1.37	1.40	1.28	1.41	1.43	1.42	1.45
AR+FCI1 (rec)	0.97	1.03	1.01	0.93	0.92	0.82	0.95	0.99	0.99	0.98	1.00	1.06	1.06	1.01	1.00
AR+FCI1 (rol)	1.26	1.40	1.40	1.26	1.39	1.11	1.26	1.26	1.25	1.20	1.31	1.31	1.28	1.22	1.25
AR+FCI2 (rec)	1.01	1.07	1.05	0.98	0.98	0.85	0.99	1.05	1.09	1.16	1.06	1.11	1.12	1.12	1.14
AR+FCI2 (rol)	1.29	1.47	1.41	1.44	1.52	1.15	1.33	1.41	1.54	1.62	1.38	1.33	1.33	1.37	1.43
AR+FCI3 (rec)	0.98	1.08	1.04	0.97	0.99	0.80	0.92	0.99	1.02	1.08	1.05	1.11	1.15	1.16	1.15
AR+FCI3 (rol)	1.30	1.54	1.73	1.62	1.61	1.00	1.20	1.26	1.32	1.37	1.33	1.32	1.40	1.40	1.42
AR+FCI4 (rec)	1.01	1.10	1.07	1.03	1.03	1.00	1.24	1.30	1.35	1.44	1.11	1.15	1.17	1.19	1.19
AR+FCI4 (rol)	1.33	1.49	1.44	1.41	1.51	1.18	1.34	1.42	1.50	1.55	1.34	1.37	1.36	1.35	1.38

Table 3b: Performance of our FCI compared to other FCIs: relative MSFEs, 2000Q1 - 2013Q3

	INFLATION				UNEMPLOYMENT				OUTPUT						
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
	VAR (no FCI)	0.0602	0.0744	0.0768	0.0767	0.0846	0.1529	0.5175	1.1206	1.9325	2.9269	0.5719	0.7079	0.7295	0.7203
VAR+PC	1.04	0.99	0.98	1.00	0.96	0.83	0.91	0.98	1.02	1.06	0.76	0.86	0.92	0.94	0.98
VAR+FCI1	1.19	1.19	1.18	1.17	1.18	0.68	0.70	0.76	0.81	0.88	0.92	0.97	1.00	0.99	1.02
VAR+FCI2	1.01	0.92*	0.95	1.03	0.99	2.13	2.40	2.44	2.41	2.41	0.83	0.84*	0.84*	0.83*	0.84*
VAR+FCI3	1.09	1.06	1.11	1.19	1.21	1.27	1.41	1.42	1.46	1.47	0.91	0.92	0.86*	0.82*	0.79*
VAR+FCI4	1.16	1.09	0.97	1.02	1.04	0.78	0.82	0.87	0.93	0.96	0.79*	0.80*	0.86*	0.88*	0.89*
TVP-VAR (no FCI)	0.87*	0.76*	0.80*	0.88*	0.99	0.64*	0.61*	0.67*	0.71*	0.75*	0.83*	0.73*	0.72*	0.72*	0.72*
TVP-VAR+PC	0.89*	0.76*	0.75*	0.89*	1.03	0.66*	0.65*	0.68*	0.70*	0.73*	0.71*	0.65*	0.66*	0.68*	0.69*
TVP-VAR+FCI1	1.22	1.07	0.95*	1.11	1.28	0.56*	0.52*	0.55*	0.61*	0.69*	0.77*	0.73*	0.74*	0.72*	0.72*
TVP-VAR+FCI2	1.09	0.85*	0.77*	0.95*	1.01*	0.75*	0.87*	1.04*	1.22	1.38	0.68*	0.62*	0.62*	0.62*	0.61*
TVP-VAR+FCI3	0.97*	0.96*	1.03	1.29	1.44	0.71*	0.72*	0.76*	0.82*	0.85*	0.86*	0.74*	0.73*	0.70*	0.71*
TVP-VAR+FCI4	1.05	0.91*	0.83*	0.84*	0.92*	0.60*	0.57*	0.63*	0.68*	0.74*	0.71*	0.68*	0.74*	0.78*	0.79*
FAVAR	1.05	0.98*	0.95*	0.98	0.90*	1.07	1.18	1.24	1.27	1.30	1.07	1.13	1.14	1.10	1.07
TVP-FAVAR	0.91*	0.76*	0.77*	0.88*	1.01*	0.67*	0.67*	0.71*	0.73*	0.76*	0.72*	0.65*	0.66*	0.65*	0.66*
TVP-FAVAR-DMA	0.88*	0.79*	0.84*	0.94*	1.09	0.55*	0.46*	0.45*	0.47*	0.48*	0.69*	0.58*	0.58*	0.61*	0.65*
TVP-FAVAR-DMS	0.90*	0.81*	0.90*	1.03*	1.24	0.60*	0.48*	0.44*	0.45*	0.47*	0.70*	0.59*	0.59*	0.61*	0.64*
TVP-FAVAR-BMA	0.89*	0.80*	0.86*	0.97*	1.13	0.51*	0.42*	0.41*	0.44*	0.46*	0.68*	0.57*	0.58*	0.61*	0.65*
TVP-FAVAR-BMS	0.89*	0.83*	0.95*	1.11	1.30	0.54*	0.44*	0.43*	0.45*	0.47*	0.69*	0.61*	0.61*	0.64*	0.67*
AR+PC (rec)	1.14	1.07	1.09	1.47	1.67	1.01	0.61*	0.52*	0.46*	0.41*	0.90*	0.71*	0.76*	0.73*	0.79*
AR+PC (rol)	1.11	0.79*	0.69*	0.84*	1.07	1.22	0.89*	0.92*	0.88*	0.70*	1.10	0.93	1.03	1.06	0.99*
AR+FCI1 (rec)	1.68	1.16	1.43	2.18	2.90	1.29	0.99*	0.89*	0.87*	0.91*	1.20	0.92*	1.14	1.67	1.57
AR+FCI1 (rol)	1.10	0.92*	1.00*	1.48	1.41	1.24	1.15	1.13	1.23	1.46	1.31	0.95*	1.27	1.97	1.49
AR+FCI2 (rec)	1.35	1.03*	1.24	1.88	1.95	1.17	0.81*	0.72*	0.63*	0.53*	1.02	0.87*	0.92*	0.94*	0.94*
AR+FCI2 (rol)	1.11	0.89*	1.02*	1.17	1.12	1.12	1.04	1.01	0.96*	1.14	1.26	1.04	1.28	1.28	1.20
AR+FCI3 (rec)	1.33	1.19	1.52	1.90	1.89	1.80	1.06	0.86*	0.77*	0.68*	0.90*	0.81*	0.79*	0.74*	0.83*
AR+FCI3 (rol)	1.07	0.96	1.03*	1.23	1.19	2.34	1.63	1.42	1.37	1.38	1.04	1.01	0.95	0.88*	0.89*
AR+FCI4 (rec)	1.29	1.16	1.18	1.64	1.80	1.49	0.90*	0.75*	0.71*	0.63*	0.92*	0.76*	0.73*	0.77*	0.82*
AR+FCI4 (rol)	1.28	0.89*	1.21	1.30	1.41	1.25	1.39	1.54	1.16	0.92*	1.32	1.38	1.64	1.89	1.31

4 Conclusions

In this paper, we have argued for the desirability of constructing a dynamic financial conditions index which takes into account changes in the financial sector, its interaction with the macroeconomy and data availability. In particular, we want a methodology which can choose different financial variables at different points in time and weight them differently. We develop DMS and DMA methods, adapted from Raftery et al (2010) and others, to achieve this aim. Using a large data set of US macroeconomic and financial variables, we find our estimated FCI to have reasonable properties. It is broadly similar to existing FCIs, but does exhibit some interesting differences.

