Decomposing Macroeconomic Uncertainty

Kevin Sheppard*

Department of Economics, University of Oxford

kevin.sheppard@economics.ox.ac.uk

January 2017

Abstract

The paper proposes a method to decompose the total uncertainty in a large panel of time series into common and idiosyncratic components. The decomposition uses factors extracted using standard econometric techniques appropriate for long, wide panels. The factors extracted are then rotated to create a set of volatility factors that satisfy appropriate orthogonality conditions. The model proposed for common factor volatility is not required to be fully heteroskedastic and so a subset of the rotated factors may be conditionally homoskedastic. Estimation of model parameters is described using either maximum likelihood estimation or composite maximum likelihood estimation. The model is applied to a set of monthly macroeconomic time series spanning 1959 until 2017. The data support between 2 and 4 volatility factors. These volatility factors drive the vast majority of the dynamics of uncertainty. Measuring common factor variance provides a much sharper measure of uncertainty than a measure based on the total volatility of the individual series.

Keywords: ARCH, Dynamic Factor Models, Macroeconomic Uncertainty, Volatility

JEL Codes: C33, C53, E37

*I am grateful for feedback received at seminars at Pompeu Fabra University (UPF) and Aarhus.

1 Introduction

The role of uncertainty in economics decisions has been an increasing focus over the past two decades. Uncertainty can increase the probability of losses when investments are irreversible and so may lead to large swings in aggregate activity. Recently, Baker, Bloom, and Davis (2016) have introduced an index of policy uncertainty in an attempt to begin to quantify this uncertainty. Jurado, Ludvigson, and Ng (2015) proposed using the volatility in a large number of series to measure macroeconomic uncertainty. Their main measure of uncertainty is defined as

$$\mathcal{U}_{jt}^{y}(h) = \sqrt{E\left[\left(y_{jt+h} - E\left[y_{jt+h}|\mathcal{F}_t\right]\right)^2 \middle| \mathcal{F}_t\right]}$$

which is the conditional volatility of the residual for a single series h periods in the future. They then define aggregate uncertainty as

$$\mathcal{U}_{t}^{y}(h) = \text{plim}_{N_y\to\infty} \sum_{j=1}^{N_y} w_j \mathcal{U}_{jt}^{y}(h) = E_w\left[\mathcal{U}_{jt}^{y}(h)\right].$$

This uncertainty measure makes use of simple techniques to estimate factors from long, wide panels of macroeconomic time series (Bai and Ng 2002; Stock and Watson 2002). These models posit that observable series are contemporaneously related to common factors so that

$$y_{it} = \Lambda_i' f_t + \eta_t^i.$$

Principal component analysis can be used to extract theses common factors up to some normalization (scale and rotation).

The measurement of macroeconomics uncertainty has been an active field with many recent contributions. Bijsterbosch and Guérin (2014) study period of very high uncertainty in a panel framework using a regime switching model to capture the state of uncertainty. Jo and Sekkel (2017) use a factor stochastic volatility model in a small panel of series that have forecasts available to measure uncertainty. They show that there is a strong factor and that this factor is best identified using early vintage data. Rossi and Sekhposyan (2015) use mistakes relative to historical accuracy to characterize uncertainty. Rossi, Sekhposyan, and Soupre (2016) propose a new measure using the density of the errors of professional forecasters to separate Knightian uncertainty from risk.

The contribution of this paper is to refine an uncertainty measure in the spirit of Jurado, Ludvigson, and Ng (2015). I propose a joint model for the factor innovations and the individual macroeconomic time series innovations. This model attempts to stay close to the structure of factor models and volatility factors are assumed to have conditional covariance that is orthogonal. This requires rotating the factors computed using the usual normalization that arises from PCA extraction. I also explicitly consider that case where the number of volatility factors is less than that number of mean factors, potentially substantially. The model shares common features with both BEKK (Engle and Kroner 1995)and GO-GARCH (Van der Weide 2002). These models share a common structure in that they are closed to rotations. This is important when modeling the conditional covariance of factors since the choice of normalization is made for computational convenience rather than driven by identification concerns.

I propose a novel estimator using composite likelihood to estimate the model parameters using information in both the factor innovations as well as in the individual macroeconomic time series that underlie the factors. This estimator improves precision although the parameter estimates are similar to those found using more standard maximum likelihood estimation.

The paper is structured as follows. Section 2 describes the model. Second 3 discusses estimation and inference including using composite likelihood. Next, a test is proposed to examine whether the number of heteroskedastic factors is lower than the total number of factors in the model. Section 5 contains a Monte Carlo study of the estimator while section 6 describes the data used. Section 7 contains the main empirical findings and section 8 concludes.

2 Model

We consider a model where observable variables are driven by contemporaneous factors, f_t, where the process may have additional dynamics through an autoregressive component,

$$y_{jt} = \upsilon_j' f_t + \sum_{i=1}^{p_j} \phi_i y_{jt-1} + \xi_t^j \tag{1}$$

where ξ_t^j is a white noise error that is weakly correlated with the factors.[1] We further assume that factors can be well described by a vector autoregression so that

$$f_t = \sum_{i=1}^{p^f} \Phi_i f_{t-i} + \xi_t^f. \tag{2}$$

Eq. (1) is not a forecasting model since f_t is only time t-measurable. Substituting in for the contemporaneous factor allows a forecasting model to be specified,

$$\begin{aligned} y_{jt} &= \upsilon_j' \sum_{k=1}^{p_f} \Phi_k f_{t-k} + \sum_{j=1}^{p_j} \phi_j y_{jt-1} + \xi_t^j \\ &= \sum_{k=1}^{p_F} \Upsilon_k f_{t-k} + \sum_{j=1}^{p_j} \phi_j y_{jt-1} + \upsilon_j' \xi_t^F + \epsilon_t^j \end{aligned}$$

Let the conditional covariance of the factor residuals and the residual from the observable series by specified as

$$\text{Cov} \begin{bmatrix} \xi_{t+1}^j \\ \xi_{t+1}^f \end{bmatrix} |\mathcal{F}_t \equiv \Sigma_{t+1} = \begin{bmatrix} \zeta_{jt+1}^2 & \Sigma_{t+1}^{jf} \\ \Sigma_{t+1}^{fj} & \Sigma_{t+1}^f \end{bmatrix}.$$

The time-t conditional variance of the forecast residuals of the observable variables y_{jt} can be decomposed into three components: idiosyncratic variance, factor covariance and any covariance between the factor shocks and the idiosyncratic shock.

[1] We assume that both y and f have unconditional mean 0. This is imposed by demean the data.

$$\textbf{Var}\left[y_{jt+t}|\mathcal{F}_t\right] \equiv \textbf{Var}\left[\xi^j_{t+t}|\mathcal{F}_t\right] = \zeta^2_{jt+1} = \tau^2_{jt+1} + \upsilon_j'\Sigma^F_{t+1}\upsilon_j + 2\upsilon_j'\Sigma^{Fj}_{t+1}.$$

The approach taken in Jurado, Ludvigson, and Ng (2015) model the total variance of the error, including the factor variance, the idiosyncratic variance and the covariance between the two. This may mask the important dynamics of macroeconomic uncertainty for two reasons. First, some of the components of the conditional variance may not be dynamic, and so averaging across these will attenuate shifts in uncertainty. Second, the variance dynamics of the total error may have complex dynamics even if the components can be modeled using a simpler specification. This decomposition allows the components of macroeconomic uncertainty to be identified using the conditional covariance of the factors and the idiosyncratic shocks. When the simplifying assumption is made that the idiosyncratic shock has no conditional correlation with the factors, then a natural measure of the importance of the common exposure can be expressed

$$\mathcal{V}^j_i + \mathcal{V}^j_f = \frac{\tau^2_{jt+1}}{\tau^2_{jt+1} + \upsilon_j'\Sigma^F_{t+1}\upsilon_j} + \frac{\upsilon_j'\Sigma^F_{t+1}\upsilon_j}{\tau^2_{jt+1} + \upsilon_j'\Sigma^F_{t+1}\upsilon_j}.$$

Measuring these quantities requires models of these three components: factor covariance, loadings on the factors and idiosyncratic variance.

2.1 Fully Heteroskedastic Factor Models

Common factor extraction methods use convenient, but arbitrary, normalization to identify the factors. For example, when using PCA is it common to normalize the factors so that so that the factors are orthogonal and factor loadings are orthonormal ($F'F$ is diagonal and $\Lambda'\Lambda = I$). When factors are conditionally heteroskedastic, it is desirable to use a simple model to specify the dynamics. Suppose there exists a rotation of the factors, denoted $\tilde{F}$, have innovations that follow scalar univariate GARCH models. Denote these innovations as ξ_t^f so that

$$\xi_t^f|\mathcal{F}_{t-1} \sim G\left(0, \begin{bmatrix} \sigma_{1t}^2 & 0 & 0 & 0 \\ 0 & \sigma_{2t}^2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_{kt}^2 \end{bmatrix}\right)$$

where G is some distribution with finite second moments. Further suppose the conditional variance for each component follows

$$\sigma_{it}^2 = \omega_i + \alpha_i \left(\xi_{it-1}^f\right)^2 + \beta_i \sigma_{i,t-1}^2. \tag{3}$$

Both the conditional and unconditional correlations are assumed to be zero in this specification. The assumption that the innovations are unconditionally correlated is insufficient to uniquely determine the space of the residuals.[2] Suppose that the unconditional covariance of the residuals, $\overline{\Sigma}$ was diagonal. Then $Q\overline{\Sigma}^{-1/2}\xi_t^f$ would have a covariance equal to the identity matrix for any orthonormal rotation Q.

Identification of the volatility factors relies on mispecification of the model when a rotation of the factors is used that differs from true rotation – imposing the wrong rotation will not produce the correct conditional covariance. In particular, assuming that the unconditional variance of each series is unity and defining $\tilde{\xi}_t^f = Q\xi_t^f$ for some rotation matrix Q different from the identity, imposing simple dynamics on the conditional covariance will result in a

[2]Many applications of factor analysis are interested in identifying the factors up to some arbitrary rotation. Here the identification of the factors would be up to a set of labels since, even if the specific series can be identified, one can always swap the position of the series by relabeling.

specification of the form

$$\tilde{\sigma}_{it}^2 = \tilde{\omega}_{it} + \tilde{\alpha}_i \left(\xi_{it-1}^f\right)^2 + \tilde{\beta}_i \tilde{\sigma}_{it-1}^2 \tag{4}$$

$$\tilde{\sigma}_{it}^2 = \tilde{\omega}_{it} + \tilde{\alpha}_i \left(\sum_{j=1}^{k} Q_{[ij]} \xi_{jt-1}^f\right)^2 + \tilde{\beta}_i \tilde{\sigma}_{it-1}^2 \tag{5}$$

When all elements of Q are non-zero, fitting the model with the incorrectly rotated innovations will include the squares and cross-products of all of the factors weighted by the elements of Q, and so it will not be possible to correctly specify the variance dynamics.

Rather than attempting to specify the rotation to correlctly fit the model, it may be possible to recover the *volatility space*. The volatility space is similar to the factor space since it is identified up to a rotation. Directly building on eq. (4), it may be feasible to fit

$$\tilde{\sigma}_{it}^2 = \tilde{\omega}_{it} + \sum_{j=1}^{k} \sum_{l=j+1}^{k} \ddot{\alpha}_{jl} \tilde{\xi}_{jt-1}^F \tilde{\xi}_{lt-1}^F + \tilde{\beta}_i \tilde{\sigma}_{it-1}^2$$

which has an innovation term that nests the innovation term in eq. (4). The extra flexibility added by allowing fully general innovation terms – in both squares and cross-products – is, however, still not sufficient to produce a correct fit except under an additional assumption that the values of β are common across factors in eq. (3).

The incorrectly rotated model is a special case of a general BEKK model (Engle and Kroner 1995). BEKK models have the unique property (among multivariate ARCH models) of being closed to rotations. The BEKK is not directly applicable if conditional correlations are assumed to be 0 since it would only naturally assume unconditional correlations are zero. Suppose that the true factor residuals had conditional covariance dynamics that could be described by

$$\Sigma_t = CC' + A\xi_{t-1}^F \left(\xi_{t-1}^F\right)' A' + B\Sigma_{t-1}B'$$

where $CC' = I_k - A - B$, I_k is the identity matrix and A and B are diagonal. The conditional variances in this specification are identical to the variance dynamics in eq. (3). However, the conditional covariance also follow similar dynamics so that

$$\sigma_{ijt} = \left(I - A_{ij}^2 - B_{ij}^2\right) \mathbb{I}_{[i=j]} + A_{ii} A_{jj} \xi_{it-1}^F \xi_{jt-1}^F + B_{ii} B_{jj} \sigma_{ijt-1}$$

where $\mathbb{I}$ is an indicator functions. When the true factor residuals are rotated by a non-identity matrix Q , then the conditional covariance

$$\begin{aligned} Q\Sigma_t Q' &= QCC'Q' + QAQ\xi_{t-1}^F \left(Q\xi_{t-1}^F\right)' A'Q' + QBQ\Sigma_{t-1}Q'B'Q' \\ \tilde{\Sigma}_t &= \tilde{C}\tilde{C}' + \tilde{A}\tilde{\xi}_{t-1}^F \left(\tilde{\xi}_{t-1}^F\right)' \tilde{A}' + \tilde{B}\tilde{\Sigma}_{t-1}\tilde{B}'. \end{aligned}$$

This model can be verified to nest the simple specification in eq. (3) if the matrices A and B are constructed such that the columns are mutually orthogonal. In this specification the parameter matrices $\tilde{C}$, $\tilde{A}$ and $\tilde{B}$ will generally be full even though the original parameters matrices were diagonal.[3] The model specification in eq. (3) also allows the possibility of have reduced dimension dynamics in this model since some of the factors' variances (and covariance) would be constant. While it is appealing in practice to consider that the volatility process is lower-dimension than the number of factor innovations, recovering the volatility space requires fitting a full model even if there was only a single dynamic factor.

