UNIVERSITY OF
BIRMINGHAM

Department of Economics

An Overview of the Factor-Augmented Error-Correction Model

Department of Economics Discussion Paper 15-03

Anindya Banerjee
Massimiliano Marcellino
Igor Masten

An Overview of the Factor-augmented Error-Correction Model*

Anindya Banerjee† Massimiliano Marcellino‡ Igor Masten§

January 2015

Abstract

The Factor-augmented Error Correction Model (FECM) generalizes the factor-augmented VAR (FAVAR) and the Error Correction Model (ECM), combining error-correction, cointegration and dynamic factor models. It uses a larger set of variables compared to the ECM and incorporates the long-run information lacking from the FAVAR because of the latter's specification in differences. In this paper we review the specification and estimation of the FECM, and illustrate its use for forecasting and structural analysis by means of empirical applications based on Euro Area and US data.

Keywords: Dynamic Factor Models, Cointegration, Structural Analysis, Factor-augmented Error Correction Models, FAVAR

JEL-Codes: C32, E17

*We would like to thank the participants at the 16th Advances in Econometrics Conference, held at CREATES, Aarhus University, for helpful comments.

†Department of Economics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom, e-mail: a.banerjee@bham.ac.uk

‡Bocconi University, IGIER and CEPR, Via Roentgen 1, 20136 Milan, Italy, e-mail: massimiliano.marcellino@unibocconi.it

§University of Ljubljana and Bank of Slovenia, Kardeljeva pl. 17, 1000 Ljubljana, Slovenia, e-mail: igor.masten@ef.uni-lj.si

1 Introduction

Banerjee and Marcellino (2009) introduced the Factor-augmented Error Correction Model (FECM) as a way of bringing together two important recent strands of the econometric literature, namely, cointegration (e.g., Engle and Granger (1988), Johansen (1995)) and large dynamic factor models (e.g., Forni et al. (2000) and Stock and Watson (2002a, 2002b)).

Several papers have emphasized the complexity of modelling large systems of equations in which the complete cointegrating space may be difficult to identify, see for example Clements and Hendry (1995). At the same time, large dynamic factor models and factor augmented VARs (FAVARs, e.g., Bernanke, Boivin and Eliasz (2005) or Stock and Watson (2005)) typically focus on variables in first differences in order to achieve stationarity. In the FECM, factors extracted from large datasets in levels, as a proxy for the non-stationary common trends, are jointly modelled with selected economic variables of interest, with which the factors can cointegrate. In this sense the FECM nests both ECM and FAVAR models, and can be expected to produce better results, at least when the underlying conditions for consistent factor and parameter estimation are satisfied, and cointegration matters.

Banerjee, Marcellino and Masten (2014a) assessed the forecasting performance of the FECM in comparison with the ECM and the FAVAR. Empirically, the relative ranking of the ECM, the FECM and the FAVAR depends upon the variables being modelled and the features of the processes generating the data, such as the amount and strength of cointegration, the degree of lagged dependence in the models and the forecasting horizon. However, in general, the FECM tends to perform better than both the ECM and the FAVAR.

Banerjee, Marcellino and Masten (2014b) evaluated the use of FECM for structural analysis. Starting from a dynamic factor model for nonstationary data as in Bai (2004), they derived the moving-average representation of the FECM and showed how the latter can be used to identify structural shocks and their propagation mechanism, using techniques similar to those adopted by the structural VAR literature.

The FECM model is related to the framework used recently to formulate testing for cointegration in panels (see for example Bai, Kao and Ng, 2008 and Gengenbach, Urbain and Westerlund, 2008). While prima facie the approaches are similar, there are several important differences. First, in panel cointegration, the dimension of the dataset (given by the set of variables amongst which cointegration is tested) remains finite and the units of the panel $i = 1, 2, ..N$ provide repeated information on the cointegrating vectors. By contrast in our framework the dataset in principle is infinite dimensional and driven by a finite number of common trends. Second, following from the first, the role of the factors (whether integrated or stationary) is also different as in the panel cointegrating framework the factors capture cross-section dependence while not being cointegrated with the vector of variables of interest. In our approach this is precisely what is allowed, since the cointegration between the variables and the factors proxies for the missing cointegrating information in the whole dataset.

Another connected though different paper is Barigozzi, Lippi and Luciani (2014). They work with a non-parametric static version of the factor model with common I(1) factors only, while in our context we have a parametric representation of a fully dynamic model where the factors can be both I(1) and I(0), which complicates the analysis, in particular for structural applications related to permanent shocks (see Banerjee et al. (2014b)). They also assume that the factors follow a VAR model, and show that their first differences admit a finite order ECM representation, which is an interesting result. In contrast, we focus on cointegration between the factors and the observable variables. The Barigozzi et al. (2014) model is also similar to the one analysed by Bai (2004). Bai did not consider impulse responses but in his context one could easily get consistent estimates of the responses based on the factors in levels (rather than using the ECM model for the factors as in Barigozzi et al. (2014)).

In this paper we review the specification and estimation of the FECM. We then illustrate its use for forecasting and structural analysis by means of novel empirical applications based on Euro Area and US data.

For the Euro Area we use 38 quarterly macroeconomic time series from the 2013 update of the Euro Area Wide Model (AWM) dataset, over the period 1975 - 2012. For forecasting, we focus on two subsets of three variables each, one real and one nominal. The real set consists of real GDP, real private consumption and real exports. The nominal system, on the other hand, contains the harmonized index of consumer prices (HICP), unit labor costs and the effective nominal exchange rate of the euro. For each set we consider forecasting over the period 2002-2012 and, to investigate the effect the Great Recession might have had on forecasting performance of competing models, we also split the forecasting sample into 2002q1-2008q3 and 2008q4-2012q4. As forecasting models, we use AR, VAR and ECM specifications, with or without factors. For the real variables, the FECM is clearly the best forecasting model, and a comparison with the FAVAR highlights the importance of including the error correction terms. For the nominal variables, the FECM also performs well if the factors are extracted from a subset of the nominal variables only, pre-selected as in Bai and Ng (2006). In terms of the effects of the crisis, the performance of the FECM generally further improves.

For the US, we use the set of monthly real and nominal macroeconomic series from Banerjee, Marcellino and Masten (2014a). The dataset contains over 100 macroeconomic series over the period 1959 - 2003. As real variables, we consider forecasting total industrial production (IP), personal income less transfers (PI), employment on non-agricultural payrolls (Empl), and real manufacturing trade and sales (ManTr). As nominal variables we focus on the producer price index, consumer price index, consumer prices without food prices and private consumption deflator. The forecasting period is 1970 - 1998, which is the same as in Stock and Watson (2002b). The results are again encouraging for the FECM.

In both the Euro Area and US forecasting applications we compare the results from our basic FECM estimation approach that requires all the idiosyncratic errors to be I(0) with an alternative method, based on variables in differences, where the idiosyncratic errors can also be I(1). We find that both methods perform similarly and this finding, in addition to the outcome of formal testing procedures that generally do not reject the hypothesis of I(0) idiosyncratic errors, provide support for our basic FECM estimation method.

Finally, as an illustration of the use of the FECM for structural analysis, we assess the effects of a monetary policy shock. Specifically, we replicate the FAVAR based analysis of Bernanke et al. (2005) in our FECM context, based on their same dataset. The shape of the impulse responses is overall similar across the models for most variables. Quantitatively, however, the responses may differ significantly due to the error-correction terms. For example, quite significant differences are observed for monetary aggregates, the yen-dollar exchange rate, and measures of consumption. Omission of the error-correction terms in the FAVAR model can thus have an important impact on the empirical results.

The paper is structured as follows. Section 2 reviews the representation and estimation of the FECM model, and then specializes the results for the cases of forecasting and structural analysis. Section 3 discusses the data and the models used in the empirical applications. Section 4 presents forecasting results, while Section 5 presents the analysis of monetary policy shocks with the FECM. Section 6 concludes.

2 Factor-augmented error-correction model

In this section we reproduce the derivation of the FECM from Banerjee, Marcellino and Masten (2014b). The starting point of our analysis is the dynamic factor model for I(1) data with both I(1) and I(0) factors, which allows us to distinguish between common stochastic trends and stationary drivers of all variables. We start by deriving the theoretical representation of the FECM. In the empirical applications of the paper, however, the FECM is used for forecasting and structural analysis. These applications require estimable versions of the FECM, which we present in turn in two separate subsections.

2.1 Representation of the FECM

Consider the following dynamic factor model (DFM) for I(1) data:

$$\begin{aligned} X_{it} &= \sum_{j=0}^{p} \lambda_{ij} F_{t-j} + \sum_{l=0}^{m} \phi_{il} c_{t-l} + \varepsilon_{it} \\ &= \lambda_i(L) F_t + \phi_i(L) c_t + \varepsilon_{it}, \qquad (1) \end{aligned}$$

where $i = 1, ..., N$, $t = 1, ..., T$, F_t is an r_1-dimensional vector of random walks, c_t is an r_2-dimensional vector of I(0) factors, $F_t = c_t = 0$ for $t < 0$, and ε_{it} is a zero-mean idiosyncratic component. $\lambda_i(L)$and $\varphi_i(L)$ are lag polynomials of orders p and m respectively, which are assumed to be finite.