We also demonstrate the usefulness of our FCI as a forecasting tool. Working with a large model space involving many TVP-FAVARs which make different choices of financial variables, we find DMA and DMS methods lead to improved forecasts of macroeconomic variables, relative to methods which use a single model. We also show the importance of allowing for time variation in parameters and heteroskedasticity in achieving good forecast performance.

References

- [1] Aiolfi, M., Capistrán, C., Timmermann, A., 2010. Forecast combinations. In: Michael P. Clements and David F. Hendry (Eds). The Oxford Handbook of Economic Forecasting. Oxford University Press, USA.
- [2] Balakrishnan, R., Danninger, S., Elekdag, S., Tytell, I., 2009. The transmission of financial stress from advanced to emerging economies. IMF Working Papers 09/133, International Monetary Fund.
- [3] Banerjee, A., Marcellino, M., Masten, I., 2008. Forecasting macroeconomic variables using diffusion indexes in short samples with structural change. CEPR Discussion Papers 6706, C.E.P.R. Discussion Papers.
- [4] Bates, B. J., Plagborg-Møller, M., Stock, J. H., Watson, M. W., 2013. Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177, 289-304.
- [5] Beaton, K., Lalonde, R., Luu, C., 2009. A financial conditions index for the United States. Bank of Canada Discussion Paper, November.
- [6] Bernanke, B., Boivin, J., Elias, P., 2005. Measuring monetary policy: A factor augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics* 120, 387-422.
- [7] Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132, 169-194.
- [8] Brave, S., Butters, R., 2011. Monitoring financial stability: a financial conditions index approach. *Economic Perspectives*, Issue Q1, Federal Reserve Bank of Chicago, 22-43.
- [9] Breitung, J., Eickmeier, S., 2011. Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, 163(1), 71-84.
- [10] Clark, T., 2009. Real-time density forecasts from VARs with stochastic volatility. Federal Reserve Bank of Kansas City Research Working Paper 09-08.
- [11] Cogley, T., Sargent, T. J., 2005. Drift and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics* 8(2), 262-302.
- [12] D'Agostino, A., Gambetti, L., Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28, 81-101.
- [13] Del Negro, M., Otrok, C., 2008. Dynamic factor models with time-varying parameters: Measuring changes in international business cycles. University of Missouri Manuscript.

- [14] Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. *Journal of Business Economics and Statistics*, 13, 134-144.
- [15] Doz, C., Giannone, D., Reichlin, L., 2011. A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics* 164, 188-205.
- [16] Eickmeier, S., Lemke, W., Marcellino, M., 2011a. Classical time-varying FAVAR models - estimation, forecasting and structural analysis. Deutsche Bundesbank, Discussion Paper Series 1: Economic Studies, No 04/2011.
- [17] Eickmeier, S., Lemke, W., Marcellino, M., 2011b. The changing international transmission of financial shocks: evidence from a classical time-varying FAVAR. Deutsche Bundesbank, Discussion Paper Series 1: Economic Studies, No 05/2011.
- [18] English, W., Tsatsaronis, K., Zoli, E., 2005. Assessing the predictive power of measures of financial conditions for macroeconomic variables. Bank for International Settlements Papers No. 22, 228-252.
- [19] Filippeli, T. and Theodoridis, K., 2013. DSGE priors for BVAR models, *Empirical Economics*, forthcoming.
- [20] Forni M., Hallin M., Lippi M., Reichlin L., 2003. Do financial variables help forecasting inflation and real activity in the EURO area? *Journal of Monetary Economics* 50, 1243-55.
- [21] Garratt, A., Koop, G., Mise, E. and Vahey, S., 2009. Real-time prediction with UK monetary aggregates in the presence of model uncertainty. *Journal of Business and Economic Statistics*, 27, 480-491.
- [22] Giannone, D., Reichlin, L., Sala, L., 2005. Monetary Policy in Real Time. NBER Chapters, in: *NBER Macroeconomics Annual 2004*, Volume 19, pages 161-224, National Bureau of Economic Research, Inc.
- [23] Gomez, E., Murcia, A., Zamudio, N., 2011. Financial conditions index: Early and leading indicator for Colombia? *Financial Stability Report*, Central Bank of Colombia.
- [24] Hatzius, J., Hooper, P., Mishkin, F., Schoenholtz, K., Watson, M., 2010. Financial conditions indexes: A fresh look after the financial crisis. NBER Working Papers 16150, National Bureau of Economic Research, Inc.
- [25] Hoeting, J. A., Madigan, D., Raftery, A. E., Volinsky, C. T., 1999. Bayesian Model Averaging: A Tutorial. *Statistical Science* 14, 382-417.

- [26] Ingram, B., Whiteman, C., 1994. Supplanting the 'Minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors. *Journal of Monetary Economics*, 34, 497-510.,
- [27] Kárný, M., 2006. *Optimized Bayesian Dynamic Advising: Theory and Algorithms*, Springer-Verlag New York Inc.
- [28] Kaufmann, S., Schumacher, C., 2012. Finding relevant variables in sparse Bayesian factor models: Economic applications and simulation results. *Deutsche Bundesbank Discussion Paper No 29/2012*.
- [29] Koop, G., Korobilis, D., 2012. Forecasting inflation using dynamic model averaging. *International Economic Review* 53, 867-886.
- [30] Koop, G., Korobilis, D., 2013. Large time-varying parameter VARs. *Journal of Econometrics*, 177, 185-198.
- [31] Korobilis, D., 2013. Assessing the transmission of monetary policy shocks using time-varying parameter dynamic factor models. *Oxford Bulletin of Economics and Statistics* 75, 157-179.
- [32] Kulhavý, R., Kraus, F., 1996. On duality of regularized exponential and linear forgetting. *Automatica* 32,1403-1416.
- [33] Lo Duca, M., Peltonen, T., 2013. Assessing systemic risks and predicting systemic events. *Journal of Banking and Finance* 37, 2183-2195.
- [34] Lütkepohl, H., 2005. *New Introduction to Multiple Time Series Analysis*. Springer: New York.
- [35] Matheson, T., 2011. Financial conditions indexes for the United States and Euro Area. *IMF Working Papers 11/93*, International Monetary Fund.
- [36] Nelson, L., Stear, E., 1976. The simultaneous on-line estimation of parameters and states in linear systems. *IEEE Transactions on Automatic Control* 21, 94 - 98.
- [37] Primiceri, G., 2005. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies* 72, 821-852.
- [38] Quintana, J.M., West, M., 1988. Time Series Analysis of Compositional Data. In *Bayesian Statistics 3*, (eds: J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith), Oxford University Press.
- [39] Raftery, A. E., Karny, M., Ettler, P., 2010. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics* 52, 52-66.