2.2 Volatility Factor Models

The initial specification will make use of a number of simplifying assumption. First, idiosyncratic shocks will be assumed to be conditionally homoskedastic. Second, the number

[3] Formally each parameter matrix will have $\left(k^2 + k\right)/2$ parameters to account for the restrictions on Q. There are also cross-parameter restrictions to account for the common rotation when directly implementing this model.

of volatility factors, m, will typically be less than the number of factors in the model. The 1-volatility-factor model is specified as

$$Q_1\xi_t^f \sim G\left(0, \begin{bmatrix} \sigma_{1t}^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_k^2 \end{bmatrix}\right)$$

where Q_1 is an orthonormal rotation matrix that will uniquely identify the first factor but notht he remaining $k-1$ factors.

Factor residuals are assumed to be unconditionally orthogonal and have unit variance. This is a simplifying assumption and has no consequence on the ability of the factor model to fit the conditional mean of the data since rotating any set of correlated factor residuals can be directly implemented by rotating the factors. When there are k mean factors (the factors used to model the conditional mean of the panel of variables) but there is only 1 volatility factor, it isn't possible to fully identify Q since many rotation matrices can be constructed that uniquely identify the volatility factor in the first position but will have an identical covariance structure for the remaining factor residuals. For example, defining

$$\ddot{Q} = Q \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{bmatrix}$$

where $\tilde{Q}$ is any non-identify rotation of size $k-1$would preserve the dynamics in the first factor while applying some other rotation to the remaining $k-1$ factors. The rotation used in the 1-volatility-factor model consists of $k-1$ rotation parameters which uniquely define a unit vector as the first column of Q and then the remainder of the matrix is assured to be orthogonal to the first factor but is otherwise unspecified.

Models with multiple factors allow additional flexibility when choosing Q. The general

m-factor volatility model has dynamics given by

$$Q_m \xi_t^f \sim G\left(0, \left[\begin{array}{cccccc} \sigma_{1t}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{mt}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{m+1}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_k^2 \end{array}\right]\right) \tag{6}$$

where Q_m uniquely identified the first m factors, and the remaining $k - m$ are assumed to be homoskedastic. In general, a m-factor model requires $mk - m(m+1)/2$ parameters when specifying Q_m. When $m = k - 1$ then the parameters full identify all elements of Q_m.

2.3 Utilizing the Full Panel

Under an assumption that the idiosyncratic shocks to the individual series are uncorrelated with the factor residuals, the joint covariance of the factors and the individual model shocks covariance can be specified as

$$\Sigma_t = \left[\begin{array}{cc} \Sigma_t^i & \Sigma_t^f \upsilon \\ \upsilon' \Sigma_t^f & \Sigma_t^f \end{array}\right]$$

where Σ_t^f is defined in eq. (6) and Σ_t^i is the conditional covariance of the combined idiosyncratic shock that include the common factors $\left(\xi_t^j\right)$. When the components of the idiosyncratic errors that is uncorrelated with the factor shocks $\left(\epsilon_t^i\right)$ are homoskedastic so that $E\left[\left(\epsilon_t^i\right)^2 | \mathcal{F}_{t-1}\right] = \tau_i^2$, the the conditional covariance of the individual series forecast residuals is

$$\sigma_{it}^2 = \upsilon' \Sigma_t^f \upsilon + \tau_i^2. \tag{7}$$

This decomposition motivates using a complete covariance model for the joint dynamics of the factor innovations, ξ_t^f and the individual series innovations, ξ_t^i. Expanding from the core model with k mean factors and m variance factors, this model will add the k by N array of factor loadings υ and the N parameters of τ^2.

3 Estimation and Inference

Conditional Mean

Estimation of all specifications is done using Gaussian Quasi MLE where residuals are assumed to be conditionally normally distributed. All model make use of residuals filtered by mean equations. In order to ensure comparability across specifications, all model residuals are computed in the same manner. k factors are extracted from the cross-section of individual time series. The number of factors utilized ranges from 3 – 6. This range of factors has commonly been found to be beneficial in forecasting time series and out-of-sample forecasting results are broadly similar. Factors are extracted from mean-zero unit variance data using principal components. All series are used to extract factors and the EM algorithm is used to address the problem of missing data The EM algorithm estimates PC by iteratively filling missing values with their expectation given the current factors, and then extracting the factors from the augmented data set. The factor residuals are computed using a first-order vector autoregression,

$$f_t = \Phi f_{t-1} + \xi_t^f.$$

Idiosyncratic residuals for observable series i are computed using a model which regresses each series on the first lag of the factors and $p_i$lags of series i,

$$y_{it} = \psi' f_{t-1} + \sum_{j=1}^{p_i} \phi_j y_{it-1} + \xi_t^i.$$

The lag length p_i was selected from $\{0,\ldots,12\}$ by minimizing the Schwartz/Bayes Information Criteria. Estimating these $n+1$ models produced the residuals $\hat{\xi}_t$ that are used in all of the volatility estimation models. For simplicity of exposition, the hat is suppressed in the discussion of the volatility estimation.

Maximum Likelihood Estimation

Model parameters are estimated using maximum likelihood or composite maximum likelihood. The latter utilizes information in the components series when estimating the parameters of the model. I will first present the estimation and inference details for the MLE before turning to the CMLE.

The model has three distinct groups of parameters. The first, $\bar{\Sigma}$, is an estimate of the long-run covariance of the factor residuals, ξ_t^f. Denote the unique parameters of this matrix as $\psi_{\bar{\Sigma}} = \text{vech}\left(\bar{\Sigma}\right)'$. The second are the parameters of the volatility dynamics, α_i, β_i for $i = 1, 2, \ldots, m$ where there are m volatility factors. These parameters are restricted to be non-negative and $\alpha_i + \beta_i < 1$ for $i = 1, 2, \ldots m$. These parameters are collectively denoted $\psi_{\alpha\beta} = (\alpha_1, \beta_1, \ldots, \alpha_m, \beta_m)'$. The final set of parameters are the angles used to rotate the m factors with conditional heteroskedasticity. When there are k mean factors and m volatility factors, the model will include $m\left(2k-m-1\right)/2$ angles where each angles $\theta_i \in (-\pi, \pi)$. Let $\psi_\theta = \left(\theta_1, \ldots, \theta_{m(2k-m-1)/2}\right)'$. The complete set of model parameters is then $\psi = \left(\psi_{\bar{\Sigma}}', \psi_{\alpha\beta}', \psi_\theta'\right)'$.

The parameters are estimated in two stages. The first stage uses a simple moment-based estimator for $\psi_{\bar{\Sigma}}$,

$$\hat{\bar{\Sigma}} = T^{-1} \sum_{t=1}^{T} \xi_t^f \left(\xi_t^f\right)'.$$

The standardized residuals are then

$$\dot{\xi}_t^f = \left(\hat{\bar{\Sigma}}^f\right)^{-1/2} \xi_t^f,$$

which have an unconditional covariance equal to an identity matrix. This transformation has no effect on the conditional mean models since these are the residuals that would occur if $\left(\hat{\bar{\Sigma}}^f\right)^{-1/2} f_t$ was used in place of f_t . As long as $\bar{\Sigma}^f$ is full rank, the fit of the models, and the estimated residuals (up to a factor $\left(\hat{\bar{\Sigma}}^f\right)^{1/2}$), are unaffected by this alternative normalization. The volatility factor model is then estimated using $\dot{\xi}_t^f$. The parameters of the model are the estimated by maximizing the log-likelihood,

$$\begin{aligned}\max_{\psi_{\alpha\beta},\psi_\theta} l\left(\dot{\xi}^f;\psi_{\alpha\beta},\psi_\theta\right) &= -\frac{1}{2T}\sum_{t=1}^{T}\left(c_k + \ln|\Sigma_t| + \left(\dot{\xi}_t^f Q_m\right)'\Sigma_t^{-1}\left(\dot{\xi}_t^f Q_m\right)\right) \qquad (8)\\ &= -\frac{1}{2T}\sum_{t=1}^{T}\left(c_k + \sum_{i=1}^{m}\ln\sigma_{it}^2 + \sum_{j=1}^{m}\frac{\left(\dot{\xi}_t^f Q_m^{(j)}\right)^2}{\sigma_{jt}^2} + (k-m)\right)\end{aligned}$$

where $Q_m^{(j)}$ is column j of Q_m and c_k is a constant. The simplification of the log-likelihood is possible since the rotated factor covariance is diagonal and the remaining unit variances of the $k-m$ static factor variances. Ignoring constants, the this log-likelihood is also the log-likelihood of the first m components – those with dynamics. This is natural since, by construction, there is no information in the remaining $k-m$ components and so these make no contribution to the log-likelihood. The m conditional variances evolve according to

$$\sigma_{it}^2 = (1-\alpha_i-\beta_i) + \alpha_i\left(\dot{\xi}_{t-1}^f Q_m^{(j)}\right)^2 + \beta_i\sigma_{it-1}^2.$$

The rotation matrix Q_m are parameterized using $km - m(m+1)/2$ angles where each angle is restricted to be in $(-\pi,\pi)$ using Given's rotations. Specifically,

$$Q_m(\psi_\theta) = \prod_{i=1}^{m}\prod_{j=i+1}^{k} G_{ij}(\theta_{ij})$$

where $G_{ij}(\theta_{ij})$, $i<j$, is an identify matrix except for elements (i,i), (i,j), (j,i) and (j,j) which

are $\cos(\theta_{ij})$, $-\sin(\theta_{ij})$, $\sin(\theta_{ij})$ and $\cos(\theta_{ij})$, respectively. Givens rotations are simple methods to parameterize orthonormal matrices. Q_m is equivalent to a full set of givens rotations across all coordinates where angles used for the missing coordinates are set to 0.

The m-volatility factor model with k mean factors has $km - m(m-3)/2$ parameters the require optimization using some non-linear optimizer. This is linear in m and $O(1)$ in N, the number of observed series. For example, a small model with $k = 3$ and $m = 1$ will have 4 parameters while a complete model with $k = m = 6$ will have 27 parameters. While this structure is sufficient to uniquely identify the parameters of the model, there can be label swapping could be an issue during estimation where two components switch order in the product $Q_m \epsilon_t$. To avoid this issue, I order the volatility parameters by sorting $\alpha_i + \beta_i$.

Under standard regularity conditions (see Comte and Lieberman (2003) and **pedersen2014multivariate** the estimated parameter vector is $\hat{\psi}$ is asymptotically normal. Using standard asymptotic theory for 2-step estimators (Newey and McFadden 1994), the asymptotic distribution is

$$\sqrt{T}\left(\hat{\psi} - \psi_0\right) \xrightarrow{d} N\left(0, A^{-1} B \left(A'\right)^{-1}\right) \tag{9}$$

where

$$A = E\begin{bmatrix} \frac{\partial g}{\partial \psi_{\bar{\Sigma}}} & 0 & 0 \\ \frac{\partial^2 l}{\partial \psi_{\bar{\Sigma}} \partial \psi'_{\alpha\beta}} & \frac{\partial^2 l}{\partial \psi_{\psi_{\alpha\beta}} \partial \psi'_{\psi_{\alpha\beta}}} & \frac{\partial^2 l}{\partial \psi_{\theta} \partial \psi_{\alpha\beta}} \\ \frac{\partial^2 l}{\partial \psi_{\bar{\Sigma}} \partial \psi'_{\theta}} & \frac{\partial^2 l}{\partial \psi_{\psi_{\alpha\beta}} \partial \psi'_{\theta}} & \frac{\partial^2 l}{\partial \psi_{\theta} \partial \psi'_{\theta}} \end{bmatrix}$$

where

$$g\left(\xi_t^f; \bar{\Sigma}\right) = T^{-1} \sum_{t=1}^{T} g_t = T^{-1} \sum_{t=1}^{T} \text{vech}\left(\xi_t^f \left(\xi_t^f\right)' - \bar{\Sigma}\right)$$

is the moment function used to estimate the unconditional covariance. $\hat{A}$ is estimated using the sample analogue to the theoretical value.

B is the asymptotic covariance of the scores, and is defined

$$B = \text{avar}\left(T^{-1/2}\sum s_t\right)$$

where $s_t = \left(g_t', \frac{\partial l_t}{\partial \psi'}\right)'$ are the moment conditions and scores of the log-likelihood. Since the moment conditions are not martingale difference sequence when the model is correctly specified, a log-run covariance estimator is used. I use the estimator of Newey and West (1987) with bandwidth proportional to $\sqrt[3]{T}$.