The loadings λ_{ij} and ϕ_{ij} are either deterministic or stochastic and satisfy the following restrictions. For $\lambda_i = \lambda_i(1)$ and $\phi_i = \phi_i(1)$ we have $E\|\lambda_i\|^4 \leq M < \infty$, $E\|\phi_i\|^4 \leq M < \infty$, and $1/N \sum_{i=0}^{N} \lambda_i \lambda_i'$, $1/N \sum_{i=0}^{N} \phi_i \phi_i'$ converge in probability to positive definite matrices. Furthermore, we assume that $E(\lambda_{ij}\varepsilon_{is}) = E(\phi_{ij}\varepsilon_{is}) = 0$ for all i, j and s.The idiosyncratic component ε_{it} can be in principle serially and cross-correlated. Specifically, for $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$ we assume that

$$\varepsilon_t = \Gamma(L)\varepsilon_{t-1} + v_t, \tag{2}$$

where v_t are orthogonal white noise errors. If the roots of $\Gamma(L)$ lie inside the unit disc for all i, the model fits the framework of Bai (2004). This assumption implies that X_{it} and F_t cointegrate. If instead ε_{it} are I(1) for some i, then our model fits the framework of Bai and Ng (2004). The following derivation of the FECM representation accommodates both cases.

To derive the FECM and discuss further assumptions upon the model that ensure consistent estimation of the model's components, it is convenient to write first the DFM in static form. To this end, we follow Bai (2004) and define

$$\tilde{\lambda}_{ik} = \lambda_{ik} + \lambda_{ik+1} + ... + \lambda_{ip}, \ k = 0, ..., p.$$

Let us in addition define

$$\tilde{\Phi}_i = [\phi_{i0}, ..., \phi_{im}].$$

Then, we can get a static representation of the DFM which has the I(1) factors isolated from the I(0) factors:

$$X_{it} = \Lambda_i F_t + \Phi_i G_t + \varepsilon_{it} \tag{3}$$

where

$$\begin{aligned}
\Lambda_i &= \tilde{\lambda}_{i0}, \\
\Phi_i &= \left[\tilde{\Phi}_i, -\tilde{\lambda}_{i1}, \ldots, -\tilde{\lambda}_{ip}\right], \\
G_t &= \left[c_t', c_{t-1}', ..., c_{t-m}', \Delta F_t', ..., \Delta F_{t-p+1}'\right]'.
\end{aligned}$$

Introducing for convenience the notation $\Psi_i = [\Lambda_i', \Phi_i']'$, the following assumptions are also needed for consistent estimation of both the I(1) and I(0) factors: $E\|\Psi_i\|^4 \leq M < \infty$ and $1/N \sum_{i=0}^{N} \Psi_i \Psi_i'$ converges to a $(r_1(p+1) + r_2(m+1)) \times (r_1(p+1) + r_2(m+1))$ positive-definite matrix.

Grouping across the N variables we have

$$X_t = \Lambda F_t + \Phi G_t + \varepsilon_t \tag{4}$$

where $X_t = [X_{1t}, ..., X_{Nt}]', \Lambda = [\Lambda_1^{'}, ..., \Lambda_N^{'}]', \Phi = [\Phi_1^{'}, ..., \Phi_N^{'}]'$ and $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$.

The serial correlation of the idiosyncratic component in (4) can be eliminated from the error process by premultiplying (3) by $I - \Gamma(L)L$. As shown in Banerjee, Marcellino and Masten (2014b), straightforward manipulation leads to the ECM form of the DFM, which is the factor-augmented error-correction model (FECM), specified as:

$$\begin{aligned}\Delta X_t = & \underbrace{-(I-\Gamma(1))(X_{t-1} - \Lambda F_{t-1})}_{\text{Omitted in the FAVAR}} + \Lambda \Delta F_t + \Gamma_1(L)\Lambda\Delta F_{t-1} \\ & + \Phi G_t - \Gamma(1)\Phi G_{t-1} + \Gamma_1(L)\Phi\Delta G_{t-1} - \Gamma_1(L)\Delta X_{t-1} + v_t,\end{aligned} \tag{5}$$

where we have used the factorization

$$\Gamma(L) = \Gamma(1) - \Gamma_1(L)(1-L).$$

Equation (5) is a representation of the DFM in (1) in terms of stationary variables. From it, we can directly observe the main distinction between a FAVAR model and the FECM. The latter contains the error-correction term, $-(I - \Gamma(1))(X_{t-1} - \Lambda F_{t-1})$, while in the FAVAR model this term is omitted, leading to an omitted variables problem.

Empirically, the error-correction term can have a signifficant role. Banerjee, Marcellino and Masten (2014b) report for the US data that 63 out of 77 equations for the I(1) variables contain a statistically significant error-correction term. For the Euro Area dataset analyzed in this paper, the score is 27 out of 32 I(1) variables. which is a very similar share to the case of the US dataset.

Note that it follows from (4) that

$$X_{t-1} - \Lambda F_{t-1} = \Phi G_{t-1} + \varepsilon_{t-1},$$

such that it would appear at first sight that the omitted error-correction term in the FAVAR could be approximated by including additional lags of the I(0) factors. However, by substituting the previous expression into (5) and simplifying we get

$$\Delta X_t = \Lambda \Delta F_t + \Phi \Delta G_t + \Delta \varepsilon_t, \tag{6}$$

which contains a non-invertible MA component. Conventional structural analysis in a FAVAR framework relies on inverting a system like (6) (see Stock and Watson (2005) and the survey in Luetkepohl, 2014). Hence, whenever we deal with I(1) data, and many macroeconomic series exhibit this feature, the standard FAVAR model produces biased results unless we use an infinite number of factors as regressors, or account explicitly for the non-invertible MA structure of the error-process.[1]

To complete the model, we assume that the nonstationary factors follow a vector random walk process

$$F_t = F_{t-1} + \varepsilon_t^F, \tag{7}$$

while the stationary factors are represented by

$$c_t = \rho c_{t-1} + \varepsilon_t^c, \tag{8}$$

where ρ is a diagonal matrix with values on the diagonal in absolute term strictly less than one. ε_t^F and ε_t^c are independent of λ_{ij}, ϕ_{ij} and ε_{it} for any i, j, t. It should be noted that the error processes ε_t^F and ε_t^c need not necessarily be *i.i.d.*. They are allowed to be serially and cross correlated and jointly follow a stable vector process:

$$\begin{bmatrix} \varepsilon_t^F \\ \varepsilon_t^c \end{bmatrix} = A(L) \begin{bmatrix} \varepsilon_{t-1}^F \\ \varepsilon_{t-1}^c \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix}, \tag{9}$$

where u_t and w_t are zero-mean white-noise innovations to dynamic nonstationary and stationary factors, respectively. Under the stability assumption, we can express the model as

$$\begin{bmatrix} \varepsilon_t^F \\ \varepsilon_t^c \end{bmatrix} = [I - A(L)L]^{-1} \begin{bmatrix} u_t \\ w_t \end{bmatrix}. \tag{10}$$

Note that, under these assumptions, we have $E \left\| \varepsilon_t^F \right\|^4 \leq M < \infty$, which implies that $\sum_{t=1}^T F_t F_t'$ converges at rate T^2, while $\sum_{t=1}^T G_t G_t'$ converges at the standard rate T. The cross-product matrices $\sum_{t=1}^T F_t G_t'$ and $\sum_{t=1}^T G_t' F_t$ converge at rate $T^{3/2}$. At these rates, the elements of the matrix composed of these four elements jointly converge to form a positive definite matrix.

Using (7), (8) and (10) we can write the VAR for the factors as

$$\begin{aligned} \begin{bmatrix} F_t \\ c_t \end{bmatrix} &= \left[\begin{bmatrix} I & 0 \\ 0 & \rho \end{bmatrix} + A(L) \right] \begin{bmatrix} F_{t-1} \\ c_{t-1} \end{bmatrix} - A(L) \begin{bmatrix} I & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} F_{t-2} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix} \\ &= C(L) \begin{bmatrix} F_{t-1} \\ c_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix}, \end{aligned} \tag{11}$$

where the parameter restrictions imply that $C(1)$ is a block-diagonal matrix with block sizes corresponding to the partition between F_t and c_t.

The FECM is specified in terms of static factors F and G, which calls for a corresponding VAR specification. Using the definition of G_t and (11) it is possible to get the

[1]The model by Barigozzi et al. (2014) is basically (6) with $\Phi = 0$, augmented with an ECM model for ΔF_t.

following representation

$$
\begin{bmatrix}
I & 0 & \cdots & & & \cdots & 0 \\
0 & I & \cdots & & & \cdots & 0 \\
\vdots & & & & & & \vdots \\
\vdots & & & & & & \vdots \\
0 & \cdots & I & 0 & & \cdots & 0 \\
-I & \cdots & 0 & I & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & I & \cdots & 0 \\
\vdots & & & & & & \vdots \\
0 & \cdots & & \cdots & & \cdots & I
\end{bmatrix}
\begin{bmatrix}
F_t \\ c_t \\ c_{t-1} \\ \vdots \\ c_{t-m} \\ \Delta F_t \\ \Delta F_{t-1} \\ \vdots \\ \Delta F_{t-p+1}
\end{bmatrix}
=
$$

$$
\begin{bmatrix}
C_{11}(L) & C_{12}(L) & 0 & \cdots & & \cdots & 0 \\
C_{21}(L) & C_{22}(L) & 0 & \cdots & & \cdots & 0 \\
0 & I & 0 & \cdots & & \cdots & 0 \\
\vdots & & \cdots & & \cdots & & \vdots \\
0 & \cdots & \cdots & I & 0 & \cdots & 0 \\
-I & \cdots & & \cdots & & \cdots & 0 \\
0 & & \cdots & & I & \cdots & 0 \\
\vdots & & & & & & \vdots \\
0 & \cdots & & \cdots & & I & 0
\end{bmatrix}
\begin{bmatrix}
F_{t-1} \\ c_{t-1} \\ c_{t-2} \\ \vdots \\ c_{t-m-1} \\ \Delta F_{t-1} \\ \Delta F_{t-2} \\ \vdots \\ \Delta F_{t-p}
\end{bmatrix}
+
\begin{bmatrix}
I & 0 \\ 0 & I \\ 0 & 0 \\ & \\ \vdots & \vdots \\ & \\ \vdots & \vdots \\ & \\ 0 & 0
\end{bmatrix}
\begin{bmatrix}
u_t \\ w_t
\end{bmatrix}
\qquad (12)
$$

With the definition of G_t, the VAR for the static factors, and premultiplying the whole expression by the inverse of the initial matrix in (12), the factor VAR can be more compactly written as

$$
\begin{bmatrix} F_t \\ G_t \end{bmatrix} = \begin{bmatrix} M_{11}(L) & M_{12}(L) \\ M_{21}(L) & M_{22}(L) \end{bmatrix} \begin{bmatrix} F_{t-1} \\ G_{t-1} \end{bmatrix} + Q \begin{bmatrix} u_t \\ w_t \end{bmatrix}, \qquad (13)
$$

where the $(r_1(p+1) + r_2(m+1)) \times (r_1 + r_2)$ matrix Q accounts for dynamic singularity of G_t. This is due to the fact that the dimension of the vector process w_t is r_2, which is smaller than or equal to $r_1 p + r_2(m+1)$, the dimension of G_t. In what follows we assume that the order of the VAR in (13) is n.