- [40] Schorfheide, F., Wolpin, K. E., 2012. On the use of holdout samples for model selection. *American Economic Review - Papers and Proceedings* 102(3), 477-481.
- [41] Stock, J. H., Watson, M. W., 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* 20,147-162.
- [42] Stock, J. H., Watson, M. W., 2009. Forecasting in dynamic factor models subject to structural instability. In: Jennifer Castle and Neil Shephard (Eds), *The Methodology and Practice of Econometrics: A Festschrift in Honour of Professor David Hendry*, Oxford: Oxford University Press.

A. Data Appendix

The following table describes the series we used to extract our Financial Conditions Index. The series are selected as described in Section 3.1 of the paper. The fourth column describes the stationarity transformation codes (Tcodes) which have been applied to each variable. Tcode shows the stationarity transformation for each variable: Tcode=1, variable remains untransformed (levels) and Tcode=5, take first log differences. The fifth column describes the source of each variable. The codes are: B - Bloomberg; D - Datastream; F - Federal Reserve Economic Data (<http://research.stlouisfed.org/fred2/>); G - Amit Goyal (<http://www.hec.unil.ch/agoyal/>); R - Board of Governors of the Federal Reserve System (<http://www.federalreserve.gov/>); U - University of Michigan (<http://www.sca.isr.umich.edu/>); W - Mark W. Watson (<http://www.princeton.edu/~mwatson/>).

Table A1: The 18 variables used for the FCI

No	Mnemonic	Description	Tcode	Source	Sample
1	S&P500	S&P 500 Stock Price Index	5	F	1970:Q1-2013:Q3
2	TWEXMMTH	FRB Nominal Major Currencies Dollar Index (Linked To EXRUS In 1973:1)	5	W/F	1970:Q1-2013:Q3
3	OILPRICE	U.S. Crude Oil Imported Acquisition Cost by Refiners (Dollars per Barrel)	5	B	1974:Q1-2013:Q3
4	TED spread	3m LIBOR - 3m Treasury Bill Rate	1	B	1981:Q4-2013:Q3
5	10/2 y spread	10-Year/2-Year Treasury Yield Spread	1	F	1976:Q3-2013:Q3
6	2y/3m spread	2-Year/3-Month Treasury Yield Spread	1	F	1976:Q3-2013:Q3
7	Commercial Paper spread	3-Month Financial Commercial Paper/Treasury Bill Spread	1	B	1971:Q2-2013:Q3
8	LOANHPI Index	Home Loan Performance Index U.S. Index Level	5	B	1976:Q1-2013:Q3
9	30y Mortgage Spread	30y Conventional Mortgage Rate - 10y Treasury Rate	1	F	1971:Q2-2013:Q3
10	CMDEBT	Households and Nonprofit Organizations; Credit Market Instruments; Liability, Level	5	F	1971:Q1-2013:Q3
11	GCI Index	Continuous Commodity Futures Price Index	5	B	1970:Q1-2013:Q3
12	MOVE Index	Merrill Lynch One-Month Treasury Options Volatility Index (MOVE)	1	B	1988:Q2-2013:Q3
13	VIX	CBOE S&P100 Volatility Index	1	B	1986:Q3-2013:Q3
14	TOTALSL	Total Consumer Credit Owned And Securitized, Outstanding	5	F	1970:Q1-2013:Q3
15	STDSCOM	SLOOS: Net percentage of banks tightening standards for commercial real estate loans	1	R	1990:Q3-2013:Q3
16	MICH1	Michigan Survey: Interest Rates/Credit Reason for Good/Bad Conditions for Buying Houses Spread	1	U	1970:Q1-2013:Q3
17	MICH2	Michigan Survey: Buying Conditions for Vehicles	1	U	1970:Q1-2013:Q3
18	MICH3	Michigan Survey: Expected Change In Financial Situation	1	U	1970:Q1-2013:Q3

B. Technical Appendix

In this appendix, we describe the econometric methods we use to estimate a TVP-FAVAR and restricted versions of it.

We write the TVP-FAVAR compactly as

$$x_t = z_t \Lambda_t + u_t, \quad u_t \sim N(0, V_t) \quad (\text{B.1})$$

$$z_t = z_{t-1} \beta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, Q_t) \quad (\text{B.2})$$

$$\lambda_t = \lambda_{t-1} + v_t, \quad v_t \sim N(0, W_t) \quad (\text{B.3})$$

$$\beta_t = \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, R_t) \quad (\text{B.4})$$

where $\lambda_t = \left((\lambda_t^y)', (\lambda_t^f)' \right)'$ and $z_t = \begin{bmatrix} y_t \\ f_t \end{bmatrix}$. We also use notation where \tilde{f}_t is the standard principal components estimate of f_t based on x_t (using data up to time t) and $\tilde{z}_t = \begin{bmatrix} y_t \\ \tilde{f}_t \end{bmatrix}$. Additionally, if a_t is a vector then $a_{i,t}$ is the i^{th} element of that vector; and if A_t is a matrix $A_{ii,t}$ is its $(i, i)^{\text{th}}$ element. Estimates of time varying parameters or latent states can be made using data available at time $t-1$ (filtering), or time t (updating) or time T (smoothing). We use subscript notation for this such that $a_{t|\tau}$ is an estimate (or posterior moment) of time-varying parameter a_t using data available through period τ .

Our estimation algorithm requires initialization of all state variables. In particular we define the following initial conditions for all system unknown parameters

$$f_0 \sim N\left(0, \Sigma_{0|0}^f\right), \quad (\text{B.5})$$

$$\lambda_0 \sim N\left(0, \Sigma_{0|0}^\lambda\right), \quad (\text{B.6})$$

$$\beta_0 \sim N\left(0, \Sigma_{0|0}^\beta\right), \quad (\text{B.7})$$

$$V_0 \equiv 1 \times I_n, \quad (\text{B.8})$$

$$Q_0 \equiv 1 \times I_{s+1}. \quad (\text{B.9})$$

The algorithm follows the following steps:

1. Given the initial conditions and $z_t = \tilde{z}_t$ obtain **filtered** estimates of $\lambda_t, \beta_t, V_t, Q_t$ using the following recursion for $t = 1, \dots, T$

(a) The Kalman filter tells us:

$$\lambda_t | \text{Data}_{1:t-1} \sim N\left(\lambda_{t|t-1}, \Sigma_{t|t-1}^\lambda\right),$$

$$\beta_t | \text{Data}_{1:t-1} \sim N\left(\beta_{t|t-1}, \Sigma_{t|t-1}^\beta\right),$$

where $\lambda_{t|t-1} = \lambda_{t-1|t-1}$, $\Sigma_{t|t-1}^\lambda = \Sigma_{t-1|t-1}^\lambda + \widehat{W}_t$, $\beta_{t|t-1} = \beta_{t-1|t-1}$ and $\Sigma_{t|t-1}^\beta = \Sigma_{t-1|t-1}^\beta + \widehat{R}_t$. The error covariances are estimated using forgetting factors as: $\widehat{W}_t = (1 - \kappa_3^{-1}) \Sigma_{t-1|t-1}^\lambda$ and $\widehat{R}_t = (1 - \kappa_4^{-1}) \Sigma_{t-1|t-1}^\beta$.