Composite Likelihood Estimation

It may be possible to improve estimation using the entire cross-section including both the factor residuals as well as the forecast errors of the individual models. The full joint model includes the factor covariance as well as the covariance of the idiosyncratic terms. The full log-likelihood is likely challenging to optimize since the idiosyncratic innovations are not assumed to have zero covariance, and the number of individual series is large relative to the sample size. Moreover, there is little interest in the off-diagonal terms of this matrix when examining macroeconomic uncertainty. Recently, Pakel et al. (2017) have proposed a composite likelihood estimator for large N, large T panels. These estimators make use of simple moment-based estimators to avoid non-linear estimation in high-dimensional spaces. They also intentionally discard the information in the full cross-section and instead focus on "submodels" of the full model to estimate the unknown parameters.

Using the decomposition in eq. (7), the conditional variance of idiosyncratic shock i can be expressed as

$$\sigma_{it}^2 = \upsilon_i' \Sigma_t^f \upsilon_i + \tau_i^2$$

When estimating factor volatilities, the natural unit to estimate the common parameters is the $k+1$ block consisting of the factors and one of the individual series.

$$Cov\left[\begin{array}{c}\xi_t^f\\ \xi_t^i\end{array}|\mathcal{F}_{t-1}\right]=\left[\begin{array}{cc}\Sigma_t & \Sigma_t\upsilon_i\\ \upsilon_i'\Sigma_t & \varsigma_{it}^2\end{array}\right].$$

The log-likelihood for this sub-model is then

$$l_i\left(\left[\begin{array}{c}\xi_t^f\\ \xi_t^i\end{array}\right];\psi_{\alpha\beta},\psi_\theta\right)=T^{-1}\sum_{t=1}^{T}\ln|\Omega_{it}|+\left(\left(\xi_t^f\right)',\xi_t^i\right)\Omega_{it}^{-1}\left(\left(\xi_t^f\right)',\xi_t^i\right)'.$$

The information across the N models can be combined by averaging each of the sub-models to construct the composite likelihood,

$$cl\left(\xi^f,\xi^i;\psi_{\alpha\beta},\psi_\theta\right)=N^{-1}\sum_{i=1}^{N}l_i.$$

There are two sets of parameters constant parameters that are identified though moment conditions. The first is $\bar{\Sigma}$ which is estimated using the standard moment estimator for the covariance. The second is υ_i. These parameters are estimated using the regression-based estimator where ξ_t^i is regressed on ξ_t^f, so that

$$\hat{\upsilon}_i=\left(\sum_{t=1}^{T}\xi_t^f\left(\xi_t^f\right)'\right)^{-1}\left(\sum_{t=1}^{T}\xi_t^f\xi_t^i\right).$$

The estimator of υ_i, which easily written in the form of a standard linear regression, is only a function of the joint unconditional covariance matrix of the factor innovations and the idiosyncratic innovation, and so the moment-based estimator used in the j$^{\text{th}}$ composite likelihood component is

$$\hat{\bar{\Omega}}=T^{-1}\sum_{t=1}^{T}\left[\left(\xi_t^f\right)'\ \xi_t^i\right]'\left[\left(\xi_t^f\right)'\ \xi_t^i\right].$$

Inference on composite likelihood parameters is similar to procedure described in the MLE estimation section. Further details are presented in Pakel et al. (2017).

4 Testing the Number of Volatility Factors

The number of volatility factors is a model parameter. When the hypothesized number of volatility factors is less than the number of factors then estimation of the model with m=k volatility factors has some novel challenges. If the correct rotation were known, testing for volatility would be standard and any ARCH-LM type test could be used. However, when the correct rotation is not known, as is the assumed case, then it is not possible to test whether the loading on recent innovations is zero since this creates an unidentified nuisance parameter problem. The problem arises since when a subset of the series is conditionally homoskedastic, any orthonormal rotation will produce the same likelihood, and so the rotation can only be consistently estimated for the heteroskedastic factors.

A general test can be developed by ignoring the structure of the problem. Let Q_m is the rotation matrix for m volatility factors. This matrix can be decomposed as $Q_m = \left[Q_m^D \, Q_m^S\right]$ where Q_m^D are the columns that rotate the factors with dynamic volatilities and Q_m^S correspond to the factors that have static (homoskedastic) volatilities. When the true number of volatility factors, there is some matrix $\tilde{Q}_{k-m}$ that only rotates the final $k-m$ factors that will produce at least one additional volatility factor. However, under the null that the number of volatility factors is m this matrix cannot be identified. A natural solution is to sidestep the identification issue by over-parameterizing the dynamics that might be present in the hypothesized static series. This over-parameterized model will parameterize the upper bound of $k - m + m\,(m+1)\,/2$ parameters in the correct model with $(k-m)^2$ parameters. In particular, neglected dynamics in the static block much have shocks of the form

$$DQ_m^s \xi_t^f \left(Q_m^s \xi_t^f\right)' D'$$

where if there are additional factors then $D = \text{diag}(\delta)\tilde{Q}_{k-m}$ so that the final form of the innovations to variance will follow the model. Substituting in for D,

$$\text{diag}(\delta)\tilde{Q}_{k-m}Q_m^s\xi_t^f\left(\tilde{Q}_{k-m}Q_m^s\xi_t^f\right)'\text{diag}\,(\delta) = \text{diag}(\delta)\left(\ddot{Q}_k\xi_t^f\right)\left(\ddot{Q}_k\xi_t^f\right)'\text{diag}(\delta)$$

where $\ddot{Q}_k$ is also a rotation matrix. However, if the data in the remaining $k-m$ factors are homoskedastic, then all elements of D will be 0. This suggests specifying additional dynamics using $(k-m)^2$ parameters in D to avoid creating an unidentified nuisance parameter problem.

Remark 1. When the data in the remaining $k-m$ factors are heteroskedastic, it would normally be the case that lagged variance would be included in the model. This poses additional issues since when the data are homoskedastic then the parameter on lagged variance is not identified (Andrews 2001). I avoid this issue by using an ARCH-LM-type test that only included lags of innovations.

I implement this test by augmenting the model to allow for a lag of volatility in the assumed static factors. Parameters are estimated using 2-step MLE where $\bar{\Sigma}$ is estimated using moment estimator and the remainder of the parameters, $\psi_{\alpha\beta}$, ψ_θ and $\psi_\delta = \text{vec}(D)$ are estimated to maximize the log-likelihood

$$\max_{\psi_{\alpha\beta},\psi_\theta} l\left(\dot{\xi}^f;\psi_{\alpha\beta},\psi_\theta,\psi_\delta\right) = -\frac{1}{2T}\sum_{t=1}^{T}\left(c_k + \sum_{i=1}^{k}\ln\sigma_{it}^2 + \sum_{j=1}^{k}\frac{\left(\dot{\xi}_t^f Q_m^{(j)}\right)^2}{\sigma_{jt}^2}\right) \tag{10}$$

where the variance dynamics for the first m series are unchanged and the variance dynamics for σ_{jt}^2, $j = m+1,\dots,k$ are

$$\sigma_{jt}^2 = 1 - \sum_{i=1}^{k-m} D_{ji}^2 + \left(\sum_{l=1}^{k-m} D_{jl}\left(\dot{\xi}_t^f Q_m^{(m+l)}\right)\right)^2.$$

The variance dynamics of the $k-m$ additional series can be equivalently parameterized as the diagonal elements of

$$I_{k-m} - DD' + D\left(Q_m^S \dot{\xi}_t^f\right)\left(\dot{\xi}_t^f Q_m^S\right)' D'.$$

Inference is implemented using the asymptotic distribution in eq. (9) where the matrices are extended to account for ψ_δ. The hypothesis test for homoskedasticity of the remaining $k-m$ components is $H_0 : \delta = 0$ against an alternative $H_1 : \delta \neq 0$. The test is implemented as a Wald test

$$T\hat{\delta}'\left(S\hat{A}^{-1}\hat{B}\hat{A}^{-1}S\right)^{-1}\hat{\delta} \sim \chi^2_{(k-m)^2}$$

where S is a selection matrix that selects the final $(k-m)^2$ rows and columns of the covariance matrix.

5 Monte Carlo

A Monte Carlo study was conducted to assess the ability of the model to correctly estimate the volatility of the model under alternative number of factors when the number of volatility factors was less than the number of mean factors. Factor model residuals were simulated by generating m independent univariate GARCH processes and augmenting these series with $k-m$ homoskedastic series. The innovations in all series were normal. The initial T by k set of residuals ξ were then rotated by a rotation matrix Q, which was randomly drawn where each of the $mk - m(m+1)/2$ angles was uniformly distributed on $(-\pi, \pi)$. The first set of simulations considers models with between 3 and 6 mean factors and only 1 volatility factor. The parameters of the volatility factor were varies across $\alpha \in (.1, .15)$ and $\alpha + \beta \in (.9, .95, .99)$. All experiments were repeated 1,000 times, used 500 observations and 200 observations were burned when generating samples.

Table 2 contains parameter bias and RMSE for the volatility parameters as well as measures of the accuracy of the rotation parameters. The panels range across persistence and the size of the innovation shock. The number of factors has a small effect on the precision of the volatility parameters when the volatility is not highly persistent. In the specification with the highest persistence, $\alpha = 0.15$, $\beta = 0.84$, there is no effect of increasing the cross-sectional dimension. In all specifications, the volatility parameters were more precisely estimated when the persistence was higher and when the parameter on the recent innovation was larger. These two features both make the conditional volatility more apparent and so this is not surprising.

The final four columns provide measures of the parameters that determine the rotation. The columns labeled θ directly examine the precision of the model parameters used to compute the rotation while the columns labeled Q^1 look at the first column of the rotation matrix Q. This column corresponds to the actual weights used to construct the volatility factor. The bias is negligible for both measures of the accuracy of the rotation. The second measure computes the angle between the true parameters and the estimates divided by π. This angle is computed as

$$\cos^{-1}\left(\frac{u'v}{\sqrt{u'u}\sqrt{v'v}}\right)/\pi$$

where u and v are the inputs. When the values coincide, this value will be 0. If the parameters were uncorrelated it would be 1. This measure indicates that both θ and the first column of the rotation matrix are accurately estimated with only a small discrepancy between the estimates. Like volatility parameters, estimating the angle parameters is easier when the volatility parameters are easier to identify. It is also more challenging to estimate these parameters when the number of mean factors k increases.

The performance in multi-factor models was also examined. The number of mean factors was fixed at $k = 3$ and the number of volatility factors was allowed to vary between 1 and 3. The parameters of the three factors were $(\alpha_i, \beta_i) \in \{(.1, .87), (.15, .73), (.2, .6)\}$ where the order was preserved so that the first factor was always the most persistent. Results from this simulation study are presented in figure 3. The top two panels show the bias and RMSE of the parameters. Increasing the model's complexity by introducing additional factors does not affect the accuracy of the estimation of other factors. If anything, larger models are easier to estimate. This happens since additional volatility factors make it simpler to identify the rotation required to orthogonalize the residuals when constructing the shocks to the volatility factors. The final panel contains measures of the accuracy of the rotated residuals which serve as shocks to the factors and the fitted variances. The columns labeled ρ measure the correlation between the estimated series and the true series. These are all very high indicating that the model is accurate. The columns labeled σ measure the standard deviation between these. Both of these measures indicate that larger models are easier to identify than smaller models. Again, this is due to the increased accuracy of estimating the rotation.

6 Data

All data are from the FRED-MD (McCracken and Ng 2016) using the August 2017 vintage. The FRED-MD is freely available from the St. Louis Federal Reserve Bank and is regularly updated. This database is an alternative to the database commonly used in factor modeling based on the Global Insights Basic Economics Database (GSI) (Stock and Watson 2005; Stock and Watson 2006). Factors extracted from FRED-MD have predictive power as those extracted from the GSI database.

The database contains 128 series from January 1959 until August 2017. The components in the FRED-MD span 8 broad categories where the number if parentheses reports the number of components in each: output and income (17), labor market (32), housing (10), consumption (14), orders and inventories (14), money and credit (15), interest and exchange rates (22), prices (21) and stock market (4).

Data were transformed according to the recommended transformation. When consider-

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 2	-0.006				
Factor 3	0.518	0.409			
Factor 4	0.176	0.107	-0.104		
Factor 5	0.515	-0.024	0.359	-0.373	
Factor 6	-0.251	0.110	-0.269	-0.053	-0.159

Table 1: Correlations between factor residuals when factor normalization requires factors to be orthogonal.

ing uncertainty timing, this has some impact since most variables are differenced or log-differenced. I align all variables in the month in which they were observable. All variables were further transformed so that they have zero mean and unit variance. Factors were extracted using principal component analysis. Variables with missing values were included in any sample where at least 120 monthly observations were not missing. Principal components were extracted using the E-M algorithm which iterates between replacing missing values by their expected value given the current value of the estimated factors and updating the factors using the augmented data. This procedure is monotonic in reducing the sum of squared residuals for a fixed number of factors and continues until the change in the factors is negligible. The E-M algorithm was initialized using the sample mean of the non-missing values for each component. While the factors extracted by PCA are normalized to be orthogonal, the main input into the analysis is the residuals from the forecasting model for the dynamic factors. These value of the correlation for the first six factors, extracted from a VAR(1) on these factors, is presented in table 1.

Many of these correlations are far from zero. These are standardized so that they are uncorrelated with unit variance by multiplying the factors (and so their residuals) by the symmetric square root of the inverse of the sample covariance matrix.