2.2 The FECM form for forecasting

The specification in (5) is not a convenient forecasting model as it is heavily parameterized, which makes it very difficult or even impossible to estimate with standard techniques when N is large. Hence, we focus on forecasting a small set of variables, as in Banerjee, Marcellino and Masten (2014a). These variables of interest, a subset of X, are denoted by X_A.

According to (5) X_{At} cointegrate with F_t, which means that we can model them with an error-correction specification. Note, however, that we need to incorporate into the model also the information in the I(0) factors G_t. Given that the FECM model (5) can be re-written also as

$$\Delta X_t = -(I - \Gamma(1))(X_{t-1} - \Lambda F_{t-1} - \Phi G_{t-1}) + \Lambda \Delta F_t + \Gamma_1(L)\Lambda \Delta F_{t-1} \\ + \Phi \Delta G_t + \Gamma_1(L)\Phi \Delta G_{t-1} - \Gamma_1(L)\Delta X_{t-1} + v_t, \tag{14}$$

this implies that G_t is best included in the cointegration space. This way the forecasting model can be written as

$$\begin{bmatrix} \Delta X_{At} \\ \Delta F_t \\ \Delta G_t \end{bmatrix} = \begin{bmatrix} \gamma_A \\ \gamma_F \\ \gamma_G \end{bmatrix} \delta' \begin{bmatrix} X_{At-1} \\ F_{t-1} \\ G_{t-1} \end{bmatrix} + B_1 \begin{bmatrix} \Delta X_{At-1} \\ \Delta F_{t-1} \\ \Delta G_{t-1} \end{bmatrix} + \cdots + B_q \begin{bmatrix} \Delta X_{At-q} \\ \Delta F_{t-q} \\ \Delta G_{t-q} \end{bmatrix} + \begin{bmatrix} \epsilon_{At} \\ \epsilon_t^F \\ \epsilon_t^G \end{bmatrix}. \tag{15}$$

(15) is clearly an approximation of the original model in (5). Its parameterization, dictated by empirical convenience for forecasting applications, deserves a few comments. First, while in the model (5) cointegration is only between each individual variable and the factors (due to the assumed factor structure of the data), we treat the cointegration coefficients δ as unrestricted. This is because (15) is only an approximation to the original model and omits potentially many significant cross-equations correlations. For a similar reason, the loading matrices γ_A, γ_F and γ_G and short-run coefficients $B_1, ..., B_q$ are also left unrestricted. The lag structure of the model in such a case cannot be directly recovered from the orders of $\Gamma(L)$ and $M(L)$, in our empirical applications it is determined by suitable information criteria. Note that the extent of the potential mis-specification of (15) depends mainly on the structure of the $\Gamma(L)$ matrix in (2), which in turn depends on the extent of the cross-correlation of the idiosyncratic errors. With a diagonal $\Gamma(L)$, hence uncorrelated idiosyncratic errors, (15) is very close to (5).

Conditional on the estimated factor space, the remaining parameters of the model can be estimated using the Johansen method (Johansen, 1995). The rank of δ can be determined, for example, either by the Johansen trace test (Johansen, 1995) or the procedure of Cheng and Phillips (2009) based on information criteria. Hence, we focus on factor estimation.

Estimation of the space spanned by the factors and of their number depend on the properties of the idiosyncratic components ε_{it}. Under the assumption of I(0) idiosyncratic errors, the number of I(1) factors r_1 can be consistently estimated using the criteria developed by Bai (2004), applied to data in levels. The overall number of static factors $r_1(p+2)+r_2(m+1)$ can be estimated using the criteria by Bai and Ng (2002), applied to the data in differences. The space spanned by the factors can be consistently estimated using principal components. F_t can be consistently estimated as the eigenvectors corresponding to the largest r_1 eigenvalues of XX' normalized such that $\widetilde{F}'\widetilde{F}/T^2 = I$. The stationary factors can be consistently estimated as the eigenvectors corresponding to the next q largest eigenvalues normalized such that $\widetilde{G}'\widetilde{G}/T = I$ (Bai, 2004).

In case some of the ε_{it} are I(1), the space spanned by F_t and G_t jointly (but not separately) can be estimated consistently using the method by Bai and Ng (2004), from data in differences.

Replacing the true factors with their estimated counterparts is permitted under the assumptions discussed above and in Bai (2004) (see Bai (2004), Lemma 3, p. 148) or Bai and Ng (2004), so that we do not have a generated-regressors problem.

Even though the FECM can accommodate either of the assumptions about the order of integration of the idiosyncratic components, we give preference in our empirical applications to the Bai (2004) setting with $I(0)$ idiosyncratic components, but also provide results with factors obtained from the data in differences as a robustness check. There are two main reasons for our choice. First, from an economic point of view, integrated errors are unlikely as they would imply that the integrated variables can drift apart in the long run, contrary to general equilibrium arguments. This is especially so in our forecasting applications, in which we consider forecasting a small set of key observable variables. Integrated variables that drift apart are likely marginal, and as such they do not contain essential information and can be dropped from the analysis. Second, whether the idiosyncratic errors ε_{it} are stationary or not is an empirical issue.The empirical applications below use two datasets. The first one is composed of the Euro Area quarterly variables used in Fagan et al. (2001), updated to cover the period 1975 - 2013. It contains 32 I(1) series. The second uses a monthly US dataset for the period 1959 - 2003, taken from Bernanke, Boivin and Eliasz (2005) with 77 I(1) series. By applying the ADF unit root test to the estimated idiosyncratic components after extracting 4 factors from each dataset (as indicated by appropriate information criteria), the unit-root null is rejected at the 5% significance level for all series in the Euro area dataset, while for the US data for a few series rejection occurs at the 10% level and at 5% for the remaining series. Moreover, the panel unit root test (Bai and Ng, 2004) rejects the null of no panel cointegration between X_{it} and F_t for both datasets. Overall, it appears that the assumption of stationary idiosyncratic errors fits well the properties of the two datasets we use.

2.3 The FECM form for structural analysis

The identification of structural shocks in a standard VAR model relies on imposing restrictions upon the parameters of the moving-average representation of the VAR and/or the variance covariance matrix of the VAR errors. An analogous approach in the case of large-scale models entails the moving-average representation of the FAVAR. In the general case, this requires the estimation of the VAR representation of the dynamic factor model (see Stock and Watson, 2005 and Lütkepohl, 2014) or, in case of large nonstationary panels with cointegration, the equations of the FECM (rather than just the approximation in (15)).

To avoid the curse of dimensionality in estimating either the FAVAR or the FECM, we need to strengthen the assumptions about the properties of the idiosyncratic components. Specifically, we assume (1) to be a strict factor model: $E(\varepsilon_{it}, \varepsilon_{js}) = 0$ for all i, j, t and s, $i \neq j$.[2] However, serial correlation of ε_{it} is still permitted in the form $\varepsilon_{it} = \gamma_i(L)\varepsilon_{it-1} + v_{it}$ with the roots of $\gamma_i(L)$ lying inside the unit disc. Under this assumption we can write the lag polynomial $\Gamma(L)$ as

$$\Gamma(L) = \begin{bmatrix} \gamma_1(L) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \gamma_N(L) \end{bmatrix}.$$

This restriction, being stronger than Bai's assumptions, leaves all of his results directly applicable to our model, as also verified by the simulation experiments reported by Banerjee, Marcellino and Masten (2014b). Under the strict dynamic factor assumption, the estimation of the parameters of the FECM model (5) is straightforward. Using the estimated factors and loadings, the estimates of the common components are $\widetilde{\Lambda}\widetilde{F}_t$, $\widetilde{\Phi}\widetilde{G}_t$, $\widetilde{\Lambda}\Delta\widetilde{F}_t$ and $\widetilde{\Phi}\Delta\widetilde{G}_t$, while for the cointegration relations it is $X_{t-1} - \widetilde{\Lambda}\widetilde{F}_{t-1}$. Finally, the estimated common components and cointegration relations can be used in (5) to estimate the remaining parameters of the FECM by OLS, equation by equation. Also in this case, replacing the true factors and their loadings with their estimated counterparts is permitted under the assumptions discussed above and in Bai (2004) (see Bai,2004, Lemma 3, p. 148) so that we do not have a generated-regressors problem.