- (b) Calculate estimates of V_t and Q_t for use in the updating step using the following EWMA specifications:

$$\widehat{V}_{i,t} = \kappa_1 V_{i,t-1|t-1} + (1 - \kappa_1) \widehat{u}_{i,t} \widehat{u}_{i,t}' \quad (\text{B.10})$$

$$\widehat{Q}_t = \kappa_2 Q_{t-1|t-1} + (1 - \kappa_2) \widehat{\varepsilon}_t \widehat{\varepsilon}_t' \quad (\text{B.11})$$

where $\widehat{u}_{i,t} = x_{i,t} - \widetilde{z}_t \lambda_{i,t|t-1}$, for $i = 1, \dots, n$, and $\widehat{\varepsilon}_t = \widetilde{z}_t - \widetilde{z}_{t-1} \beta_{t|t-1}$.

- (c) Update λ_t and β_t given information at time t using the Kalman filter update step

- Update $\lambda_{i,t}$ for each $i = 1, \dots, n$ using

$$\lambda_{it} | \text{Data}_{1:t} \sim N(\lambda_{i,t|t}, \Sigma_{ii,t|t}^\lambda),$$

where $\lambda_{i,t|t} = \lambda_{i,t|t-1} + \Sigma_{ii,t|t-1}^\lambda \widetilde{z}_t' \left(\widehat{V}_{ii,t} + \widetilde{z}_t \Sigma_{ii,t|t-1}^\lambda \widetilde{z}_t' \right)^{-1} (x_t - \widetilde{z}_t \lambda_{t|t-1})$

and $\Sigma_{ii,t|t}^\lambda = \Sigma_{ii,t|t-1}^\lambda - \Sigma_{ii,t|t-1}^\lambda \widetilde{z}_t' \left(\widehat{V}_{ii,t} + \widetilde{z}_t \Sigma_{ii,t|t-1}^\lambda \widetilde{z}_t' \right)^{-1} \widetilde{z}_t \Sigma_{ii,t|t-1}^\lambda$.

- Update β_t from

$$\beta_t | \text{Data}_{1:t} \sim N(\beta_{t|t}, \Sigma_{t|t}^\beta),$$

where $\beta_{t|t} = \beta_{t|t-1} + \Sigma_{t|t-1}^\beta \widetilde{z}_{t-1}' \left(\widehat{Q}_t + \widetilde{z}_{t-1} \Sigma_{t|t-1}^\beta \widetilde{z}_{t-1}' \right)^{-1} (\widetilde{z}_t - \widetilde{z}_{t-1} \widehat{\beta}_{t|t-1})$

and $\Sigma_{t|t}^\beta = \Sigma_{t|t-1}^\beta - \Sigma_{t|t-1}^\beta \widetilde{z}_{t-1}' \left(\widehat{Q}_t + \widetilde{z}_{t-1} \Sigma_{t|t-1}^\beta \widetilde{z}_{t-1}' \right)^{-1} \widetilde{z}_{t-1} \Sigma_{t|t-1}^\beta$.

- (d) Update V_t and Q_t given information at time t using the EWMA specifications as follows:

$$V_{i,t|t} = \kappa_1 V_{i,t-1|t-1} + (1 - \kappa_1) \widehat{u}_{i,t|t} \widehat{u}_{i,t|t}' \quad (\text{B.12})$$

$$Q_{t|t} = \kappa_2 Q_{t-1|t-1} + (1 - \kappa_2) \widehat{\varepsilon}_{t|t} \widehat{\varepsilon}_{t|t}' \quad (\text{B.13})$$

where $\widehat{u}_{i,t|t} = x_{i,t} - \widetilde{z}_t \lambda_{i,t|t}$, for $i = 1, \dots, n$, and $\widehat{\varepsilon}_{t|t} = \widetilde{z}_t - \widetilde{z}_{t-1} \beta_{t|t}$.

2. Obtain **smoothed** estimates of $\lambda_t, \beta_t, V_t, Q_t$ using the following recursions for $t = T - 1, \dots, 1$

- (a) Update λ_t and β_t given information at time $t + 1$ using the fixed interval smoother

- Update $\lambda_{i,t}$ for each $i = 1, \dots, n$ from

$$\lambda_{it}|Data_{1:T} \sim N(\lambda_{i,t|T}, \Sigma_{ii,t|T}^\lambda),$$

where $\lambda_{i,t|T} = \lambda_{i,t|t} + C_t^\lambda (\lambda_{i,t+1|T} - \lambda_{i,t+1|t})$, $\Sigma_{ii,t|T}^\lambda = \Sigma_{ii,t|t}^\lambda + C_t^\lambda (\Sigma_{ii,t+1|T}^\lambda - \Sigma_{ii,t+1|t}^\lambda) C_t^{\lambda'}$
and $C_t^\lambda = \Sigma_{ii,t|t}^\lambda (\Sigma_{ii,t+1|t}^\lambda)^{-1}$.

- Update β_t from

$$\beta_t|Data_{1:T} \sim N(\beta_{t|T}, \Sigma_{t|T}^\beta),$$

where $\beta_{t|T} = \beta_{t|t} + C_t^\beta (\beta_{t+1|T} - \beta_{t+1|t})$, $\Sigma_{t|T}^\beta = \Sigma_{t|t}^\beta + C_t^\beta (\Sigma_{t+1|T}^\beta - \Sigma_{t+1|t}^\beta) C_t^{\beta'}$
and $C_t^\beta = \Sigma_{t|t}^\beta (\Sigma_{t+1|t}^\beta)^{-1}$.

- (b) Update V_t and Q_t given information at time $t + 1$ using the following equations

$$V_{t|t+1}^{-1} = \kappa_1 V_{t|t}^{-1} + (1 - \kappa_1) V_{t+1|t+1}^{-1}, \quad (\text{B.14})$$

$$Q_{t|t+1}^{-1} = \kappa_2 Q_{t|t}^{-1} + (1 - \kappa_2) Q_{t+1|t+1}^{-1}. \quad (\text{B.15})$$

3. Means and variances of f_t given appropriate estimates of $\lambda_t, \beta_t, V_t, Q_t$ described in the preceding steps can be obtained using the standard **Kalman filter and smoother**.

Treatment of missing values

In our application our sample is unbalanced, since it contains many financial variables which have been collected only after the 1970s or the 1980s. Similar issues are faced by organizations which monitor FCIs. For instance, the Chicago Fed National FCI comprises 100 series where most of them have different starting dates. Although specific computational methods for dealing with such issues exist (e.g. the EM algorithm or Gibbs sampler with data augmentation), our focus is on averaging over many models which means such methods are computationally infeasible. Accordingly, similar to our purpose of developing a simulation-free and fast algorithm for parameter estimation, we want to avoid simulation methods for estimating the missing data in x_t . Additionally, methods such as interpolation can work poorly when missing values are at the beginning of the sample.

Since the missing data in x_t are in the beginning, we make the assumption that the factor (FCI) is estimated using only the observed series. The estimation algorithm above allows for such an approach in a straightforward manner by just replacing missing values with zeros. The loadings λ (whether time-varying, or constant) will become equal to 0, thus removing from the estimate of f_t the effect of the variables in x_t which have missing values at time t . This feature holds both

for the initial principal components estimate \tilde{f}_t , as well as the final Kalman filter estimate.