7 Results

All models were estimated on the full sample of 128 series. Models ranging from 3 to 6 mean factors were estimated using MLE where the number of volatility factors was allowed to vary between 1 and the number of mean factors. Volatility parameter estimated for a subset of the models estimated are presented in table 4. The first three rows include 3 mean factors and vary the number of volatility factors from 1 to 3. The parameters for the factors do not meaningfully as the number of volatility factors is increased. The first factor has a total persistence of $\alpha_1 + \beta_1 \approx 0.97$ with a large innovation parameter ($\alpha_1 \approx 0.18$). While the persistence is similar to what is found in financial data, the innovation sensitivity is higher. The second factor, only included when $m = 2$ or 3, is also highly persistent with a larger innovation parameters of $\alpha_2 =$0.25. The third factor, which only was included in one specification, is meaningfully less persistent with $\alpha_3 + \beta_3 = 0.87$. Inference on the parameter indicates all are strongly identified with very small standard errors. Increasing the number of volatility factors does not have an effect on the precision of the more important factors.

The bottom three rows of table 4 contain volatility parameters for a fixed number of volatility factors, $m = 2$, varying the number of mean factors from 4 to 6. These results show some important regularity – larger model have volatility factors that more persistent with smaller innovation parameters (α_1). The persistence rises of 0.99 when 6 mean factors are included. The volatility parameters of the second factor do not change as the number of factors changes, and all are highly statistically significant.

One concern when estimating factor models with different orders is that the estimated factors are not comparable across parameterizations. The model is parameterized using angles which are converted to orthonormal weighting matrices where column j contains the weights for volatility factor j. Table 5 contains weight parameters for two factors from models with k between 3 and 6 and $m = 2$. This shows that the first factor changes dramatically when the number of factors included in the model rises above 4. When the model has $k = 3$, the first factor has large weights on the residuals from the first and third component. This weighting will have positive weights on real variables and negative weights on rates. For larger models where $k \geq 4$, the first component is dominated by the fourth factor, although it also has non-negligible weights on the first and third components. The fourth component is especially sensitive to interest rates. When $k = 5$ or 6, the first factor also receives non-negligible weight. The second factors in all models are broadly comparable, with a large weight on the second factor. These results indicate that, at least from the point-of-view of modeling the conditional variance of the factors, models with $k = 3$ factors may be missing important components of uncertainty.

An alternative method to examine weight stability is to hold the number of mean factors, k, constant while varying the number of volatility factors. Table 6 contains three panels which examine stability in this dimension. The top panel contains the weights on the first factor as the number of volatility factors increases from 1 to 3. These parameters are virtually identical indicating that the factor identification, conditional on the number of mean factors, is stable. The bottom two panel contain similar results for a model with 6 mean factors. Both of these panels indicate that the choice of volatility factor is not sensitive to the number of included volatility factors.

Figure 1 contains plots for the estimated factor volatility for k between 3 and 6 when the models included 2 volatility factors. The fitted values for $k = 3$ are considerably different from the fitted values when $k = 4$, 5 or 6. The first factor, in particular, is considerably different. In the larger models, this factor has a very large, pronounced volatility spike in the early 1980s. It also has smaller spikes in the mid 1970s as well as the early 2000s and during the financial crisis. When $k = 3$ the first factor also indicates considerable volatility in the early 1980s and during the financial crisis, but differs in the early 2000s as well as in the early years where this component indicates considerable volatility that does not appear in the models with $k \geq 4$. The second component is virtually identical across the four specifications. This isn't surprising since both the estimated volatility parameters and the weights were very similar.

While factor volatilities can be interesting when if the factors can be clearly identified, a more natural measure of the volatility is the volatility that appears in the series used to construct the factors. The innovations of these series depend on the factor innovations as well as an idiosyncratic shock. The dependence on the factor innovations is measures by $\hat{\upsilon}_i$ which is computed from a regression of the idiosyncratic shock on the factor shocks. Using this decomposition, the variance attributable to the common factors for series i is

$$\upsilon_i' \Sigma_t \upsilon_i,$$

and the average total volatility is then

$$\bar{\sigma}_{f,t}^2 = N^{-1} \sum_{i=1}^{N} \hat{\upsilon}_i' \hat{\Sigma}_t \hat{\upsilon}_i.$$

Figure 2 contains four panels plotting $\sqrt{\bar{\sigma}_{f,t}^2}$ corresponding to $k = 3$, 4, 5 and 6. In each panel, the dynamics for specification with m between 1 and k are plotted, although $m = 2$ is highlighted. In all specification setting $m = 1$ produced total volatility series which nearly constant aside from the early 1980s. Expanding this to two factors produced total volatility series that show spikes for a number of events. When $k \geq 4$ the fitted total volatility series are virtually identical. Increasing the number of volatility factors changes the fit by improving the fit during low volatility periods which not affecting the fit during spikes. The notable exception to this is the spike in volatility in the early 1960s which is not evident unless $m > 2$ when $k \geq 4$. A closely related measure is the share of total volatility that can be attributed to common factors. This is computed as

$$\bar{s}_{f,t}^2 = N^{-1} \sum_{i=1}^{N} \frac{\hat{\upsilon}_i' \hat{\Sigma}_t \hat{\upsilon}_i}{\tau_i^2 + \hat{\upsilon}_i' \hat{\Sigma}_t \hat{\upsilon}_i}$$

and accounts for the different constant in each of the individual series variances. Figure 5 shows the plot of this series for $m = 2$ where k ranges from 3 to 6. The model with $k = 3$ is again different from the larger models, which are all nearly indistinguishable from each other. There are three very large spikes in these series – mid-1970s, early 1980s and the financial crisis.

The model was also estimated using composite likelihood (CMLE). Estimating model parameters using this method allows the model to utilize both the factor residuals as well as the individual residuals when estimating the dynamics of factor variances. Results from fitting the model using CMLE are presented in tables 7 and 8. The estimated parameters are similar to those from the MLE estimates with one important distinction. All estimates indicate higher persistence than the ML estimates. The parameters on the first factor across all specification had persistence very close to unity. The second factor was also highly persistent with a typical persistence of 0.98. In the one model that included the first component, it was also highly persistent with an average persistence 0.97. Sensitivities, especially for the first factor, were generally larger. The increase in sensitivity was especially pronounced for specifications based on 4 or more factors. The estimated factor loadings were broadly similar to those in the models fit using ML. The first factor differed substantially once the fourth factor was allowed to enter.

Figure 3 and 4 contain the factor and total volatility from the models estimated using composite likelihood. Like the parameter estimates, these are similar to the corresponding figures for the models estimated using ML. The fitted factor volatilities or total volatilities are virtually indistinguishable when k is larger than 4.

While labeling factors is difficult, figure 6 contains a heat map of factor loadings in both the original PCA-extracted mean factors and in the final rotated volatility factors from a model with four mean and four volatility factors $(k = m = 4)$. There are large differences in the important blocks in the two figures. In the mean factors the first and second components are broadly real and nominal factors. The third has positive weights on a number of interest and exchange rates and negative loadings on real factors. The fourth has positive loadings on interest rates and negative loadings on housing-related measures. When re-rotated through the volatility model the importance of the components shifts. The first volatility factor has large positive loadings on interest rates and relatively small loadings on real components. The second factor has positive exposure to real components. Both the first and the second have negative loadings to monetary policy variables and exchange rates. The third volatility factor is nearly identical to the second mean factor, and the fourth factor has positive loadings on monetary policy variables and negative loadings on housing. The differences in the loadings show that prices are less important for uncertainty measurement and that interest rates – the largest set of series that are financial in nature – play an important role.

7.1 Recursive Estimation

Volatility models in many financial time series are not stable. This has been attributed to number of factors including structural change. The stability of these models was examined by recursively estimating the models starting in 1989 after 30 years of data was available. Table 9 contains the results for models estimated using 30, 40, and 50 years as well as the full sample. The left panel contains model estimates for 3 mean factors while the right panel contains results for 6 mean factors. The number of volatility factors was fixed at 2 in both panels. The first factor is fairly stable across specifications. In the larger model there is some evidence of decline in the sensitivity to recent news (α_1) in the longer samples, although the persistence is consistent across sample sizes. The second factor appears to be less well identified in the sub-samples. Specifically the persistence is fairly low until the sub-sample that ended in 2009. This is not surprising considering the time-series dynamics of the second volatility factor which was highly variable post-2000 but considerably less so in the early sample. The larger model stabilizes a bit sooner and shows high persistence from 1999 for both volatility factors.

Figure 7 contains the fitted total total volatilities for the two models estimated using the 4 sub-samples. The highlighted portion indicate the segment estimated using the most recent data. In both plots the total variance is not meaningfully altered using parameters estimated on longer series. This is due to the model capturing the salient factor volatility in each period even though all factors were not equally well measured. Moreover, the factors are weighted by the exposure of the individual series v_i when computing total volatility and so small exposures make factors less important.

7.2 Forecasting

An alternative method to assess model performance is to examine how well the model forecasts future volatility. This forecasting exercise examines the effect of the number of volatility factors on the ability of the model to forecast the volatility of future factor covariance. Parameters were estimated using a 30-year moving window. In each window the subset of available data was constructed subject to a minimum of 120 non-missing observations for each series. The series were standardized by subtracting the mean and dividing through by the standard deviation. The standardized series were used to estimate factors and the weights used to construct the factors. The factor residuals were then estimated using a VAR(1). Finally, the volatility model was estimated on the factor residuals.

The models were evaluated using 1-step ahead data. Since the data are factor residuals, care was required to construct meaningful residuals. Each set of 360 monthly observation was extended by one observation. These were then standardized using the in-sample mean and variance used in the original factor construction. These standardized values were then multiplied by the factor factor weights to produce out-of-sample factors. Finally, the evaluation residual was constructed taking by subtracting the $t + 1$ set of factors from the period t set of factors multiplied by the in-sample VAR(1) coefficients. This produced an evaluation sample of 343 observations. The residual construction was repeated for each k between 3 and 6 since the residuals are constructed by a VAR, and so are not directly comparable. These residuals are used to proxy for the unknown conditional covariance.

Forecasts for factor residual covariance were constructed by taking the diagonal conditional covariance matrix of the rotated residual and rotating them back to the original series. These were then rotated so that the one-step ahead forecast is

$$\Omega_{t+1|t} = \Sigma_t \hat{Q}_m \hat{\hat{\Sigma}}^{1/2}.$$

Forecasts were evaluated using the QLIK loss function was has been found to provide good power in model evaluation (Patton and Sheppard 2009; Laurent, Rombouts, and Violante 2014) and to be robust to the use of a proxy for covariance. The loss was computed as

$$\mathcal{L}\left(\Omega_{t+1}, \xi^f_{t+1}\right) = \ln|\Omega_{t+1}| + \left(\xi^f_{t+1}\right)' \Omega^{-1}_{t+1} \xi^f_{t+1}$$

Two tests were performed. The first is a Diebold-Mariano-Giacomini-White test which examines the pairwise difference in loss differential, $\delta_t = \mathcal{L}\left(\Omega^A_{t+1}, \xi^f_{t+1}\right) - \mathcal{L}\left(\Omega^B_{t+1}, \xi^f_{t+1}\right)$ (Diebold and Mariano 1995; Giacomini and White 2006). The test statistic is the mean loss difference divided by a consistent estimator of the long run variance,

$$t = \frac{\bar{\delta}}{\sqrt{Var\left(\bar{\delta}\right)}}.$$

The long-run variance was estimated using the stationary bootstrap of Politis and Romano 1994 using an average window size of 18 periods. Since residuals depend on the size of the model, the models were only compared with a particular k. Results of the pairwise comparison are presented in table 10. In each panel, negative values indicate the model in the row (indicated by the value of m) was superior. When $k = 3$ the one factor model is statistically superior to either the two or the three. In he larger models the one-factor model is also superior to the 2 factor model for $k = 4$ and 6. When $k = 4$ there were no other rejections although the model with 4 factors was work than all others. In the largest two models the 4 factor model performed well and was better than all others, although the performance difference was not statistically significant. In general models with more volatility factors performed better than models with fewer volatility factors.

The model confidence set of Hansen, Lunde, and Nason 2011 was used to jointly test all specifications. While this test is likely less powerful than pairwise DMGW tests, it has the advantage of controlling family-wise error rate. The MCS was applies to all models for each k. Results of using the MCS are presented in table 11. In the smaller models the single factor model was superior while in the larger models the 4 factor models were the best. The MCS was only able to remove a small number of models in each group and there are no clear winners.

7.3 Comparing Alternative Measures of Uncertainty

One natural measure of uncertainty in a large panel is the average volatility of all of the data. This measure has been explored in Jurado, Ludvigson, and Ng 2015 as the average volatility of the forecast residuals. In the original version, this was computed using a standard one-factor log stochastic volatility model. I computed a similar measure using univariate GARCH(1,1) models where each series has independent volatility. I also consider a pooled model where all series share common parameters but have different long-run levels of volatility. In either case, the conditional for the i^{th} series is given by

$$\varsigma_{it}^2 = \omega_i + \alpha_i \left(\xi_{t-1}^i\right)^2 + \beta_i \varsigma_{it-1}^2.$$

When composite likelihood is used to estimate parameters, $\alpha_i = \alpha_j$ and $\beta_i = \beta_j$ for all i, j. This can be used to measure total volatility by averaging cross-sectionally and then taking the square root. The fitted volatility computed using this method is produced in the top panel figure 8. The two method produce similar fits although the composite method appears less noisy. The dynamics of this series appear to be very similar to those of the total volatility that can be attributed to common factors. The average parameters values when estimating the model while allowing for heterogeneity was $\bar{\alpha} = 0.283$, $\bar{\beta} = 0.599$. When the parameters were pooled, the estimates were considerably larger $\hat{\alpha} = 0.226$, $\hat{\beta} = 0.709$. Both of these estimates indicate less persistence that was seen in the factor volatility. This difference is the root of the discrepancies between total factor volatility and total volatility. The individual models are conflating dynamics of factors and shocks.