The FECM model (5) and the corresponding factor VAR representation (13) are in reduced form. The identification of structural shocks in VAR models usually rests on imposing restrictions upon the parameters of the moving-average representation of the VAR. For vector-error correction models, the derivation of the moving-average representation uses the Granger representation theorem. The generalization of the Granger representation theorem to large dynamic panels is provided by Banerjee, Marcellino and Masten (2014b) who show that the moving-average representation of the FECM is

$$\begin{bmatrix} X_t \\ F_t \\ G_t \end{bmatrix} = \begin{bmatrix} \Lambda \\ I_{r_1} \\ 0_{r_2 \times r_1} \end{bmatrix} \omega \sum_{i=1}^{t} u_t + C_1(L) \begin{bmatrix} v_t + [\Lambda, \Phi] Q [u_t', w_t']' \\ Q \begin{bmatrix} u_t \\ w_t \end{bmatrix} \end{bmatrix}, \tag{16}$$

where $C_1(L)$ is a stable matrix polynomial and the remaining notation is as above.

Our model contains I(1) and I(0) factors with corresponding dynamic factors innovations. From the MA representation (16), we can observe that the innovations in the first group have permanent effects on X_t, while the innovations in the second group have only transitory effects, which makes the FECM a very useful model also for the application of

[2]Stock and Watson (2005) show on the US dataset that the strict factor model assumption is generally rejected but is of limited quantitative importance.

long-run identifications schemes.

For the purposes of the identification of structural dynamic factor innovations, we assume that they are linearly related to the reduced-form innovations as

$$\tilde{\eta} = \begin{bmatrix} \eta_t \\ \mu_t \end{bmatrix} = H \begin{bmatrix} u_t \\ w_t \end{bmatrix}, \tag{17}$$

where H is a full-rank $(r_1 + r_2) \times (r_1 + r_2)$ matrix. η_t are r_1 permanent structural dynamic factor innovations and μ_t are r_2 transitory structural dynamic factor innovations. It is assumed that $E\tilde{\eta}_t\tilde{\eta}_t' = I$ such that $H\Sigma_{u,w}H' = I$.

3 Data and empirical applications

The empirical applications below illustrate the performance of the FECM in forecasting and structural analysis of monetary policy shocks. The forecasting application is based on Euro Area data, coming from the 2013 update of the Euro Area Wide Model dataset of Fagan et al. (2001). It contains 38 quarterly macroeconomic series for the period 1975-2012.[3] 32 out of 38 series are $I(1)$. Data are seasonally adjusted at source. The only exception is the consumer price index, which we seasonally adjust using the X-11 procedure. On this dataset we investigate also whether the method of factors extraction - either from the levels or differences of the data - affect the forecasting performance of the FECM.

Further evidence on this matter is subsequently provided in a second forecasting application, which is based on data from Stock and Watson (2005), containing 132 monthly series, 104 of which are treated as I(1). We used this dataset in Banerjee, Marcellino and Masten (2014b), extracting the factors from the variables in levels. Here we get and compare FECM forecasts obtained with factors extracted from differences, to allow for possible I(1) idiosyncratic errors.

The structural application is an analysis of the transmission of monetary policy shocks, based on the FAVAR study of Bernanke, Boivin and Eliasz (2005) and for comparability we use their dataset for the US. It contains 120 monthly variables, spanning the period 1959-2003. 77 variables are by the authors treated as I(1).[4]

Bai (2004) IPC2 information criterion indicates $r_1 = 2$ for both the US and EA datasets. The choice of the total number of estimated factors for the Euro Area, r, is instead based on Bai and Ng (2004). Their PC3 criterion indicates 4 factors in total. For the US, in the choice of the total number of estimated factors r we follow Bernanke et

[3]The data and the corresponding list of variables can be downloaded from the Euro area business cycle network webpage (www.eabcn.org/area-wide-model).

[4]The structure of the Euro Area data is not rich enough to implement this structural analysis. In particular, it does not contain a sufficient number of fast-moving variables (those that react contemporaneously to the monetary policy shock). In addition, the comparison of the FECM with the Bernanke et al. (2005) FAVAR based results is of interest by itself.

al. (2005) and set it to 3. Including the federal funds rate as an observable factor, as in Bernanke et al. (2005), gives a total number of factors equal to 4, as for the Euro Area application. However, as in their case, the main findings are robust to working with more factors. [5]

The datasets contain both I(1) and I(0) variables. The I(0) variables in the panel are treated in the empirical analysis in the following way. At the stage of factor estimation all variables are used. The space spanned by F_t and G_t is estimated by the principal components of the data in levels containing both the I(1) and I(0) variables (Bai, 2004), whose good finite sample performance is confirmed by a simulation experiment in Banerjee, Marcellino and Masten (2014b). The structure of the FECM equations, however, needs to be adapted for the purposes of the structural analysis.

Denote by X_{it}^1 the $I(1)$ variables and by X_{it}^2 the $I(0)$ variables. Naturally, the issue of cointegration applies only to X_{it}^1. As a consequence, the $I(1)$ factors load only to X_{it}^1 and not to X_{it}^2. In other words, the fact that X_{it}^2 are assumed to be $I(0)$ implies that the $I(1)$ factors F_t do not enter the equations for X_{it}^2, which is a restriction that we take into account in model estimation. Our empirical FECM is then:[6]

$$\Delta X_{it}^1 = \alpha_i(X_{it-1}^1 - \Lambda_i F_{t-1}) + \Lambda_i^1(L)\Delta F_t + \Phi_i^1(L)G_t + \Gamma^1(L)\Delta X_{it-1}^1 + v_{it}^1 \quad (18)$$

$$X_{it}^2 = \Phi_i^2(L)G_t + \Gamma^2(L)\Delta X_{it-1}^2 + v_{it}^2 \quad (19)$$

The model for the I(1) variables in (18) is the FECM, while the model for the I(0) variables in (19) is a FAVAR. Note that these FAVAR equations differ from standard applications. The initial model from which we derived the FECM is the DFM for $I(1)$ data. In such a model the $I(1)$ factors by definition cannot load onto $I(0)$ variables. This restriction is explicit in (19), while the FAVAR application of Bernanke, Boivin and Eliazs (2005), for example, uses the following form of the FAVAR

$$\Delta X_{it}^1 = \Lambda_i^1(L)\Delta F_t + \Phi_i^1(L)G_t + v_{it}^1 \quad (20)$$

$$X_{it}^2 = \Lambda_i^2(L)\Delta F_t + \Phi_i^2(L)G_t + v_{it}^2 \quad (21)$$

As discussed above, the main difference between the FECM and the FAVAR is that the latter does not contain the error-correction term.

[5]For the US dataset used in second forecasting application and based on Stock and Watson (2005), none of the Bai and Ng (2004) criteria give conclusive evidence about r. For comparability with the EA dataset and our previous analysis with US data in Banerjee, Marcellino and Masten (2014a), we also set the total number of factors for this US dataset to 4.

[6]Note that only levels of G_t enter (18) while also differences are present in (5). Given that G_t is $I(0)$ we can see $\Gamma^1(L)$ as coming from a repramaterization of the polynomial $\Gamma_1(L)\Phi$ in (5).

4 Forecasting macroeconomic variables

We start with the presentation of forecasting results for selected Euro Area variables. We consider two systems of three variables each, one real and one nominal. The real set consists of the real GDP, real private consumption and real exports. The nominal set, on the other hand, contains the harmonized index of consumer prices (HICP), unit labor costs and the effective nominal exchange rate of the euro. For each set, we consider forecasting over the whole forecast period 2002 - 2012 and, to investigate the effects the Great Recession might have had on forecasting performance of competing models, we also split the forecasting sample into 2002q1-2008q3 and 2008q4-2012q4.

Forecasting is performed using the following set of competing models. First, we use three models that are all based on the observable variables only: an autoregressive model (AR), a vector autoregression (VAR) and an error-correction model (ECM). In order to assess the forecasting role of the additional information, the second set of models augments the first set with factors extracted from the larger set of available variables: FAR, FAVAR and FECM specifications are factor-augmented AR, VAR and ECM models, respectively.

For the FECM model we use two approaches to factor extraction. As argued above, our primary choice is estimation with PCA from the data in levels. As a robustness check, commented in the next subsection, we use the factors estimated from the data in differences, using the method of Bai and Ng (2004). Such a FECM model is denoted $FECM_{BN}$.

The numbers of $I(1)$ and $I(0)$ factors, both set at 2, are kept fixed over the forecasting period, but their estimates are updated recursively. Each forecasting recursion also includes model selection. The lag lengths are determined by the BIC information criterion.[7] As for the cointegration test for determining the cointegration ranks of the ECM and the FECM, we have considered two approaches: the Johansen trace test (Johansen, 1995) and the Cheng and Phillips (2009) semi-parametric test based on the BIC. The two methods gave very similar results (details available upon request), but, due to its lower computational burden and also its ease of implementation in practice, we gave preference to the method of Cheng and Phillips.[8]

The levels of all variables are treated as $I(1)$ with a deterministic trend, which means that the dynamic forecasts of the differences of (the logarithm of) the variables h steps ahead produced by each of the competing models are cumulated in order to obtain the forecasts of the level h steps ahead. We consider four different forecast horizons, $h = 1, 2, 4, 8$. In contrast to our use of iterated h-step-ahead forecasts (dynamic forecasts), Stock and

[7] We have also checked and confirmed the robustness of the results when using the Hannan-Quinn (HQ) criterion (details are available upon request).

[8] The simulation results provided by Cheng and Phillips (2009) show that using the BIC tends to lead to underestimation of the rank when the true rank is not very low, while it performs best when the true cointegration rank is very low (0 or 1). Given that BIC model selection is generally preferred for model selection for forecasting, we chose to use it for testing for cointegration rank as well. However, our results (available upon request) are robust to the use of HQ too.