Estimation of a single TVP-FAVAR

Given the algorithm above, we can estimate the TVP-FAVAR by choosing values of $\kappa_1, \kappa_2, \kappa_3, \kappa_4 < 1$. For the main results in the paper we set $\kappa_1 = \kappa_2 = 0.96$ and $\kappa_3 = \kappa_4 = 0.99$. The restricted special cases of the TVP-FAVAR listed in Section 2.1 can be obtained by setting forgetting and/or decay factors to particular values. If we set $\kappa_3 = 1$ and $\kappa_4 < 1$ then we can obtain a TVP-VAR augmented with factors estimated with constant loadings¹². Setting $\kappa_3 = 1$ and $\kappa_4 = 1$ leads to the constant parameter FAVAR with heteroskedastic covariances (assuming that $\kappa_1, \kappa_2 < 1$). If we additionally set $\kappa_1 = \kappa_2 = 1$ we can estimate homoskedastic versions of the various models, since in that case $V_t = V_{t-1} = \dots = V_1 = V_0$ and $Q_t = Q_{t-1} = \dots = Q_1 = Q_0$. Nevertheless, as discussed in the main body of the paper, this is a case which is always dominated (in terms of forecast performance) by the heteroskedastic case. We provide more evidence for this in Appendix C.

Estimation of multiple models (DMA/DMS)

In order to implement the DMA/DMS exercises we run the algorithm described above for each of the $2^{17} = 131,072$ models. Note that for a specific DMA exercise all models are nested, and the only thing that changes is the number of variables in the vector x_t that we use in order to extract the FCI. Given our discussion about how missing values are treated by the Kalman filter, in order to estimate a specific model which uses, say, the 1st, 3rd and 15th series in x_t , we simply multiply all but the 1st, 3rd and 15th columns of x_t with zeros. In that case, we remove at all times t the effects of all 15 variables we do not use for estimation of the specific model, and at the same time we still have as a dependent variable a 18×1 dimensional vector (and programming is greatly simplified).

The most important feature of DMA is that, unlike many Bayesian model selection and averaging procedures which use MCMC methods, there is no dependence in estimating each model and iterations using “for” loops are independent. That means that it is trivial to adapt our code to use features such as parallel computing, thus taking advantage of the widespread availability of modern multi-core processors (or large clusters of PCs). In MATLAB this is as easy as replacing the typical “for” loop which would run for models 1 to 131,072, with a “parfor” loop.

The reader is encouraged to look at our code which is available on <https://sites.google.com/site/dimitriskorobilis/matlab>, which also has the option to call the Parallel Processing Toolbox in a MATLAB environment.

¹²In a previous version of this paper we have named this model the factor augmented TVP-VAR or FA-TVP-VAR.

C. Sensitivity analysis

In this section we present further results for several different choices of prior hyperparameters, which can reflect different beliefs about time variation in the model parameters as well as beliefs about variation over time in the choice of the optimal model.

C.1. Comparison of relatively noninformative with training sample priors

In the main body of the paper results are presented for a subjectively-elicited but relatively noninformative prior. An interesting alternative is to choose all prior hyperparameters using a training sample of data. In the context of TVP-VARs, Primiceri (2005) suggests such a prior which is based on splitting the data into a training sample, and a testing sample where estimation occurs. OLS estimation of a constant coefficient model using the training sample provides parameter estimates which are used as prior hyperparameters for the testing sample. Such training sample priors are commonly-used in Bayesian analysis, and in the context of TVP models help provide regularized posterior estimators which can also help numerical stability. This latter feature is important in the case of TVP-VARs and TVP-FAVARs estimated with MCMC - see the discussion in Section 4.1 of Primiceri (2005).

In this section, we introduce such a training sample prior for our FAVAR and TVP-FAVAR models. We use the first 10 years of data (1970Q1-1979Q4) in our original sample as the training sample. We estimate a FAVAR with constant parameters using OLS methods (replacing the factor with its principal component estimate) on this training sample. These OLS estimates are used to in the initial conditions for the estimation sample, 1980Q1-2013Q3, as follows

$$\begin{aligned}
 f_0 &\sim N\left(\widehat{f}_T^{TS}, 0.1\right), \\
 \lambda_0 &\sim N\left(0, 4 \times \text{var}\left(\widehat{\lambda}^{TS}\right)\right), \\
 \beta_0 &\sim N\left(0, 4 \times \text{var}\left(\widehat{\beta}^{TS}\right)\right), \\
 V_0 &\equiv 1 \times \widehat{V}^{TS}, \\
 Q_0 &\equiv 1 \times \widehat{Q}^{TS}, \\
 \pi_{0|0,j} &= \frac{1}{J},
 \end{aligned}$$

where parameters with a hat and a superscript TS denote OLS estimates of the respective parameters in the time-invariant FAVAR fitted using the training sample.

Other settings used in the main body of the paper remain the same, e.g. we use four lags everywhere and the decay and forgetting factors that define each of the

models (DMA vs BMA, or FAVAR vs TVP-FAVAR) are exactly the ones specified in Section 3.2. The forecast evaluation sample is 1990Q1-2013Q3. Table C.1 presents the results of this exercise. It is divided in three blocks, each one corresponding to one of the three variables of interest (inflation, unemployment, output growth). For each block the first line shows the APL and MSFE of the benchmark TVP-FAVAR used in the main body of the paper. The APLs and MSFEs of all other models are relative to the APL and MSFE of the benchmark TVP-FAVAR, which implies that relative APL higher than one (similarly, relative MSFE lower than one) is an indication of better performance relative to the benchmark.

Table C1: Comparison of benchmark with training sample (TS) priors, 1990Q1-2013Q3

Forecast Metric	APL					MSFE				
	INFLATION									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	1.1233	0.9741	0.9066	0.8194	0.7380	0.0416	0.0454	0.0466	0.0562	0.0696
FAVAR	0.76	0.77	0.76	0.78	0.80	1.06	1.14	1.10	1.02	0.94
TVP-FAVAR-TS	1.01	1.04	1.06	1.10	1.09	0.93	0.94	0.87	0.87	0.97
FAVAR-TS	0.87	0.89	0.91	0.96	1.00	0.93	1.08	1.10	0.98	0.94
TVP-FAVAR-DMA-TS	1.13	1.13	1.15	1.16	1.17	0.94	0.84	0.89	0.90	0.86
TVP-FAVAR-BMA-TS	1.13	1.14	1.16	1.17	1.18	0.87	0.82	0.87	0.86	0.85
	UNEMPLOYMENT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	0.9170	0.5868	0.4191	0.3229	0.2639	0.0906	0.2741	0.5950	1.0105	1.5359
FAVAR	0.64	0.61	0.61	0.63	0.64	1.68	2.01	2.09	2.14	2.17
TVP-FAVAR-TS	0.81	0.74	0.74	0.77	0.80	1.58	1.84	1.77	1.62	1.45
FAVAR-TS	0.50	0.51	0.56	0.62	0.66	2.44	2.85	2.82	2.47	2.21
TVP-FAVAR-DMA-TS	1.14	1.15	1.15	1.19	1.19	1.01	0.87	0.77	0.75	0.75
TVP-FAVAR-BMA-TS	1.15	1.16	1.16	1.19	1.22	1.00	0.86	0.77	0.75	0.75
	OUTPUT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	0.4643	0.4344	0.4206	0.4156	0.4131	0.3548	0.4085	0.4204	0.3972	0.3854
FAVAR	0.71	0.71	0.72	0.73	0.74	1.43	1.64	1.57	1.48	1.40
TVP-FAVAR-TS	0.80	0.82	0.81	0.79	0.78	1.64	1.64	1.68	1.09	1.03
FAVAR-TS	0.72	0.76	0.74	0.66	0.65	2.04	1.84	2.01	1.21	1.37
TVP-FAVAR-DMA-TS	1.05	1.01	1.04	1.02	0.93	1.45	1.40	1.26	1.50	1.18
TVP-FAVAR-BMA-TS	0.97	0.95	1.06	0.93	0.89	1.44	1.40	1.26	1.50	1.18