The same model was estimated on forecast residual shocks purged of the effect of the common factors, defined as $\eta_t^i = \xi_t^i - \upsilon_i \xi_f^f$. The bottom panel of figure 8 shows that average volatility of these series computed using the same methods. The volatility spikes in these series are substantially mitigated. This difference is not due to a reduction in persistence since $\bar{\alpha} = 0.228$ and $\bar{\beta} = 0.636$ when individual models were fit and $\hat{\alpha} = 0.180$, $\hat{\beta} = 0.727$ when composite likelihood was used. These parameters indicate more persistence in the factor-free shocks than in the total shocks. Instead the difference is due to less commonality in the factor-free volatility series which produces cross-sectional averages which has less pronounced dynamics.

Figure 9 plots the three series together. The overall dynamics of the common factor volatility and the idiosyncratic volatility are very similar. However, it can be clearly seen that the volatility of the idiosyncratic residual contaminates the common factor volatility and makes large changes in uncertainty more difficult to detect. The volatility of the idiosyncratic residual is nearly flat with two important exceptions – that early period in the 1960s and the financial crisis. Both of these also appear in the common factor volatility and are much more pronounced.

8 Conclusions

I propose a new method to measure and decompose macroeconomic uncertainty. Factors extracted from macroeconomic panels are typically normalized using a convenient but ultimately arbitrary normalization. This motivates searching for an alternative normalization that emits volatility factors – rotations of the original factors – that have conditional covariance that are uncorrelated and have simple volatility dynamics. The model allows the dimension of the heteroskedasticity to be smaller than number of factors in the series. This is an important consideration when modeling macroeconomic data which has often been hypothesized to have homoskedastic residuals. Conditional on a specific rotation, the model has simple dynamics where each volatility factor has dynamics that follow a standard GARCH model. The permits simple likelihood-based estimation. I also propose a composite likelihood-based estimator of the model parameters that utilizes information in the full panel of macroeconomics variables.

Monte Carlo evidence demonstrates that the model parameters can be accurately recovered from empirically relevant sample sizes. They also show that many aspects of the estimation are easier when the heteroskedasticity is pervasive rather than limited to a small number of volatility factors. This occurs since the additional heteroskedasticity improves the identification of the rotation of the original factors to produce the volatility factors.

The model is applied to a panel of 128 monthly macroeconomic series from the FRED-MD. The model suggests macroeconomic volatility is both pervasive and highly persistent. Composite likelihood estimation suggests the macroeconomic volatility components are more persistence than the maximum likelihood estimates, although both show high persistence. Construction of total common factor volatility shows that between 2 and 4 volatility components capture the lion share of the dynamics in the series. This is confirmed by comparing the total volatility in the series with idiosyncratic-only volatility where the latter is relatively homoskedastic. The proposed measure of common factor volatility provides much sharper identification of large changes in macroeconomic uncertainty than models using the full forecast residual from the individual series.

References

Andrews, Donald W K (2001). "Testing when a parameter is on the boundary of the maintained hypothesis". *Econometrica* 69.3, pp. 683–734.

Bai, Jushan and Serena Ng (2002). "Determining the number of factors in approximate factor models". *Econometrica* 70.1, pp. 191–221.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016). "Measuring economic policy uncertainty". *The Quarterly Journal of Economics* 131.4, pp. 1593–1636.

Bijsterbosch, Martin and Pierre Guérin (2014). *Characterizing Very High Uncertainty Episodes*. Tech. rep. European Central Bank.

Comte, F. and O. Lieberman (2003). "Asymptotic theory for multivariate GARCH processes". *Journal of Multivariate Analysis* 84.1, pp. 61–84. ISSN: 0047-259X.

Diebold, Francis X and Roberto S Mariano (July 1995). "Comparing Predictive Accuracy". *Journal of Business & Economic Statistics* 13.3, pp. 253–263.

Engle, Robert F and Kenneth F Kroner (1995). "Multivariate Simultaneous Generalized {ARCH}". *Econometric Theory* 11.1, pp. 122–150.

Giacomini, Raffaella and Halbert White (Nov. 2006). "Tests of Conditional Predictive Ability". *Econometrica* 74.6, pp. 1545–1578.

Hansen, Peter R, Asger Lunde, and James M Nason (2011). "The model confidence set". *Econometrica* 79.2, pp. 453–497.

Jo, Soojin and Rodrigo Sekkel (2017). "Macroeconomic Uncertainty Through the Lens of Professional Forecasters". *Journal of Business & Economic Statistics*, pp. 0–0. ISSN: 0735-0015.

Jurado, Kyle, Sydney Ludvigson, and Serena Ng (2015). “Measuring Uncertainty”. *American Economic Review* 2.1, pp. 2–7. ISSN: 1467-9639.

Laurent, Sébastien, Jeroen V K Rombouts, and Francesco Violante (2014). “Consistent Ranking of Multivariate Volatility Models”. *Journal of Econometrics*.

McCracken, Michael W. and Serena Ng (2016). “FRED-MD: A Monthly Database for Macroeconomic Research”. *Journal of Business and Economic Statistics* 34.4, pp. 574–589. ISSN: 1537-2707.

Newey, Whitney K and Daniel McFadden (1994). “Large Sample estimation and Hypothesis Testing”. *Handbook of Econometrics*. Vol. 4. Elsevier North Holland.

Newey, Whitney K and Kenneth D West (1987). “A Simple, Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. *Econometrica* 55.3, pp. 703–708.

Pakel, Cavit et al. (Sept. 12, 2017). *Fitting Vast Dimensional Time-Varying Covariance Models*. Tech. rep. NYU Working Paper No. FIN-08-009.

Patton, Andrew and Kevin Sheppard (Dec. 2009). “Evaluating Volatility Forecasts”. *Handbook of Financial Time Series*. Ed. by Torben Gustav Andersen et al. Springer, p. 750.

Politis, D N and J P Romano (1994). “The stationary bootstrap”. *Journal of the American Statistical Association* 89.428, pp. 1303–1313.

Rossi, Barbara and Tatevik Sekhposyan (2015). “Macroeconomic uncertainty indices based on nowcast and forecast error distributions”. *The American Economic Review* 105.5, pp. 650–655.

Rossi, Barbara, Tatevik Sekhposyan, and Matthieu Soupre (2016). *Understanding the sources of macroeconomic uncertainty*. Tech. rep. Universitat Pompeu Fabra.

Stock, James H. and Mark W. Watson (2002). “Macroeconomic Forecasting Using Diffusion Indexes”. *Journal of Business and Economic Statistics* 20.2, pp. 147–162.

Stock, James H. and Mark W. Watson (July 2005). *Implications of Dynamic Factor Models for VAR Analysis*. Working Paper 11467. National Bureau of Economic Research.

Stock, James H. and Mark W. Watson (Jan. 1, 2006). “Forecasting with many predictors”. *Handbook of Economic Forecasting*. Vol. 1. North Holland, pp. 515–554.

Van der Weide, Roy (2002). “GO-GARCH: a multivariate generalized orthogonal GARCH model”. *Journal of Applied Econometrics* 17.5, pp. 549–564.

Single Factor Monte Carlo Results

	α		β		θ		Q^1	
k	Bias	RMSE	Bias	RMSE	Bias	Angle	Bias	Angle
				$\alpha = 0.1, \beta = 0.8$				
3	0.010	0.047	-0.035	0.149	-0.001	0.039	-0.001	0.060
4	0.012	0.051	-0.038	0.164	0.010	0.059	0.000	0.076
5	0.014	0.054	-0.033	0.162	0.013	0.075	0.001	0.092
6	0.023	0.058	-0.057	0.187	-0.014	0.089	-0.001	0.105
				$\alpha = 0.1, \beta = 0.85$				
3	0.004	0.039	-0.017	0.090	-0.005	0.027	0.001	0.041
4	0.008	0.045	-0.024	0.110	0.005	0.042	0.003	0.056
5	0.008	0.042	-0.020	0.097	0.001	0.061	-0.001	0.062
6	0.012	0.048	-0.031	0.116	0.002	0.079	-0.001	0.077
				$\alpha = 0.15, \beta = 0.84$				
3	-0.000	0.037	-0.002	0.041	0.001	0.011	0.001	0.014
4	0.002	0.037	-0.004	0.042	0.005	0.022	0.000	0.018
5	0.004	0.038	-0.006	0.043	-0.006	0.037	0.000	0.021
6	0.002	0.034	-0.004	0.039	-0.002	0.045	0.001	0.024

Table 2: Results from one volatility factor ($m = 1$), 3 mean factor ($k = 3$) models. The three panels differ in the specification of volatility model dynamics. In each panel, the left four columns report bias and RMSE of the estimated volatility parameters. The columns corresponding to θ report the bias in the estimated rotation angles and the average angle divided by π between the estimates of $\hat{\theta}$ and the true values. The final two columns report the bias and average angle divided by π between the first column of the rotation matrix, the weights on the volatility factor, and the first column of the true rotation matrix.

Multi-Factor Monte Carlo Results

Bias

m	α_1	β_1	α_2	β_2	α_3	β_3
1	-0.000	-0.006				
2	-0.001	-0.004	0.006	-0.020		
3	-0.003	-0.000	0.005	0.002	0.005	-0.031

RMSE

m	α_1	β_1	α_2	β_2	α_3	β_3
1	0.035	0.065				
2	0.034	0.051	0.049	0.115		
3	0.034	0.049	0.048	0.087	0.057	0.139

Correlations

m	$\rho_{\xi f}$	$\sigma_{\xi f}$	ρ_{σ^2}	σ_{σ^2}
1	0.990	0.056	0.966	0.072
2	0.992	0.014	0.965	0.071
3	0.995	0.007	0.971	0.048

Table 3: Results from Monte Carlo simulations for $k = 3$ varying m from 1 to 3. The top panel reports the bias of the estimated volatility parameters and the middle panel reports the RMSE of the estimated volatility parameters. The left two columns of the bottom panel reports correlations between the estimated and true volatility factor weights and the variance of the factor weights. The right two columns report the correlation between the true factor variances and the estimated factor variances, and the standard deviation between these two.

Volatility Model Parameters

k	m	α_1	β_1	α_2	β_2	α_3	β_3
3	1	0.183	0.789				
		(0.052)	(0.054)				
3	2	0.181	0.791	0.256	0.708		
		(0.053)	(0.054)	(0.047)	(0.057)		
3	3	0.181	0.791	0.256	0.708	0.187	0.688
		(0.051)	(0.057)	(0.047)	(0.057)	(0.103)	(0.216)
4	2	0.193	0.782	0.254	0.713		
		(0.048)	(0.058)	(0.047)	(0.057)		
5	2	0.169	0.815	0.253	0.716		
		(0.052)	(0.061)	(0.046)	(0.055)		
6	2	0.150	0.840	0.261	0.706		
		(0.035)	(0.040)	(0.047)	(0.056)		

Table 4: Volatility parameter estimated using MLE. The first three rows report the parameters for models based on 3 mean factors ($k = 3$), varying the number of volatility factors between 1 and 3. The final 3 lines report the model parameters holding the number volatility factors (m) constant while varying the number of mean factors from 4 to 6.

Volatility Model Factor Weights

First Factor

k	w_1	w_2	w_3	w_4	w_5	w_6
3	0.800	-0.004	-0.600			
4	0.234	0.064	-0.294	0.924		
5	0.297	0.012	-0.214	0.887	-0.280	
6	0.303	0.083	-0.224	0.865	-0.276	-0.167

Second Factor

k	w_1	w_2	w_3	w_4	w_5	w_6
3	0.018	1.000	0.017			
4	0.024	0.981	0.191	-0.013		
5	0.047	0.983	0.177	0.006	-0.023	
6	0.043	0.931	0.355	-0.015	-0.043	0.055

Table 5: Estimated volatility factor weights on the first and second volatility factors for models based on between 3 and 6 mean factors (k). All models estimated included 2 volatility factors.

Comparison of Volatility Model Factor Weights

$k = 3$, First Factor

m	w_1	w_2	w_3
1	0.800	0.020	-0.600
2	0.800	-0.004	-0.600
3	0.798	-0.004	-0.603

$k = 6$, First Factor

m	w_1	w_2	w_3	w_4	w_5	w_6
2	0.303	0.083	-0.224	0.865	-0.276	-0.167
3	0.307	0.049	-0.162	0.863	-0.333	-0.145
4	0.290	0.048	-0.161	0.870	-0.329	-0.149
5	0.292	0.045	-0.151	0.870	-0.333	-0.145
6	0.288	0.051	-0.164	0.874	-0.321	-0.142

$k = 6$, Second Factor

m	w_1	w_2	w_3	w_4	w_5	w_6
2	0.043	0.931	0.355	-0.015	-0.043	0.055
3	0.040	0.943	0.328	-0.009	-0.026	0.043
4	0.026	0.941	0.335	-0.002	-0.026	0.040
5	0.031	0.937	0.344	-0.002	-0.025	0.039
6	0.024	0.939	0.338	0.009	-0.009	0.066

Table 6: Estimated volatility factor weights.The top panel shows the weights in the first volatility factor for a model estimated with 1, 2 or 3 volatility factors. The bottom two panels show the weights used to construct the first and second volatility factors for models that included between 2 and 6 volatility factors.