Watson (1998, 2002a,b) adopt direct h-step-ahead forecasts, while Marcellino, Stock and Watson (2006) find that iterated forecasts are often better, except in the presence of substantial misspecification.[9] In our FECM framework, such forecasts are easier to construct than their h-step-ahead equivalents, and the method of direct h-step-ahead forecasts and our iterative h-step-ahead forecasts produce similar benchmark results on a common estimation and evaluation sample.

The results of the forecast comparisons are presented in Tables 1 to 8, where we list the MSEs of the competing models relative to the MSE of the AR at different horizons for each variable under analysis, with asterisks indicating when the MSE differences are statistically significant according to the Clark and West (2007) test. The tables also report information on the cointegration rank selection and the number of lags in each model.

4.1 Forecasting results for the Euro area

Our basic results are presented in Table 1 for real variables and in Table 2 for nominal variables. For real variables we can in general observe a very good performance of the FECM. It results to be the best performing model in 11 out of 12 cases. In all of the cases the gains over the AR model are statistically significant according to the Clark and West (2007) test. Only for private consumption at 2-year horizon the forecasting precision of the FECM is lower than than that of the AR model. Moreover, the gains in forecasting precision relative to the basic AR model are rather stable across forecast horizons and range between 10 and 25%. The maximum gain is about 50% for real exports at 2-year horizon.

Other models perform considerably worse. The VAR, the ECM and the FAVAR outperform the AR model only for GDP, but are never better than the FECM. The FAR model turns out to be worse than the AR model in nearly all cases.

A final observation goes to the results of cointegration testing. As we can observe, the Cheng and Phillips test fails to find cointegration between the three variables under evaluation. By adding factors to the system to get the FECM, the test consistently signals cointegration. Such a results is in line with the analysis of Banerjee and Marcellino (2009) who point that adding factors to the ECM proxies for the potentially missing cointegration relations. In combination with the superior information set this results in a better forecasting performance. By comparing the results of the FAVAR and the FECM we see that the FECM is consistently more precise in forecasting. This difference can be attributed to the error-correction term that the FAVAR model omits.

The results on forecasting nominal variables show a very different image. The FECM results to be the best performing model in only one instance, HICP at 2-year horizon, but the gains with respect to the AR model are small. This is in line with similar findings for the case of the US in Banerjee, Marcellino and Masten (2014a). The AR model is regularly

[9] Our use of iterated h-step-ahead forecasts implies that the FAR is essentially a FAVAR containing only one variable of interest and factors.

Table 1: Forecasting real variables for the Euro area over 2002 - 2012

		RMSE		**MSE relative to MSE of AR model**				
h	**Variable**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
	GDP	0.006	1.39	0.94 ***	0.75 ***	0.92 *	0.74 *	0.64 **
1.00	Consumption	0.004	1.51	1.36	1.42	1.05	0.78 *	1.02
	Exports	0.007	0.95	1.06	1.02	1.04	0.90 **	0.82 **
	GDP	0.013	1.17	0.92 ***	0.85 ***	0.85 *	0.72 *	0.75 **
2.00	Consumption	0.007	1.71	1.59	1.60	1.18	0.76 *	1.08
	Exports	0.014	1.10	1.14	1.14	1.13	0.88 *	0.91 **
	GDP	0.024	1.05	0.93	0.94	0.90 *	0.78 **	0.86 ***
4.00	Consumption	0.013	1.66	1.62	1.68	1.36	0.91 *	1.22
	Exports	0.028	1.20	1.21	1.21	1.23	0.77 *	0.80 *
	GDP	0.040	1.01	0.94	0.99	0.95 **	0.84 ***	0.95
8.00	Consumption	0.025	1.50	1.49	1.56	1.34	1.04	1.19
	Exports	0.052	1.23	1.24	1.22	1.33	0.51 *	0.66 *
		AR	0.81	2.51	1.24			
	Lags	FAR	0.27	0.27	1.00			
		VAR	FAVAR	ECM	FECM	FECM$_{BN}$		
		1.00	1.00	1.99	1.00	1.00		
				Cointegration rank				
	ECM			FECM			FECM$_{BN}$	
mean	min	max	mean	min	max	mean	min	max
0.00	0.00	0.00	1.36	1.00	2.00	1.43	1.00	2.00

Notes: The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model.

beaten by the VAR, the FAVAR, the ECM and the FECM for unit labor costs, with the ECM being the best performing model overall.[10] The gains in forecasting precision are significant, reaching also levels above 50% with the ECM model above the 1-year horizon. Similarly to the case of the HICP, also for the nominal exchange rate the AR model results to be consistently the best. No other model yielded a smaller RMSE over the evaluation period.

To further explore the problems with forecasting nominal variables we consider two modifications to the factor extraction procedure. First, we consider extracting factors from subpanels containing either real or nominal variables only. Both subpanels in such a case contain 19 variables. Second, we use variable preselection based on correlation with target variables as in Boivin and Ng (2006). The correlation threshold was set to 0.75.

Table 3 reports the results with factors extracted from the sub-panel of nominal variables.[11] Such a modification results in marked improvement in the performance of the FAR model that turns out to be the best performing model in 4 cases, 3 of which are for the nominal exchange rate at horizons larger than 1 quarter. Similar improvements, but of smaller magnitude, are observed for the FAVAR model and, to an even smaller extent, for the FECM.

In Table 4 we constrain the sub-panel of variables for factors extraction even further by imposing a 0.75 correlation threshold with either of the three modelled variables. This

[10] Note that because there is no cointegration between the variables identified by the Cheng and Phillips (2008) test (see results in Table 2) the ECM model is essentially a VAR. Its results differ from those of the conventional VAR, however, because of the differences in the lag structure.

[11] Extracting factors from the real subpanel did not lead to tangible improvements.

Table 2: Forecasting nominal variables for the Euro area over 2002 - 2012

h	**Variable**	**RMSE of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
			MSE relative to MSE of AR model					
	HICP	0.004	4.05	2.46	3.08	2.15	2.08	2.14
1.00	ULC	0.006	1.61	1.03	1.57	0.85 **	1.06	0.96 *
	Nominal XR	0.021	1.05	0.99	1.11	1.03	1.15	1.09
	HICP	0.007	2.53	1.78	1.73	1.45	1.41	1.54
2.00	ULC	0.010	1.70	0.76 *	1.12	0.61 **	0.89 ***	0.73 **
	Nominal XR	0.036	1.00	1.03	1.15	1.04	1.29	1.14
	HICP	0.012	2.07	1.44	1.46	1.19	1.04	1.56
4.00	ULC	0.020	1.76	0.63 **	0.77 ***	0.46 ***	0.70 ***	0.59 ***
	Nominal XR	0.056	0.99	1.02	1.10	1.02	1.50	1.17
	HICP	0.022	2.15	1.57	1.67	1.18	0.98 ***	2.29
8.00	ULC	0.039	1.96	0.60 **	0.58 **	0.46 **	0.51 **	0.62 **
	Nominal XR	0.084	1.00	0.99	1.02	0.98	1.53	0.98
		AR	4.93	4.00	1.00			
	Lags	FAR	1.00	0.29	0.32			
		VAR	FAVAR	ECM	FECM	FECM$_{BN}$		
		1.43	1.00	1.98	1.00	1.00		
				Cointegration rank				
	ECM			FECM			FECM$_{BN}$	
mean	min	max	mean	min	max	mean	min	max
0.00	0.00	0.00	1.51	1.00	2.00	1.06	0.25	2.00

Notes: The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model.

way the number of variables in the sub-panel shrinks to 14. We can observe that the results improve further to the benefit of factor-based models. The FAR model and the FECM outperform the AR model in 9 out of 12 cases. Both consistently for the unit labor costs and the nominal exchange rate, and the FECM also for the HICP at 1-year and 2-year horizons. The FAVAR model outperforms the AR model in 8 cases. The FECM model also turns out to be the best performing model in 5 cases, which cannot be observed without preselecting variables. Note, that the inclusion of the error-correction term is in our application conditional upon testing for rank in each forecasting recursion. If the test indicates rank zero the error-correction term is not included. This turns to be consistently the case for the ECM and to a large extent also for the FECM. We see from the bottom of Table 4 that the FECM cointegration rank is on average 0.34, which implies that on average in less than half of the recursions the Cheng and Phillips test indicated non-zero rank. However, even this low fraction of cases where the error correction terms matter is sufficient for the FECM to work substantially better than the FAVAR.

4.2 Forecasting before and in the Great Recession

We now split the forecast evaluation sample into the period before (2002q1 - 2008q3) and with/after the Great Recession (2008q4 - 2012q4). Results are presented in Tables 5 and 6 for real and nominal variables, respectively. It can be observed from the tables that the RMSE of the benchmark AR model generally increased in the crisis both for real and nominal variables. The relative performance of the FECM, however, generally improved also for both sets of variables.