Notes: APL is the average predictive likelihood (not in logarithms), and MSFE is the mean squared forecast error. Model's forecast performance is better when APL (MSFE) is higher (lower). For each variable (inflation, unemployment, output) the first line shows the APL and MSFE of the benchmark model for each forecast horizon h . All other models' APL and MSFE are relative to that of the benchmark model. Values of APL (MSFE) higher (lower) than 1 signify better performance than the benchmark.

In Table C1 we see that the training sample prior is not particularly helpful for forecasting unemployment and output, while there is marginal improvement of APLs and MSFEs for inflation. This result suggests that globally (for all three variables of interest) our relatively noninformative prior is a sensible choice and avoids problems with training samples such as that noted by Schorfheide and Wolpin (2012):

“[...] from a Bayesian perspective the use of holdout samples is suboptimal because the computation of posterior probabilities should be based on the entire sample and not just on a subsample.” Schorfheide and Wolpin (2012)

From our own experience with Bayesian VARs (and FAVARs), we can argue that training sample priors are very important in cases where numerical stability is an issue. For example, the decision of Primiceri (2005) and others to use training sample priors when estimating TVP-VAR with MCMC works well because serious numerical issues can occur when relatively noninformative priors and diffuse initial conditions are used in the full sample. In the present paper, where we examine high dimensional TVP-FAVARs, we do not have such numerical issues due to the computational simplicity of our algorithm (which does not involve the use of Monte Carlo or other iterative methods). This exact advantage of our estimation methods justifies our decision to use relatively noninformative priors in the full available sample as the benchmark case. In other datasets, e.g. macroeconomic data for the Euro-Area, training samples might not be available at all. In this case one can either use a subjectively elicited prior or a prior which is informed by economic theory. For example, economic theory restrictions can enter our prior distributions for the (FA)VAR part of our model, in the spirit of Filippeli and Theodoridis (2013) and Ingram and Whiteman (1994). Examining such restrictions is beyond the scope of our main aim in this paper, which is to show the general role of time-variation in models and parameters when extracting an FCI.

C.2. A Different Way to Purge Macroeconomic Conditions from the FCI

We have stressed in the main body of the paper that a good FCI should be purged of macroeconomic conditions. However, it might be argued that these macroeconomic conditions should not reflect the current situation, but rather current expectations of future macroeconomic conditions. In the body of the paper, we always use the current values of the three macroeconomic variables of interest (y_t). It is possible that y_t is a poor proxy of future macroeconomic expectations. In order to investigate this possibility, in this appendix we replace y_t by an explicit measure of macroeconomic expectations by using forecasts of the macroeconomic variables. In particular, we replace first equation of our TVP-VAR by

$$x_t = \lambda_t^y \tilde{y}_t + \lambda_t^f f_t + u_t, \quad (\text{C.1})$$

where \tilde{y}_t contains one-year ahead forecasts of the macroeconomic variables made provided by the Survey of Professional Forecasters (SPF).¹³ For the second model

¹³This data is available at <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/historical-data/mean-forecasts.cfm>

equation, we maintain the assumption that the FCI f_t and the three macroeconomic variables of interest y_t follow a VAR of the form

$$\begin{bmatrix} y_t \\ f_t \end{bmatrix} = c_t + B_{t,1} \begin{bmatrix} y_{t-1} \\ f_{t-1} \end{bmatrix} + \dots + B_{t,p} \begin{bmatrix} y_{t-p} \\ f_{t-p} \end{bmatrix} + \varepsilon_t. \quad (\text{C.2})$$

In Table C2, we present the forecasting results for this model, which we name the TVP-FAVAR-SPF to denote the fact that we use the SPF in the measurement equation, and for the TVP-FAVAR and FAVAR as implemented in Section 3.4. Both TVP-FAVAR models have $\kappa_1 = \kappa_2 = 0.96$ and $\kappa_3 = \kappa_4 = 0.99$, but we additionally show results for the FAVAR-SPF which is the special case with $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 = 1$ (a constant parameter FAVAR with the SPF variables used to purge the FCI). We can see that using the SPF variable in order to purge the FCI, results in a slight deterioration of forecast performance. It is only when using DMA with the TVP-FAVAR-SPFs, that we see improvement over our benchmark TVP-FAVAR, but DMA led to similar improvements in the body of the paper (i.e. it is the use of DMA which is causing these improvements, not the inclusion of the SPF variables).

Table C2: Comparison with the FAVAR with macroeconomic expectations (SPF), 1990Q1-2013Q3

Forecast Metric	APL					MSFE				
	INFLATION									
	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
TVP-FAVAR	1.1233	0.9741	0.9066	0.8194	0.7380	0.0416	0.0454	0.0466	0.0562	0.0696
FAVAR	0.76	0.77	0.76	0.78	0.80	1.06	1.14	1.10	1.02	0.94
TVP-FAVAR-SPF	1.01	0.99	1.00	1.00	0.99	1.01	0.99	1.01	0.98	1.03
FAVAR-SPF	1.00	0.98	0.99	1.01	1.02	1.16	1.22	1.23	1.20	1.15
TVP-FAVAR-DMA-SPF	1.12	1.12	1.13	1.15	1.16	1.05	1.04	1.10	1.09	1.05
TVP-FAVAR-BMA-SPF	1.14	1.08	1.15	1.15	1.12	1.16	1.05	1.11	1.20	1.18
	UNEMPLOYMENT									
	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
TVP-FAVAR	0.9170	0.5868	0.4191	0.3229	0.2639	0.0906	0.2741	0.5950	1.0105	1.5359
FAVAR	0.64	0.61	0.61	0.63	0.64	1.68	2.01	2.09	2.14	2.17
TVP-FAVAR-SPF	0.99	0.99	0.99	1.00	0.98	1.02	1.05	1.04	1.04	1.04
FAVAR-SPF	0.96	0.89	0.91	0.93	0.93	1.27	1.43	1.45	1.45	1.43
TVP-FAVAR-DMA-SPF	1.14	1.14	1.15	1.18	1.16	0.90	0.84	0.80	0.81	0.82
TVP-FAVAR-BMA-SPF	1.09	1.09	1.10	1.09	1.18	0.91	0.83	0.76	0.74	0.75
	OUTPUT									
	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
TVP-FAVAR	0.4643	0.4344	0.4206	0.4156	0.4131	0.3548	0.4085	0.4204	0.3972	0.3854
FAVAR	0.71	0.71	0.72	0.73	0.74	1.43	1.64	1.57	1.48	1.40
TVP-FAVAR-SPF	0.99	0.99	1.00	1.00	1.00	1.04	1.02	1.01	1.02	1.01
FAVAR-SPF	0.89	0.90	0.92	0.95	0.96	1.18	1.32	1.25	1.19	1.17
TVP-FAVAR-DMA-SPF	1.08	1.09	1.09	1.09	1.08	1.02	1.01	1.01	1.04	1.05
TVP-FAVAR-BMA-SPF	1.02	1.11	1.12	1.04	1.11	1.04	0.98	0.96	1.00	1.02