Composite Likelihood Volatility Model Parameters

k	m	α_1	β_1	α_2	β_2	α_3	β_3
3	1	0.199	0.801				
3	2	0.197	0.803	0.292	0.698		
3	3	0.209	0.791	0.290	0.699	0.263	0.715
4	2	0.247	0.753	0.287	0.703		
5	2	0.227	0.771	0.289	0.703		
6	2	0.204	0.794	0.294	0.698		

Table 7: Volatility model parameter estimates from model which use Composite Likelihood to estimate model parameters.

Composite Likelihood Volatility Model Weights

			First Factor			
k	w_1	w_2	w_3	w_4	w_5	w_6
3	0.809	0.000	-0.587			
4	0.424	0.125	-0.391	0.808		
5	0.345	0.054	-0.233	0.890	-0.180	
6	0.342	0.084	-0.137	0.871	-0.290	-0.118

			Second Factor			
k	w_1	w_2	w_3	w_4	w_5	w_6
3	0.041	0.998	0.057			
4	0.041	0.974	0.209	-0.071		
5	0.074	0.975	0.198	-0.046	-0.049	
6	0.075	0.913	0.388	-0.080	-0.062	-0.026

Table 8: Volatility model parameter estimates from model which use Composite Likelihood to estimate model parameters. Estimated volatility factor weights on the first and second volatility factors for models based on between 3 and 6 mean factors (k). All models estimated included 2 volatility factors.

Recursively Estimated Volatility Model Parameters

	$k = 3$				$k = 6$			
	α_1	β_1	α_2	β_2	α_1	β_1	α_2	β_2
1989	0.173	0.807	0.275	0.493	0.206	0.785	0.415	0.473
	(0.035)	(0.090)	(0.059)	(0.034)	(0.039)	(0.031)	(0.053)	(0.030)
1999	0.128	0.853	0.263	0.540	0.161	0.828	0.234	0.718
	(0.035)	(0.077)	(0.057)	(0.033)	(0.035)	(0.025)	(0.049)	(0.027)
2009	0.199	0.767	0.201	0.760	0.156	0.833	0.209	0.763
	(0.026)	(0.118)	(0.051)	(0.031)	(0.044)	(0.024)	(0.024)	(0.032)
2017	0.181	0.791	0.256	0.708	0.150	0.840	0.261	0.706
	(0.049)	(0.030)	(0.026)	(0.115)	(0.042)	(0.027)	(0.025)	(0.034)

Table 9: Parameter estimates using a recursive estimation scheme. Parameters were estimated using 30 (ending in 1989), 40 (1999), 50 (2009) years of data as well as the full sample. The left panel shows parameter estimates for $k = 3$, $m = 2$ model. The right panel shows results from $k = 6$, $m = 2$.

Pairwise Comparison of Forecasting Models

$k = 3$

	2	3
1	$-5.54^{\star}$	$-4.15^{\star}$
2		-0.59

$k = 4$

	2	3	4
1	$-2.03^{\star}$	-1.29	-1.51
2		0.55	-0.07
3			-1.32

$k = 5$

	2	3	4	5
1	0.74	0.91	$2.15^{\star}$	0.71
2		-0.04	0.82	-0.14
3			$2.78^{\star}$	-0.52
4				$-3.50^{\star}$

$k = 6$

	2	3	4	5	6
1	$-1.99^{\star}$	-1.07	1.03	0.91	-0.53
2		1.39	1.63	$2.56^{\star}$	$3.21^{\star}$
3			1.18	1.35	1.05
4				-0.94	-1.07
5					-1.20

Table 10: Pairwise Diebold-Mariano-Giacomini-White test statistics. Negative values indicate the model with the number of volatility factors indicated on the row (left) was superior to the model indicated on the column (top). Positive number indicate the that the column model is superior to the row model.

Best Forecasting Model and Models in MCS

k	Best m	MCS Excluded m
3	1	2,3
4	1	
5	4	3,5
6	4	2

Table 11: The left columns identified the best number of volatility factors, m, for each model dimensions k. The right columns indicates the models which were excluded from a Model Confidence Set using a size of 5%.

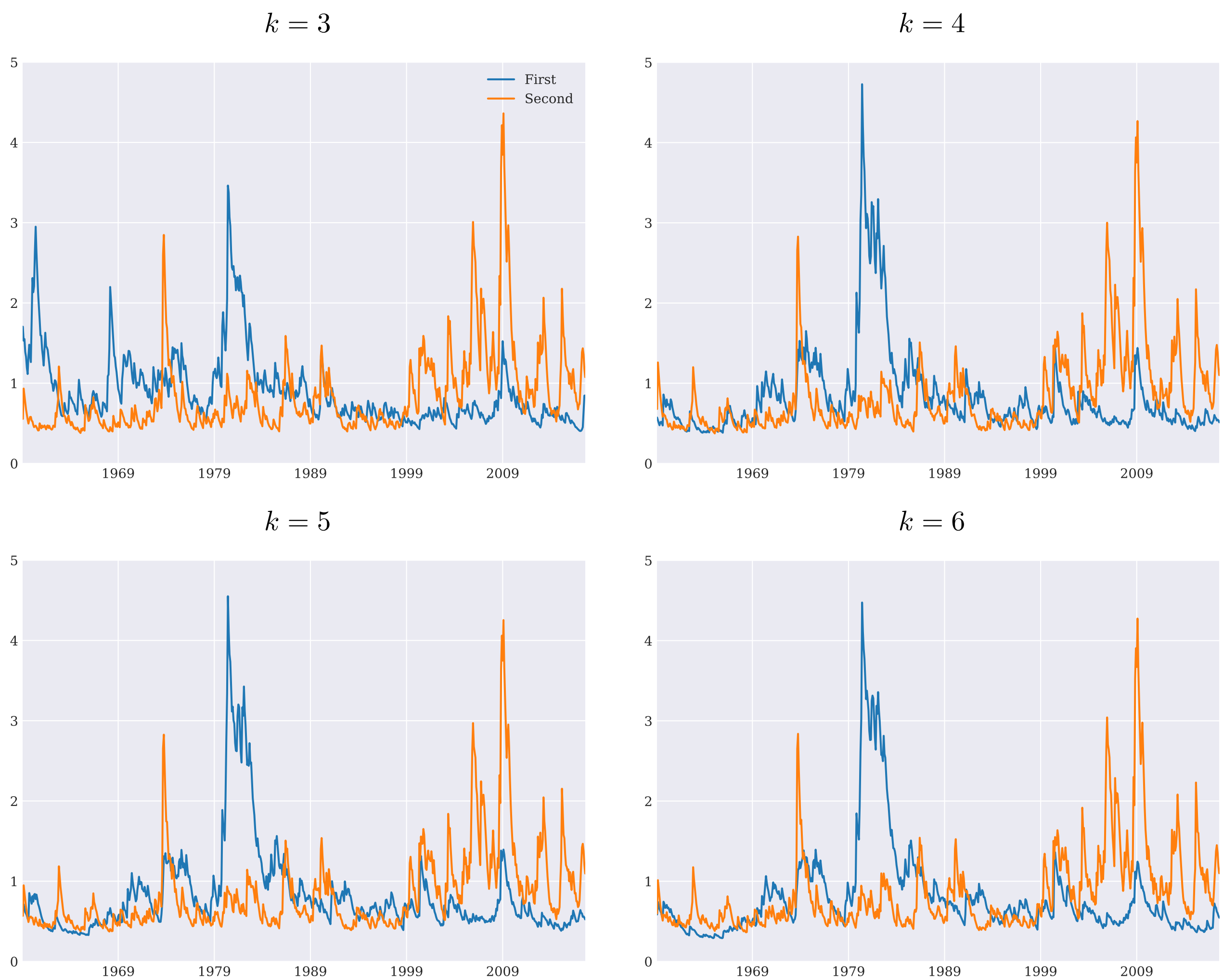

Figure 1: Estimated factor volatilities for the first and second volatility factor for models with mean factors varying from 3 to 6.

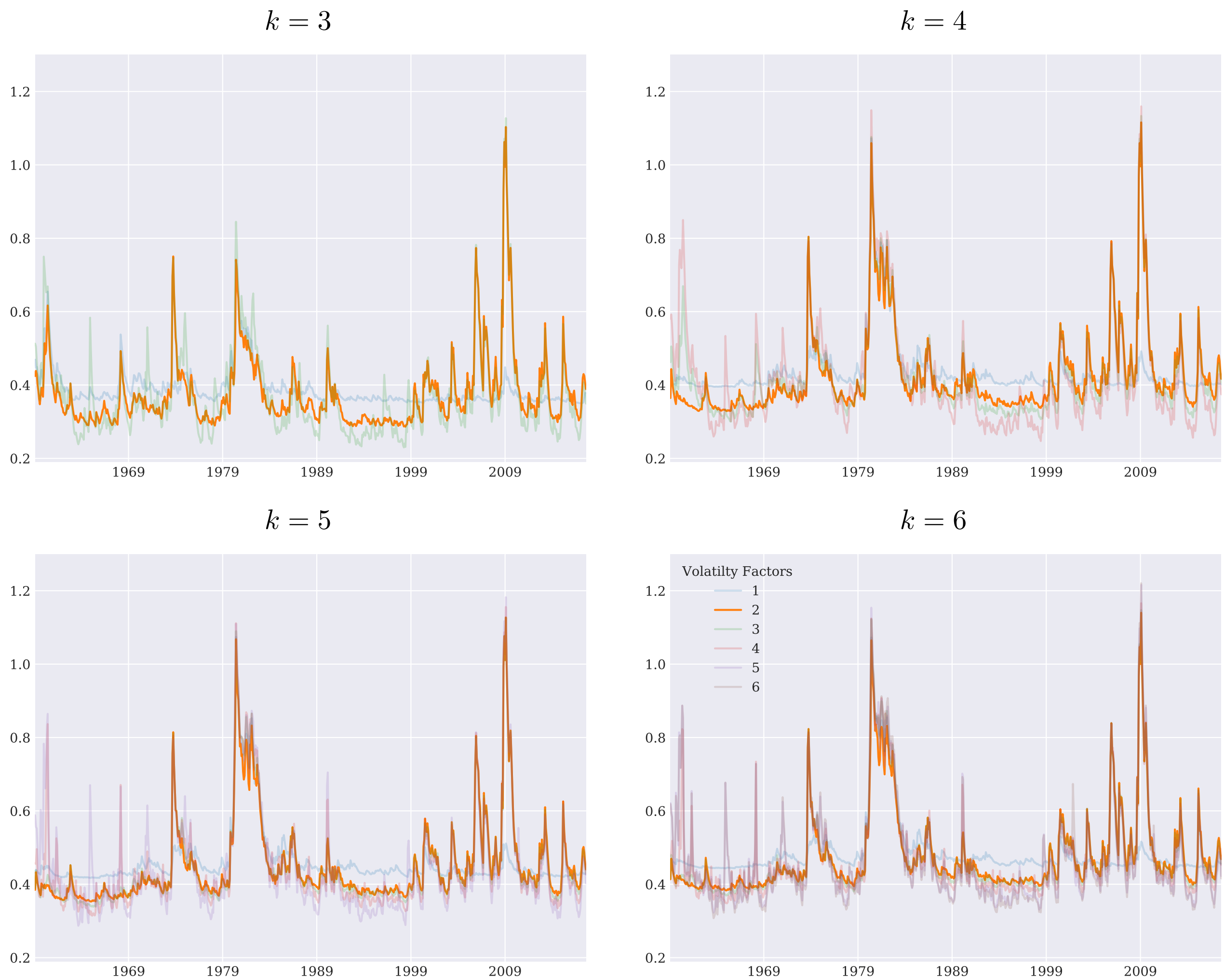

Figure 2: Average total volatility attributable to common factor exposure. Models included mean factors (k) ranging from 3 to 6. In each panel the total volatility is plotted for the number volatility factors (m) ranging between 1 and the number of mean factors in the model. In each panel, the 2-volatilty factor model is highlighted.