Table 3: Forecasting nominal variables for the Euro area over 2002 - 2012, factors extracted from nominal subpanel

		RMSE		**MSE relative to MSE of AR model**				
h	**Variable**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
	HICP	0.004	3.40	2.46	2.61	2.15	1.97	1.88
1.00	ULC	0.006	0.71 *	1.03	1.30	0.85 **	1.11	1.04
	Nominal XR	0.021	1.04	0.99	0.98	1.03	1.04	0.96 **
	HICP	0.007	2.07	1.78	1.43	1.45	1.15	1.26
2.00	ULC	0.010	0.65 ***	0.76 *	0.85 **	0.61 **	0.79 ***	0.65 **
	Nominal XR	0.036	0.95 ***	1.03	0.97 ***	1.04	1.08	0.86 *
	HICP	0.012	1.78	1.44	1.41	1.19	0.97 **	1.31
4.00	ULC	0.020	0.75 ***	0.63 **	0.70 ***	0.46 ***	0.62 ***	0.51 ***
	Nominal XR	0.056	0.89 *	1.02	0.90 *	1.02	1.05	0.87 *
	HICP	0.022	2.15	1.57	1.81	1.18	1.19	1.43
8.00	ULC	0.039	1.02	0.60 **	0.64 **	0.46 **	0.42 **	0.44 **
	Nominal XR	0.084	0.84 ***	0.99	0.87 ***	0.98	0.99	0.90
		AR	4.93	4.00	1.00			
	Lags	FAR	1.00	0.99	1.00			
		VAR	FAVAR	ECM	FECM	FECM$_{BN}$		
		1.43	1.00	1.98	1.00	1.00		
Cointegration rank								
	ECM			FECM			FECM$_{BN}$	
mean	min	max	mean	min	max	mean	min	max
0.00	0.00	0.00	0.09	0.00	2.00	1.25	0.00	2.00

*Notes:*The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model.

Table 4: Forecasting nominal variables for the Euro area over 2002 - 2012, factors extracted from nominal subpanel with Bai and Ng (2006) pre-selected variables

		RMSE		**MSE relative to MSE of AR model**				
h	**Variable**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
	HICP	0.004	1.97	2.46	1.91	2.15	1.78	1.77
1.00	ULC	0.006	1.07	1.03	1.27	0.85 **	1.16	1.17
	Nominal XR	0.021	0.99	0.99	0.98 ***	1.03	0.99	0.99
	HICP	0.007	0.99 ***	1.78	0.98 ***	1.45	1.00 **	0.98 **
2.00	ULC	0.010	0.78 ***	0.76 *	0.86 ***	0.61 **	0.76 ***	0.77 ***
	Nominal XR	0.036	0.95 **	1.03	0.99	1.04	1.02	1.00
	HICP	0.012	0.81 **	1.44	0.88 **	1.19	0.78 *	0.76 *
4.00	ULC	0.020	0.48 ***	0.63 **	0.52 ***	0.46 ***	0.50 ***	0.51 ***
	Nominal XR	0.056	0.97	1.02	0.99	1.02	0.95 ***	0.95 ***
	HICP	0.022	0.98 **	1.57	1.04	1.18	0.82 ***	0.81 ***
8.00	ULC	0.039	0.32 **	0.60 **	0.30 **	0.46 **	0.29 **	0.30 **
	Nominal XR	0.084	1.00	0.99	1.02	0.98	0.98	0.98
		AR	4.93	4.00	1.00			
	Lags	FAR	1.00	1.00	1.00			
		VAR	FAVAR	ECM	FECM	FECM$_{BN}$		
		1.43	1.00	1.98	1.00	1.00		
Cointegration rank								
	ECM			FECM			FECM$_{BN}$	
mean	min	max	mean	min	max	mean	min	max
0.00	0.00	0.00	0.34	0.00	1.00	0.32	0.00	1.00

*Notes:*The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model. Variables pre-selected as in Boivin and Ng (2006) using 0.75 threshold for the correlation coefficient.

For real variables and the period before the crisis, the ECM results to be the best model in 10 out of 12 cases. In the remaining cases the best model is the benchmark AR model. The FECM is more precise than the AR in half of the cases. the FAVAR performs worse, outperforming the AR model only once. The FAR never beats the AR.

In the crisis period results are fundamentally different. The FECM is the best performing model in 11 out of 12 cases. The ECM, on the other hand, is never the best, and beats the AR model only for the GDP. Similar observations apply to the FAVAR and the VAR models. This implies that in the crisis period the importance of the information contained in factors increased for real variables. However, given that the FAVAR is not significantly better than the VAR model, it is important that the information embedded in factors enters via cointegration relations.

Table 5: Forecasting real variables for the Euro area before and in the Great Recession

h	**Log of**	**RMSE of AR**	**MSE relative to MSE of AR model** **FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
			Before crisis - 2002:1-2008:3					
	GDP	0.004	1.18	0.88 *	0.95 ***	0.84 *	0.94	0.76 *
1.00	Consumption	0.003	1.09	1.06	1.55	0.87 *	1.13	1.24
	Exports	0.004	1.38	1.07	1.44	0.98	1.29	1.24
	GDP	0.007	1.04	0.91 *	0.96	0.82 *	0.93 ***	0.67 *
2.00	Consumption	0.006	1.12	1.17	1.47	0.90 **	0.99	1.07
	Exports	0.007	1.42	1.03	1.44	0.88 *	1.10	1.41
	GDP	0.011	1.02	0.98	1.02	0.90 ***	0.96	0.61 *
4.00	Consumption	0.009	1.10	1.26	1.48	1.07	1.08	0.99
	Exports	0.018	1.27	1.06	1.29	0.82 **	0.79 *	1.18
	GDP	0.018	1.00	0.99	1.02	0.95	1.26	0.52 **
8.00	Consumption	0.017	1.07	1.17	1.31	1.05	1.52	0.42 *
	Exports	0.046	1.16	1.05	1.18	0.82 *	0.28 *	0.64 *
			Crisis - 2008:4-2012:4					
	GDP	0.009	1.47	0.97	0.71 ***	0.95 **	0.71 *	0.61 ***
1.00	Consumption	0.005	1.83	1.55	1.35	1.20	0.54 *	0.84 ***
	Exports	0.010	0.86	1.06	0.93	1.06	0.81 *	0.72 ***
	GDP	0.019	1.19	0.92	0.84 ***	0.86 *	0.71 *	0.76 **
2.00	Consumption	0.009	2.01	1.80	1.68	1.33	0.68 *	1.08
	Exports	0.020	1.03	1.17	1.07	1.19	0.84 **	0.76 **
	GDP	0.036	1.06	0.92 ***	0.93	0.90 *	0.76 *	0.88 ***
4.00	Consumption	0.018	1.87	1.75	1.76	1.47	0.88 *	1.35
	Exports	0.037	1.18	1.27	1.19	1.38	0.81 *	0.67 *
	GDP	0.058	1.01	0.93 ***	0.98	0.95 *	0.81 *	0.99
8.00	Consumption	0.034	1.67	1.62	1.66	1.46	1.02	1.50
	Exports	0.058	1.31	1.42	1.31	1.82	0.78 *	0.60 *

*Notes:*The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model.

Similar conclusions about the role of information extracted from large datasets and cointegration can be obtained also by examining the results for nominal variables in Table 6. As in Table 4 also here the factors are extracted from the nominal subpanel and with variable preselection. For the period before the crisis the AR model is the best performing on average: 5 out 12 cases. Among the competing models only the VAR turns out to perform similarly, being the best in 3 out of 4 cases for the nominal exchange rate. The FAR follows, being the best twice and outperforming the AR model in half of the cases. The remaining models outperform the AR in only 3 cases or less.

Table 6: Forecasting nominal variables for the Euro area before and in the Great Recession, factors extracted from nominal sub panel

h	Variable	RMSE of AR	FAR	VAR	FAVAR	ECM	FECM	FECM$_{BN}$
			MSE relative to MSE of AR model					
				Before crisis - 2002:1-2008:3				
	HICP	0.003	2.49	3.37	2.59	2.60	1.79	2.11
1.00	ULC	0.004	1.30	1.66	1.42	1.02	1.40	1.19
	Nominal XR	0.019	0.97 ***	0.98	1.05	1.07	1.25	1.01
	HICP	0.005	1.45	2.93	1.53	1.82	1.04	1.36
2.00	ULC	0.006	1.09	1.68	1.18	1.04	1.11	0.84 *
	Nominal XR	0.032	0.95 **	0.93 ***	1.05	0.95	1.19	0.97 ***
	HICP	0.008	1.08	2.46	1.11	1.44	0.82 **	0.83 *
4.00	ULC	0.012	0.86 *	1.97	1.03	1.23	1.02	0.71 *
	Nominal XR	0.057	0.90 *	0.89 *	0.94 **	0.89 **	1.06	0.87 *
	HICP	0.010	2.25	5.20	1.65	2.37	1.52	1.07
8.00	ULC	0.020	0.84 *	2.35	0.82 *	1.09	1.56	0.51 *
	Nominal XR	0.096	0.87 *	0.83 *	0.84 *	0.83 *	0.86 ***	0.81 **
				Crisis - 2008:4-2012:4				
	HICP	0.005	1.68	1.96	1.67	1.96	1.49	1.41
1.00	ULC	0.008	1.01	0.78 ***	1.12	0.78 ***	0.97	0.97
	Nominal XR	0.025	1.03	0.99	0.92 ***	0.99	0.95	0.94 ***
	HICP	0.009	0.80 ***	1.28	0.84 ***	1.28	0.86 *	0.82 *
2.00	ULC	0.014	0.71 ***	0.47 **	0.75 ***	0.47 **	0.63 ***	0.63 ***
	Nominal XR	0.040	1.02	1.13	1.00	1.13	1.05	1.01
	HICP	0.018	0.76 **	1.10	0.84 **	1.10	0.72 *	0.69 *
4.00	ULC	0.028	0.38 ***	0.26 **	0.40 ***	0.26 **	0.37 ***	0.38 ***
	Nominal XR	0.055	1.17	1.22	1.13	1.22	1.03	1.02
	HICP	0.033	0.84 *	1.06	0.91 *	1.01	0.66 *	0.66 *
8.00	ULC	0.056	0.19 *	0.25 *	0.19 *	0.33 *	0.18 *	0.19 *
	Nominal XR	0.059	1.67	1.62	1.70	1.60	1.45	1.42

Notes: The FECM and the FAVAR contain 4 factors extracted from data in levels. FECM$_{BN}$ uses factors extracted from differences. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1975:1 - 2012:4, forecasting: 2002:1 - 2012:4. *,** and *** indicate the significance at 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to AR model. Variables pre-selected as in Boivin and Ng (2006) using 0.75 threshold for the correlation coefficient.