Notes: APL is the average predictive likelihood (not in logarithms), and MSFE is the mean squared forecast error. Model's forecast performance is better when APL (MSFE) is higher (lower). For each variable (inflation, unemployment, output) the first line shows the APL and MSFE of the benchmark model for each forecast horizon h . All other models' APL and MSFE are relative to that of the benchmark model. Values of APL (MSFE) higher (lower) than 1 signify better performance than the benchmark.

In general, from all the models we have attempted (also for different values of forgetting/decay factors or priors), we found that the benchmark specification consistently performs better than the specification in equations (C.1)-(C.2) which utilizes survey information.

C.3. Comparison of different rates model switching

Calculation of the DMA/DMS time-varying probabilities depends on selection of a hyperparameter α , which is a forgetting factor that determines how fast we “forget” past observations. Thus, this hyperparameter α controls how much data we use for the calculation of time-varying probabilities at time t and, thus, determines the rate of model switching. That is, as α gets lower only the most recent data are used and older data are discounted towards zero at a faster rate, resulting in model switching occurring at a faster rate.

We remind the reader that $\alpha = 1$ leads to standard BMA. The benchmark results reported in the body of the paper, using $\alpha = 0.99$, allow for slightly more rapid model switching. In this appendix we provide results for choices of α which reflect beliefs about even faster model switches, namely $\alpha = 0.96$, as well as the extreme case of $\alpha = 0.001$. Note that this latter case is close to the case of model averaging using equal weights, since it is trivial to prove that as $\alpha \rightarrow 0$ then the model weights degenerate to $\pi_{t|t-1,j} \rightarrow 1/J$ for all $j = 1, \dots, J$. Naive model averaging schemes using equal weights have been shown in many cases to perform better than more elaborate econometric techniques that perform estimation of the model averaging weights; see Aiolfi, Capistrán and Timmermann (2010).

Table C3 shows the predictive likelihoods and mean squared forecast errors for the TVP-FAVAR model with DMA implemented for various values of forgetting factor α . As in the previous tables in this appendix, the first line shows the results for the benchmark TVP-FAVAR where all 18 variables are used to extract the FCI. All models have the benchmark prior described in the main text, plus the benchmark choices for decay factors of $\kappa_1 = \kappa_2 = 0.96$ and $\kappa_3 = \kappa_4 = 0.99$. We can immediately observe that the case $\alpha = 0.001$ results, in general, in much higher relative MSFEs for most horizons, for all three variables. The case $\alpha = 0.96$ sometimes does better than $\alpha = 0.99$ implying that there are periods in our sample that allowing for faster model switching would be beneficial. Therefore, there could potentially be further improvements in forecast accuracy by estimating α using grid-search methods, or even allowing a different value of α for each forecasting equation. We remind the reader that we do not perform such a search due to the already high computational demands of our empirical exercise, and we refer to Koop and Korobilis (2013) for an example of using such a procedure.

Table C3: Comparison of faster/slower model switching, 1990Q1-2013Q3

Forecast Metric	APL					MSFE				
	INFLATION									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	1.1233	0.9741	0.9066	0.8194	0.7380	0.0416	0.0454	0.0466	0.0562	0.0696
TVP-FAVAR-BMA ($\alpha = 1$)	1.01	1.02	1.03	1.02	1.00	1.03	1.00	1.06	1.04	1.00
TVP-FAVAR-BMS ($\alpha = 1$)	1.06	1.06	1.07	1.08	1.08	1.11	1.07	1.19	1.21	1.16
TVP-FAVAR-DMA ($\alpha = 0.99$)	1.09	1.09	1.10	1.12	1.12	1.03	1.00	1.05	1.03	0.99
TVP-FAVAR-DMS ($\alpha = 0.99$)	1.12	1.10	1.12	1.14	1.14	1.12	1.09	1.20	1.23	1.17
TVP-FAVAR-DMA ($\alpha = 0.96$)	1.05	1.07	1.09	1.06	1.07	1.04	1.00	1.05	1.02	0.98
TVP-FAVAR-DMS ($\alpha = 0.96$)	1.10	1.12	1.12	1.14	1.16	1.18	1.11	1.21	1.22	1.19
TVP-FAVAR-DMA ($\alpha = 0.001$)	1.02	1.01	1.05	1.03	1.04	1.47	1.42	1.41	1.21	1.15
TVP-FAVAR-DMS ($\alpha = 0.001$)	1.06	1.08	1.10	1.09	1.13	1.47	1.25	1.42	1.45	1.40
	UNEMPLOYMENT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	0.9170	0.5868	0.4191	0.3229	0.2639	0.0906	0.2741	0.5950	1.0105	1.5359
TVP-FAVAR-BMA ($\alpha = 1$)	1.01	1.01	0.97	0.96	0.90	0.91	0.85	0.82	0.83	0.84
TVP-FAVAR-BMS ($\alpha = 1$)	1.07	1.05	1.04	1.03	1.00	0.97	0.88	0.83	0.83	0.84
TVP-FAVAR-DMA ($\alpha = 0.99$)	1.11	1.10	1.09	1.10	1.07	0.92	0.86	0.83	0.84	0.85
TVP-FAVAR-DMS ($\alpha = 0.99$)	1.08	1.09	1.09	1.11	1.10	0.99	0.86	0.79	0.80	0.81
TVP-FAVAR-DMA ($\alpha = 0.96$)	1.05	1.09	1.06	1.08	1.10	0.89	0.82	0.80	0.80	0.82
TVP-FAVAR-DMS ($\alpha = 0.96$)	1.13	1.12	1.13	1.19	1.14	0.95	0.84	0.76	0.74	0.75
TVP-FAVAR-DMA ($\alpha = 0.001$)	0.99	1.00	1.01	0.95	1.02	1.08	0.89	0.82	0.81	0.83
TVP-FAVAR-DMS ($\alpha = 0.001$)	1.08	1.04	1.07	1.12	1.03	1.03	0.88	0.77	0.76	0.76
	OUTPUT									
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
TVP-FAVAR	0.4643	0.4344	0.4206	0.4156	0.4131	0.3548	0.4085	0.4204	0.3972	0.3854
TVP-FAVAR-BMA ($\alpha = 1$)	0.95	0.94	0.93	0.91	0.91	0.97	0.95	0.94	0.97	0.98
TVP-FAVAR-BMS ($\alpha = 1$)	1.00	1.00	0.99	0.99	0.99	1.01	0.99	0.98	1.01	1.00
TVP-FAVAR-DMA ($\alpha = 0.99$)	1.03	1.04	1.03	1.03	1.04	0.96	0.94	0.94	0.97	0.98
TVP-FAVAR-DMS ($\alpha = 0.99$)	1.05	1.06	1.06	1.08	1.09	0.99	0.95	0.95	1.00	1.00
TVP-FAVAR-DMA ($\alpha = 0.96$)	1.03	1.04	1.04	1.02	1.01	1.01	1.00	1.00	1.03	1.04
TVP-FAVAR-DMS ($\alpha = 0.96$)	1.10	1.07	1.08	1.12	1.11	1.04	0.99	0.97	1.01	1.03
TVP-FAVAR-DMA ($\alpha = 0.001$)	0.96	0.98	0.96	0.96	0.94	1.06	1.03	1.04	1.05	1.10
TVP-FAVAR-DMS ($\alpha = 0.001$)	1.08	0.97	0.99	1.06	1.04	1.08	1.03	1.00	1.03	1.05