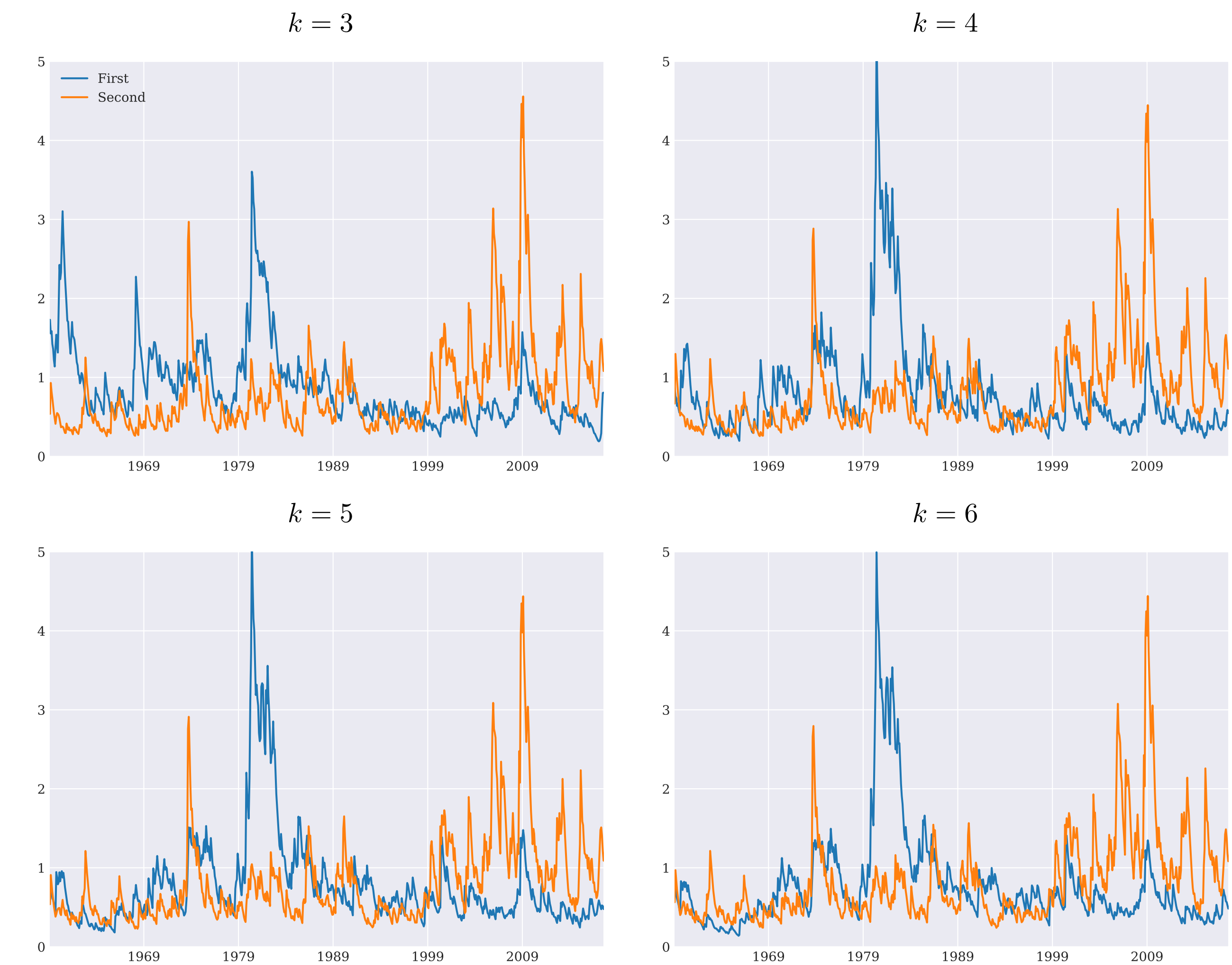

Figure 3: Estimated factor volatilities using composite likelihood for the first and second volatility factor for models with mean factors varying from 3 to 6.

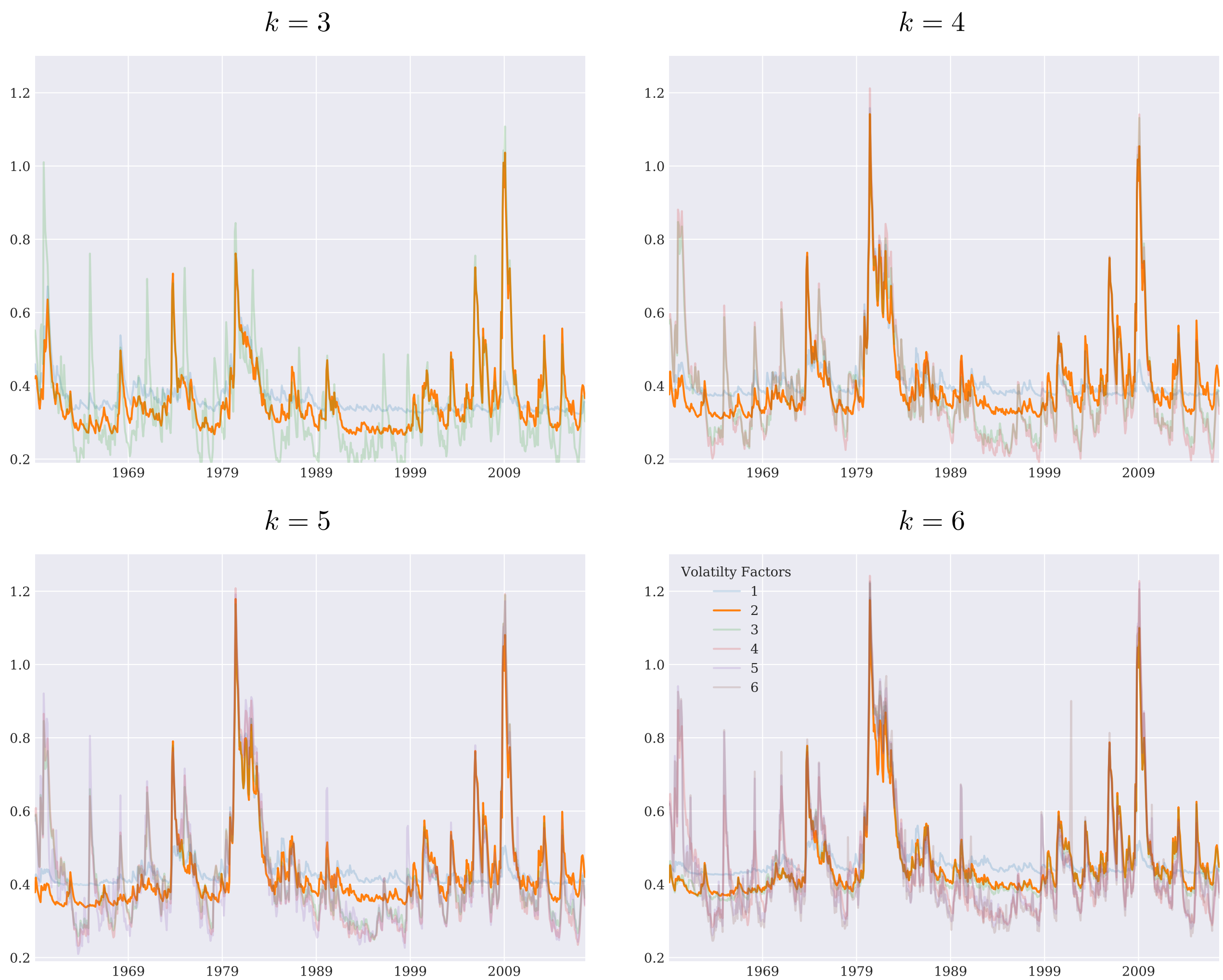

Figure 4: Average total volatility attributable to common factor exposure using models estimated by composite maximum likelihood. Models included mean factors (k) ranging from 3 to 6. In each panel the total volatility is plotted for the number volatility factors (m) ranging between 1 and the number of mean factors in the model. In each panel, the 2-volatilty factor model is highlighted.

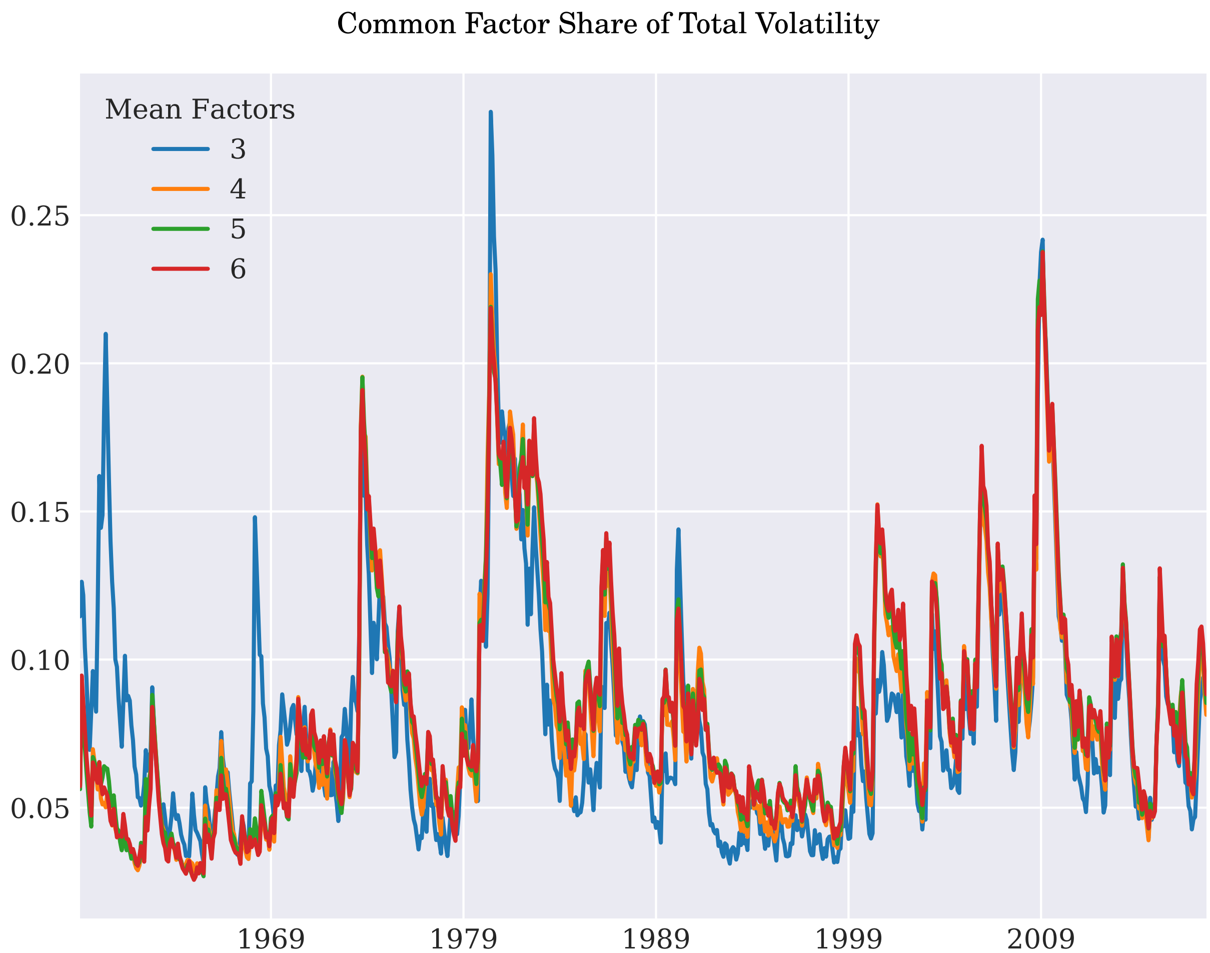

Figure 5: Share of total volatility attributable to common factor volatility. Each model uses 2 volatility factors ($m = 2$). The number of mean factors (k) was varied between 3 and 6.

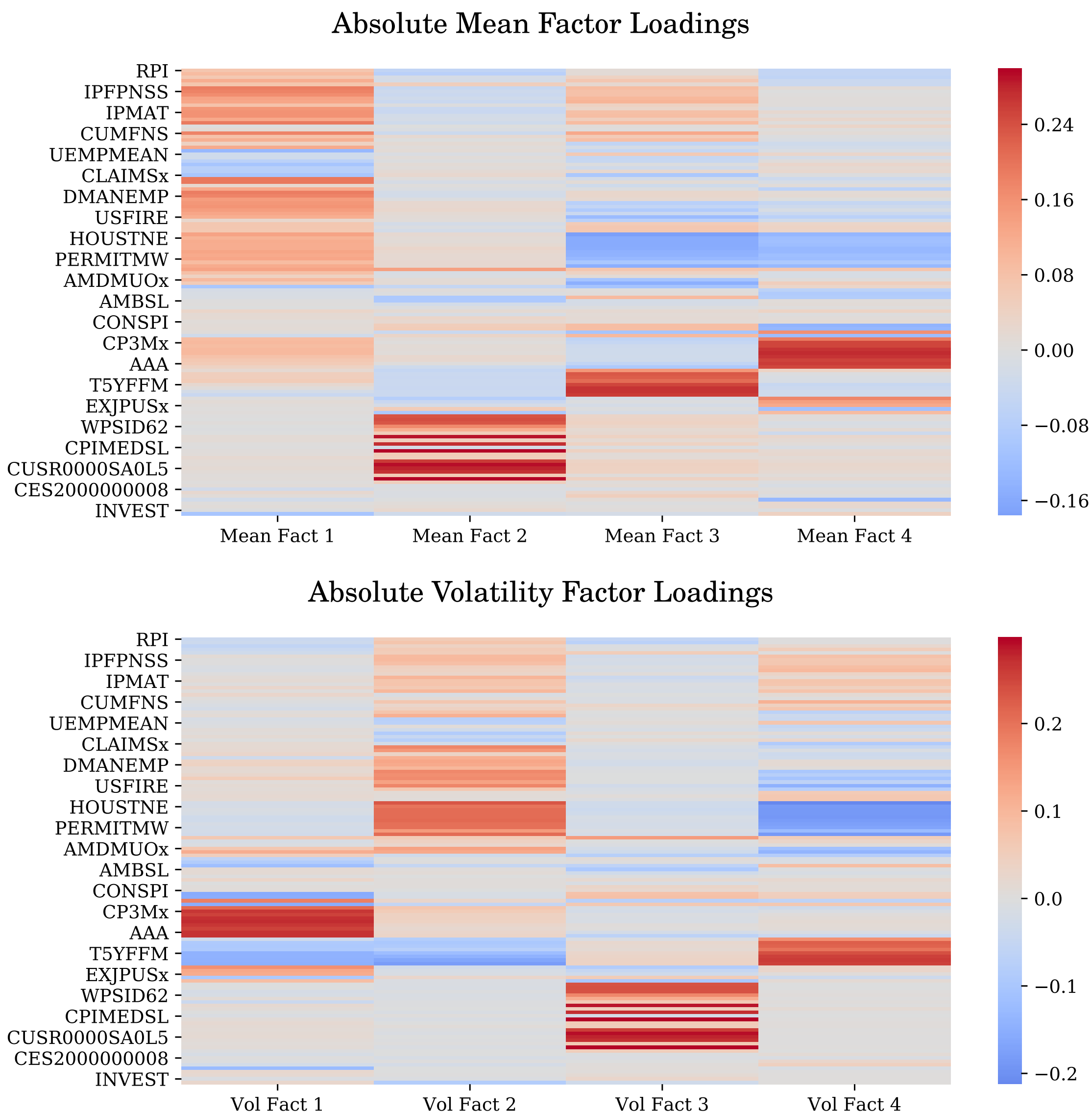

Figure 6: Factor loadings of the mean factors, as extracted using PCA on the full data set, and the volatility factor loadings, which are weighted averages of the original factors.