In the crisis period the AR model remains the best performing in 4 cases. Similarly to the period before the crisis, the FAR model outperforms the AR model in half of the cases and is best overall in one. The performance of the FAVAR model that uses the information from large panels in a system of variables improves considerably relative to the AR, outperforming it in 7 out 12 cases. The ECM, exploiting the error-correction mechanism also improves, as it outperforms the AR model in 5 instances (2 before the crisis). The FECM that incorporates both the information from large datasets and cointegration exhibits the largest improvement. It outperforms the AR in 8 out of 12 cases (2 before the crisis) and is best in 3 cases (only once before the crisis). Some of the gains in forecasting precision relative to the AR are significant: above 30% and above 80% for the HICP and for unit labor costs, respectively, at 2-year horizon.

4.3 Robustness check to I(1) idiosyncratic errors

As we argued above, in the Euro Area dataset we cannot reject the hypothesis that the idiosyncratic components of the data are I(0). Nevertheless, we also assess the forecasting performance of the FECM with the factors estimated from the data in differences, and cumulated to obtain the estimate of the space spanned by I(1) and I(0) factors. If our primary assumption of I(0) idiosyncratic components was violated, then the estimated factor space from the data in levels as in Bai (2004) would be incosistent. Estimating the factors from differences would in such a case provide consistent estimates of the factors space, and should consequently also improve the forecasting performance.

In Tables 1 to 6, the relevant results are in the last columns, labeled $FECM_{BN}$. In general the results are fairly robust to the factor estimation method, justifying the initial assumption of I(0) idiosyncratic errors. Across the tables there is no systematic indication that the $FECM_{BN}$ model would either outperform the $FECM$ model or be inferior to it. Moreover, the relative performance with respect to other competing models is also virtually unchanged.

Table 7: **Forecasting US real variables, evaluation period 1970 - 1998**

		RMSE		**MSE relative to MSE of AR model**				
h	**Variable**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECM$_{BN}$**
	PI	0.01	1.01	0.94 ***	0.92 *	0.91 *	0.90 *	0.93 *
1.00	ManTr	0.01	1.03	0.97 *	0.93 *	1.01	0.97 *	0.96 *
	IP	0.01	0.99 *	1.08	0.94 *	1.01	0.99 *	0.98 *
	Empl	0.00	1.09	1.34	1.18	1.27	1.17	1.24
	PI	0.01	1.02	0.91 *	0.87 *	0.91 *	0.80 *	0.87 *
3.00	ManTr	0.02	1.01	1.01	0.95 *	1.09	0.90 *	0.96 *
	IP	0.02	0.95 *	1.03	0.93 *	1.02	0.98 *	0.95 *
	Empl	0.00	1.13	1.51	1.39	1.47	1.26	1.35
	PI	0.02	1.00	0.93 *	0.92 *	0.97 **	0.88 *	0.94 *
6.00	ManTr	0.03	1.01	1.01	0.97 **	1.10	0.90 *	0.98 *
	IP	0.03	0.96 *	0.99	0.95 *	1.00	1.03	0.95 *
	Empl	0.01	1.10	1.34	1.32	1.37	1.30	1.31
	PI	0.03	0.99 ***	0.96 *	0.96 *	0.98 ***	0.91 *	1.01
12.00	ManTr	0.05	1.01	0.99 *	0.97 *	1.02	0.80 *	0.96 *
	IP	0.05	0.97 **	0.99 ***	0.98 ***	0.96 *	0.98 *	0.99 ***
	Empl	0.02	1.02	1.10	1.12	1.16	1.17	1.20
	PI	0.04	1.00	0.97 *	0.98 *	1.03	0.97 **	1.06
18.00	ManTr	0.06	1.00	0.99 ***	0.99 ***	1.04	0.78 *	0.98 ***
	IP	0.06	0.99	0.99 ***	1.00	1.02	0.99 *	1.03
	Empl	0.03	0.96 ***	0.99 ***	1.00	1.07	1.09	1.12
	PI	0.04	1.01	0.98 *	1.00	1.02	0.94 *	1.09
24.00	ManTr	0.07	1.00	0.99	1.01	0.98 *	0.70 *	1.00
	IP	0.08	1.00	0.98 ***	1.00	1.02	0.89 *	1.05
	Empl	0.04	0.91 *	0.91 **	0.92 **	0.97 **	0.95 *	1.06
		AR	0.99	0.66	1.81	3.15		
	Lags	FAR	1.94	1.85	1.94	1.94		
		VAR	FAVAR	ECM	FECM	FECM$_{BN}$		
		1.33	1.00	1.00	1.00	1.00		
Cointegration rank								
	ECM			FECM	FECM$_{BN}$			
mean	min	max	mean	min	max	mean	min	max
3.60	2.00	4.00	3.87	1.17	4.00	4.00	4.00	4.67

Notes: The FECM contains 4 I(1) factors, while an additional I(0) factor is added to the FECMc. The FAVAR includes 5 I(0) factors. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1960:1 - 1998:12, forecasting: 1970:1 - 1998:12.
Variables: IP - Industrial production, PI - Personal income less transfers, Empl - Employees on non-aggr. payrolls, ManTr - Real manufacturing trade and sales
*,** and *** indicate the significance 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to the AR model.

The Euro Area dataset contains only 32 I(1) variables, which makes it easier to satisfy the I(0) idiosyncratic component assumption. For this reason we provide another robustness check using a considerably wider panel. In particular, we take the example of forecasting US monthly real and nominal macroeconomic series from Banerjee, Marcellino and Masten (2014a) and augment them with the $FECM_{BN}$ model. The dataset contains over 100 macroeconomic series over the period 1959 - 2003. As in the basic example of Banerjee, Marcellino and Masten (2014a), we consider forecasting total industrial production (IP), personal income less transfers (PI), employment on non-agricultural payrolls (Empl), and real manufacturing trade and sales (ManTr) as real variables. As nominal variables we consider the producer price index (PPI), consumer price index (CPI), consumer prices without food prices (CPI no food) and private consumption deflator (PCE). We forecast recursively over the period 1970 - 1998, which is the same as in Stock and Watson (2002b).

In essence, the results in Tables 7 and *8* provide evidence of the FECM as an efficacious forecasting model, providing substantial gains in a large number of cases. For real variables, the FECM is the best model in 11 out of 24 cases. Gains in forecasting precision tend to increase with the forecast horizon and can reach levels close to or even above 20%. The FAVAR model is best in only 4 out of 24 cases, while the ECM produces the lowest MSE only once.

Table 8: **Forecasting US nominal variables, evaluation period 1970 - 1998**

h	**Log of**	**RMSE of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**$FECM_{BN}$**
			MSE relative to MSE of AR model					
	PPI	0.005	1.03	1.05	1.03	0.90 *	0.90 *	1.08
1.00	CPI all	0.002	1.04	1.01	1.09	0.95 *	0.86 *	1.08
	CPI no food	0.002	0.99 *	0.94 *	0.99 *	0.91 *	0.93 *	1.08
	PCE defl	0.002	1.04	0.97 *	1.04	1.03	0.92 *	1.10
	PPI	0.005	1.13	1.12	1.14	0.89 *	0.93 *	1.13
3.00	CPI all	0.003	1.09	1.08	1.13	1.06	0.82 *	1.16
	CPI no food	0.003	1.02	1.01	1.05	0.98 *	0.90 *	1.24
	PCE defl	0.002	1.13	1.12	1.16	1.39	1.18	1.36
	PPI	0.005	1.15	1.14	1.20	1.03	0.97 *	1.24
6.00	CPI all	0.003	1.19	1.17	1.22	1.35	1.01	1.33
	CPI no food	0.003	1.04	1.02	1.06	1.13	0.97 *	1.29
	PCE defl	0.002	1.12	1.10	1.14	1.67	1.25	1.38
	PPI	0.005	1.11	1.11	1.16	0.93 *	0.91 *	1.17
12.00	CPI all	0.003	1.06	1.06	1.09	1.16	0.84 *	1.14
	CPI no food	0.003	1.02	1.01	1.04	1.00	0.86 *	1.15
	PCE defl	0.002	1.06	1.05	1.08	1.41	0.95 *	1.18
	PPI	0.006	1.08	1.07	1.13	0.95 **	0.96 *	1.17
18.00	CPI all	0.003	1.05	1.04	1.08	1.06	0.87 *	1.11
	CPI no food	0.004	1.02	1.01	1.04	0.99 *	0.94 *	1.19
	PCE defl	0.003	1.05	1.04	1.07	1.22	0.97 **	1.18
	PPI	0.006	1.11	1.12	1.16	0.76 *	0.84 *	1.17
24.00	CPI all	0.004	1.10	1.09	1.12	0.84 *	0.81 *	1.11
	CPI no food	0.004	1.04	1.02	1.05	0.78 *	0.84 *	1.13
	PCE defl	0.003	1.09	1.07	1.10	1.03	0.86 **	1.20
		AR	5.10	4.70	4.38	5.12		
	Lags	FAR	2.00	1.99	2.00	1.95		
		VAR	FAVAR	ECM	FECM	$FECM_{BN}$		
		2.53	1.70	0.00	0.00	0.00		
				Cointegration rank				
	ECM			FECM	$FECM_{BN}$			
mean	min	max	mean	min	max	mean	min	max
4.00	4.00	4.00	4.00	4.00	4.50	4.00	4.00	4.67

Notes: The FECM contains 4 I(1) factors, while an additional I(0) factor is added to the FECMc. The FAVAR includes 5 I(0) factors. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1960:1 - 1998:12, forecasting: 1970:1 - 1998:12.
Variables: Inflations of producer price index (PPI), consumer price index of all items (CPI all), consumer price index less food (CPI no food) and personal consumption deflator (PCE defl)
*,** and *** indicate the significance 10, 5, and 1 percent respectively of the Clark and West (2007) test of equal predictive accuracy relative to the AR model.