Notes: APL is the average predictive likelihood (not in logarithms), and MSFE is the mean squared forecast error. Model's forecast performance is better when APL (MSFE) is higher (lower). For each variable (inflation, unemployment, output) the first line shows the APL and MSFE of the benchmark model for each forecast horizon h . All other models' APL and MSFE are relative to that benchmark model. Values of APL (MSFE) higher (lower) than 1 signify better performance than the benchmark.

C.4. Comparison of different rates of parameter change

Similar to the hyperparameters that control model switching, $\kappa_1, \kappa_2, \kappa_3, \kappa_4$ control the amount of time-variation in the error covariances (V_t, Q_t) , as well as the time-varying loadings λ_t and the VAR coefficients β_t . Table C4 presents results for different choices of these decay and forgetting factors. We only present results for the TVP-FAVAR models with all 18 variables used to extract the FCI, and not for the more computationally intensive DMA/DMS variants of the TVP-FAVAR (which allow selection of the optimal number of variables to include in the FCI). The first line again shows the benchmark TVP-FAVAR with $\kappa_1 = \kappa_2 = 0.96$ and $\kappa_3 = \kappa_4 = 0.99$. As

for the other tables in this appendix, we present relative APLs and relative MSFEs, for the evaluation period 1990Q1-2013Q3.

Table C4: Comparison of faster/slower parameter switching, 1990Q1-2013Q3

Forecast Metric	APL					MSFE				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
	INFLATION									
FAVAR (1,1)	0.8584	0.7474	0.6924	0.6427	0.5934	0.0438	0.0524	0.0542	0.0604	0.0686
TVP-FAVAR (0.96,1)	1.36	1.32	1.34	1.30	1.30	1.12	1.05	1.01	0.99	1.00
TVP-FAVAR (0.94,1)	1.49	1.45	1.49	1.47	1.46	1.12	1.02	0.98	0.94	0.98
TVP-FAVAR (0.92,1)	1.53	1.51	1.52	1.52	1.50	1.11	1.02	0.99	0.97	1.00
TVP-FAVAR (1,0.99)	1.01	1.01	1.02	1.00	1.00	0.97	0.96	0.97	0.98	1.02
TVP-FAVAR (0.96,0.99)	1.34	1.34	1.35	1.32	1.29	0.95	0.86	0.86	0.93	1.01
TVP-FAVAR (0.94,0.99)	1.46	1.44	1.48	1.45	1.43	1.05	0.95	0.92	0.88	0.95
TVP-FAVAR (0.92,0.99)	1.50	1.47	1.49	1.49	1.45	1.08	0.99	0.96	0.94	0.99
	UNEMPLOYMENT									
FAVAR (1,1)	0.5914	0.3567	0.2557	0.2024	0.1676	0.1200	0.4102	0.9406	1.6578	2.5544
TVP-FAVAR (0.96,1)	1.65	1.77	1.83	1.87	1.88	0.70	0.59	0.56	0.56	0.56
TVP-FAVAR (0.94,1)	1.74	1.90	1.95	2.00	1.99	0.71	0.60	0.56	0.55	0.55
TVP-FAVAR (0.92,1)	1.74	1.89	1.96	2.01	2.03	0.73	0.62	0.58	0.56	0.56
TVP-FAVAR (1,0.99)	1.09	1.13	1.18	1.21	1.23	1.03	1.01	0.99	0.97	0.97
TVP-FAVAR (0.96,0.99)	1.60	1.73	1.76	1.74	1.75	0.69	0.59	0.55	0.54	0.54
TVP-FAVAR (0.94,0.99)	1.73	1.85	1.88	1.90	1.87	0.65	0.53	0.51	0.51	0.52
TVP-FAVAR (0.92,0.99)	1.77	1.94	1.95	1.98	1.95	0.65	0.53	0.50	0.50	0.51
	OUTPUT									
FAVAR (1,1)	0.3276	0.3071	0.3031	0.3016	0.3041	0.3694	0.5209	0.5542	0.5181	0.4967
TVP-FAVAR (0.96,1)	1.54	1.56	1.51	1.53	1.50	0.91	0.78	0.77	0.81	0.84
TVP-FAVAR (0.94,1)	1.63	1.63	1.62	1.62	1.62	0.90	0.78	0.76	0.80	0.84
TVP-FAVAR (0.92,1)	1.66	1.68	1.66	1.68	1.68	0.92	0.77	0.76	0.80	0.83
TVP-FAVAR (1,0.99)	1.14	1.14	1.13	1.12	1.11	1.13	1.01	0.95	0.92	0.91
TVP-FAVAR (0.96,0.99)	1.51	1.51	1.49	1.48	1.46	0.91	0.75	0.72	0.75	0.77
TVP-FAVAR (0.94,0.99)	1.59	1.58	1.57	1.55	1.55	0.91	0.76	0.74	0.79	0.83
TVP-FAVAR (0.92,0.99)	1.62	1.64	1.62	1.60	1.61	0.91	0.75	0.74	0.78	0.82

Notes: APL is the average predictive likelihood, and MSFE is the mean squared forecast error. Model's forecast performance is better when APL (MSFE) is higher (lower). For each variable (inflation, unemployment, output) the first line shows the APL and MSFE of the benchmark model for each forecast horizon h . All other models' APL and MSFE are relative to that of the benchmark model. Values of APL (MSFE) higher (lower) than 1 signify better performance than the benchmark.

It is interesting to note that our conservative benchmark prior on time-variation in the volatilities ($\kappa_1 = \kappa_2 = 0.96$) is globally dominated by a prior which allows faster switching in volatilities ($\kappa_1 = \kappa_2 = 0.92$). This result is not surprising, as there is ample evidence for a high degree of volatility in macroeconomic and financial data (usually modeled in the literature as a geometric random walk, or a persistent AR(1) process). It would be surprising to find support for large time-variation in the time-varying parameters λ_t, β_t . However, we do not find such support. In terms of predictive likelihoods results are similar for $\kappa_3 = \kappa_4 = 1$ (constant parameters) and $\kappa_3 = \kappa_4 = 0.99$, while MSFEs seem to favor a small

degree of time-variation. However, values of κ_3 and κ_4 of 0.98 or lower, forecasting results deteriorate dramatically, showing the consequences of allowing too large a degree of time variation in the VAR coefficients and factor loadings.