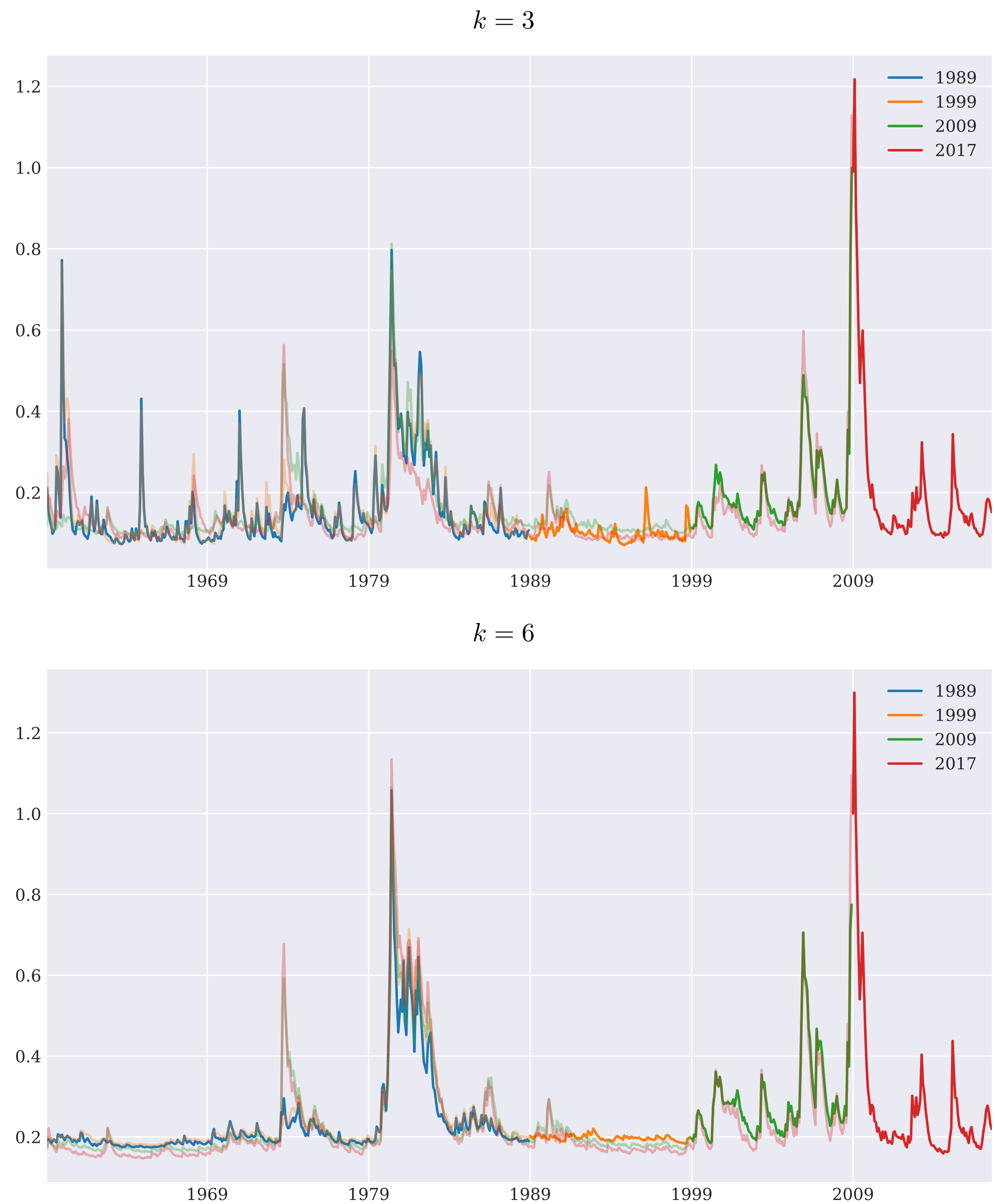

Figure 7: Recursively estimated volatilities. The top panel uses 3 mean factors (k) while the bottom panel uses 6. All models included 2 volatility factors ($m = 2$). Models were estimated using 30, 40 and 50-years of data as well as the full sample.

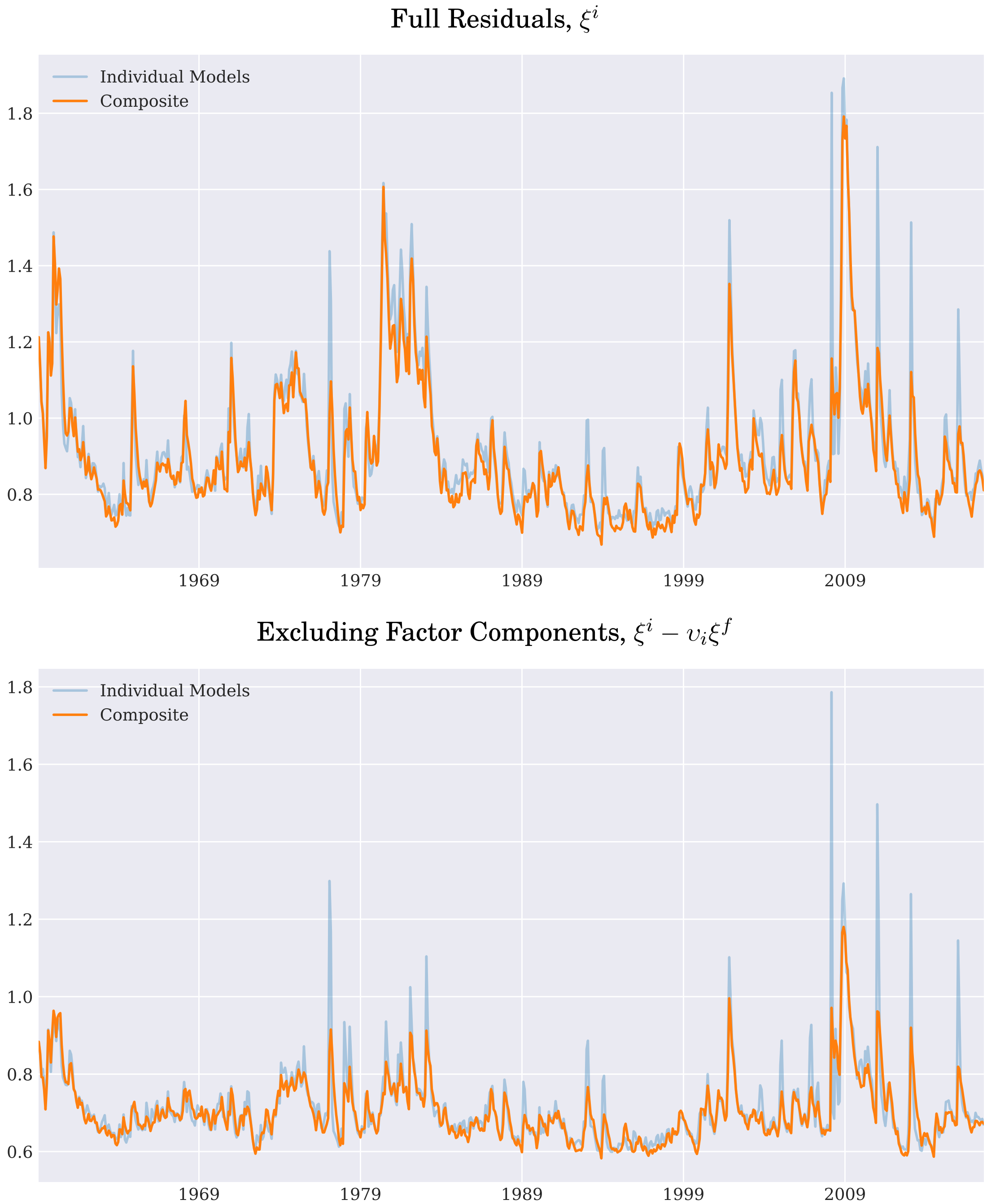

Figure 8: Average volatility when estimated using individual model shocks. The top penal uses the entire shock while the bottom uses only the idiosyncratic component assuming loading are constant. Model parameters were estimated in two way. The first fit n individual models to the residuals. The second used composite likelihood to estimate a single set of common dynamic parameters across all models.

Dynamics of Alternative Measures of Uncertainty

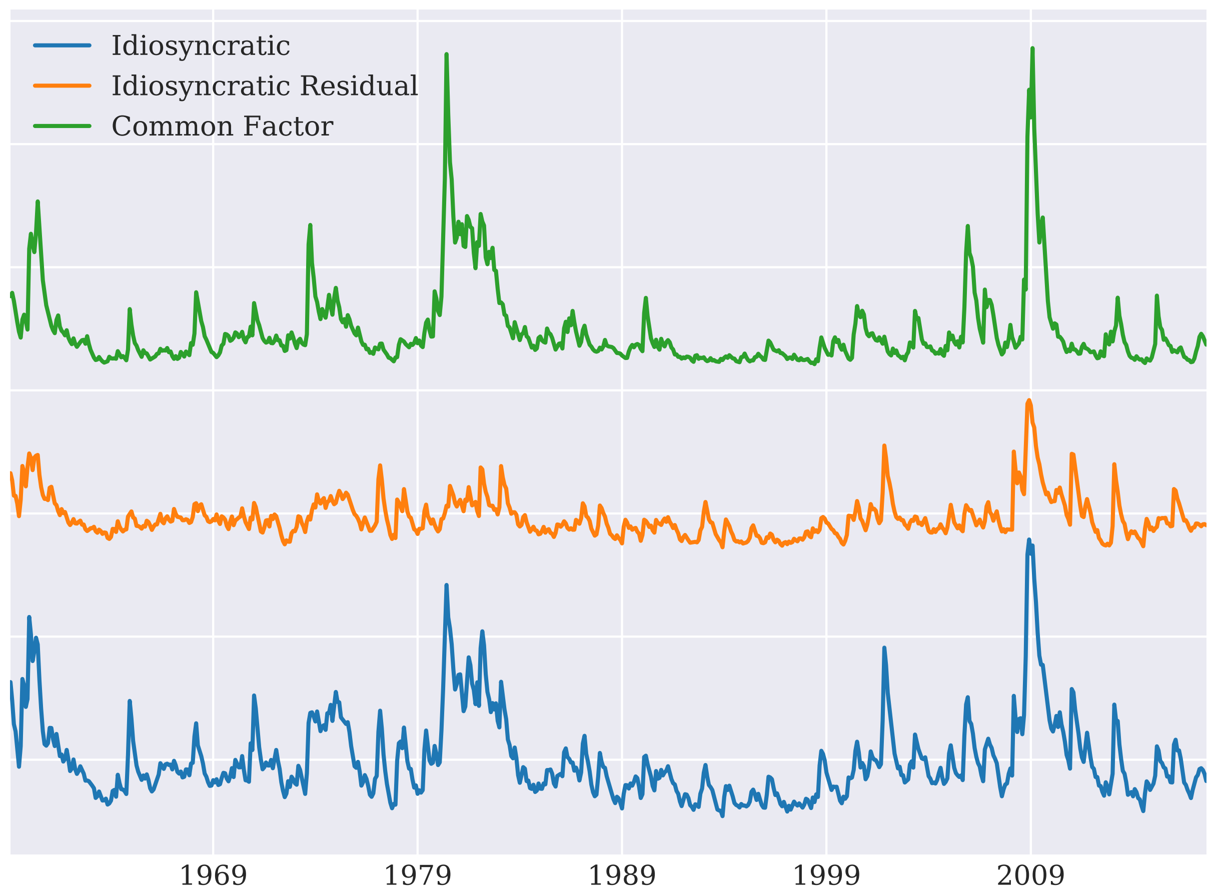

Figure 9: These three series plot the average idiosyncratic volatility as computed from individual model residuals (ξ^i), the volatility of the component of the individual model residuals that is orthogonal to factor shocks $(\xi^i - v_i\xi^f)$, and total volatility that can be explained by common factors. Both idiosyncratic volatility series are from models estimated using composite likelihood. The common factor volatility is from a model with $m = k = 4$.