For nominal variables the results are presented in Table 8. Also in this case a good forecasting performance of the FECM is confirmed. The FECM model with factors extracted from the levels of variables turns out to be the best performing in 15 out of 24 cases reported in the table. In addition, it outperforms the benchmark AR model in 22 (out of 24) cases. The FAVAR is never the best performing model and improves over the benchmark model AR model only once. The ECM is the second best performing model outperforming the AR in 11 instances, of which 5 are the best overall.

The last columns of Tables 7 and 8 illustrate the effect of extracting the factor from differences of the data. In general, they offer similar findings about the effects of factors extraction from the differences in the FECM as in the case of the Euro Area data. In the majority of cases of forecasting real variables, the relative MSEs of the $FECM_{BN}$ are close to those of the FECM model, while on average they are higher. This again confirms that, from the point of view of forecasting precision, extraction of factors from levels of the data provides valid results. Such a conclusion can be derived also from the last column of Table 8. In the case of nominal variables the $FECM_{BN}$ performs even worse. It's relative MSEs are consistently above one. The FECM using factors extracted from levels of data thus performs consistently better than the one with factors estimated from differences.

5 Monetary policy shocks in the FECM

The first analysis of monetary policy shocks in large panels, based on a FAVAR model, was developed by Bernanke, Boivin and Eliasz (2005, BBE). The essence of their approach is in the division of variables into two blocks: slow-moving variables that do not respond contemporaneously to monetary policy shocks and fast-moving variables that do. In addition, BBE treat the policy instrument variable, the federal funds rate, as one of the observed factors. They consider two estimation methods, namely Bayesian estimation and principal components analysis. In the latter approach, most frequently used in the literature and in practice, they estimate K factors from the whole panel and from the subset of slow-moving variables only (slow factors). They then rotate the factors estimated from the whole panel around the federal funds rate by means of a regression of these factors on the slow-factors and the federal funds rate. As a result of this rotation of the factors, the analysis proceeds with $K + 1$ factors, namely the K rotated estimated factors and the federal funds rate imposed as an observable factor.

Identification of monetary policy shocks is obtained in the VAR model of rotated factors assuming a recursive ordering with the federal funds rate ordered last.

$$E(\varphi_t \varphi_t') = H\Sigma_{u,w}H' = I, \tag{22}$$

where H^{-1} is lower triangular. The impulse responses of the observed variables of the panel are then estimated by multiplying the impulse responses of the factors by the loadings obtained from OLS regressions of the variables on the rotated factors.

The identification scheme for the analysis of monetary policy shocks can be easily adapted to the FECM, which enables us to study the role of error-correction mechanism in propagation of monetary policy shocks. We need to introduce one modification that makes the results obtained with the FECM directly comparable to those of the FAVAR. The difference is at the stage of factor estimation. Namely, in order to capture cointegration as in Bai (2004), we estimate the factors from the data in levels, while BBE estimate the factors from data transformed (if necessary) to I(0).[12] This gives us the estimates of the space spanned by r_1 I(1) factors and $r-r_1$ stationary factors. As in BBE, the federal funds rate is treated as one observable factor and the estimated factors are rotated accordingly. Because their method entails identifying the monetary policy shocks from a stationary factor VAR, the first r_1 nonstationary factors are differenced. Identification of monetary policy shocks is then obtained from a VAR of stationary factors.

The dataset is, for comparability, the same as in BBE. It contains 120 variables for the US, spanning over the period 1959 - 2003 at monthly frequency. 77 variables are by the authors treated as I(1). Bai's (2004) IPC2 information criteria indicates $r_1 = 2$. In the choice of the total number of estimated factors r we follow Bernanke et al. (2005) and set it to 3. However, as in their case, the main findings are robust to working with more factors. Including the federal funds rate, the total number of factors is 4.

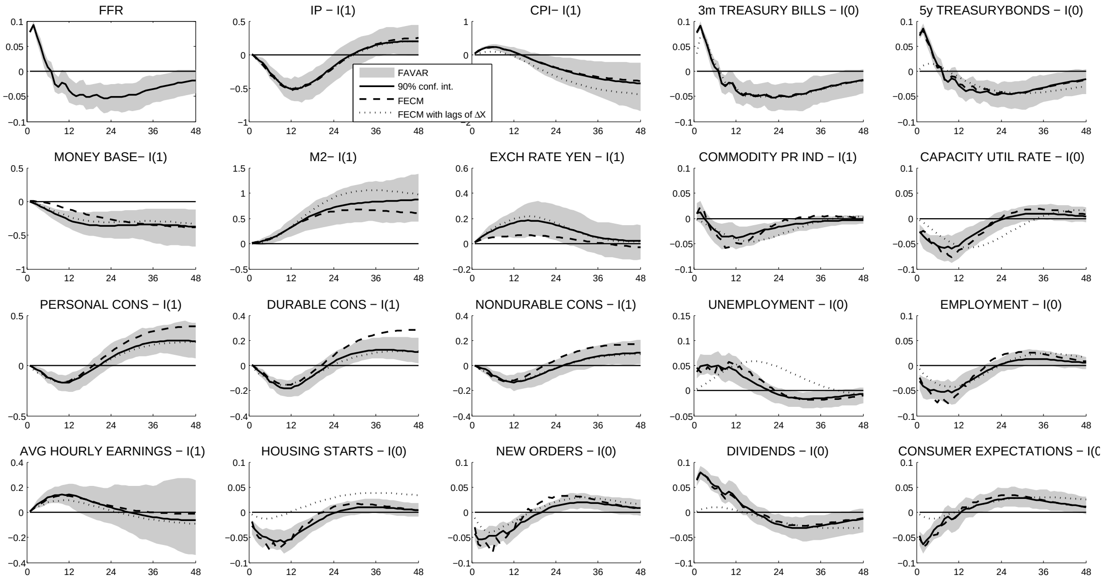
The basic results are presented in Figure 1. It contains the impulse responses for the same set of variables as in Bernanke et al. (2005) obtained from the conventional FAVAR model and the FECM model. In line with the lag structure chosen by Bernanke et al. (2005) we only include contemporaneous values of factors in equations (18)-(21). They differ in the presence of the error-correction term for the variables that are treated as I(1) in levels. Some variables are assumed to be I(0). These are the interest rates, the capacity utilization rate, unemployment rate, employment, housing starts, new orders and consumer expectations. For these variables the FAVAR and the FECM also differ. Consistent with (19), the FECM for I(0) variables excludes the I(1) factors. As a robustness check, in the figure we additionally plot impulse responses obtained with a more general FECM specification in which 6 lags of ΔX_{it} are added to the model equations.[13]

What we observe is coherence in terms of the basic shape of the impulse responses between the models. Quantitatively, however, the responses may differ significantly due to the error-correction terms. The responses of the industrial production, the CPI and wages are very similar. Quite significant differences are observed for money and the yen-dollar exchange rate. The same is true for measures of consumption. Omission of the error-correction terms in the FAVAR model can thus have an important impact on the empirical results. It is worth mentioning that these differences are observed conditional

[12] Both approaches deliver similar estimates of monetary policy shocks. However, as the factors are estimated on datasets of different order of integration, they are not numerically identical.

[13] The basic shapes of impulse responses are robust to the specification of endogenous lags and lags of factors. Result available upon request.

Figure 1: Impulse responses to monetary policy shock - FAVAR Vs FECM with factors extracted from levels



upon a shock that accounts for only a limited share of variance. Banerjee, Marcellino and Masten (2014b) present an analysis of real stochastic trends, where the differences between FAVAR and FECM responses become even more pronounced, and the shock is a considerably more important source of stochastic variation in the panel.

The impulse responses of I(0) variables are very similar across models. This means that imposing the restriction that the differences of I(1) factors do not load to I(0) variables has only a limited quantitative impact, which is consistent with the FECM specification of the model. In the FECM the restriction is evident. In the FAVAR, which makes no distinction in the structure of the loadings of factors to I(1) and I(0) variables, such a restriction cannot be directly determined.

Finally, including lags of the endogenous variables in the FECM (green lines in Figure 1) confirms our basic findings that the omission of the error-correction term is the main source of differences in the impulse responses between the FAVAR and the FECM model.

6 Conclusions

The Factor Augmented Error Correction Model (FECM) offers two important advantages for empirical modelling. First, the factors proxy for missing cointegration information in a standard small scale ECM. Second, the error correction mechanism can also be inserted in the context of a large dataset. From a theoretical point of view, since the FECM nests both the FAVAR and the ECM, it can be expected to provide better empirical results, unless either the error correction terms or the factors are barely significant, or their associated coefficients are imprecisely estimated due to small sample size, or the underlying assumptions that guarantee consistent factor and parameter estimation are not satisfied.

In our forecasting application, the FECM is clearly the best forecasting model for Euro Area real variables, and a comparison with the FAVAR highlights the importance of including the error correction terms. For nominal variables, the FECM also performs well if the factors are extracted from a subset of variables pre-selected as in Bai and Ng (2006). Moreover, the performance of the FECM generally further improves during the crisis. Overall, these results are in line with those for the US reported in Banerjee et al. (2014a). We have also seen that the forecasting performance for the US and the Euro Area is not substantially affected if the factors are estimated from the variables in levels or in differences, with a better empirical performance in general of the former method, which suggests that the hypothesis of I(0) idiosyncratic errors is not stringent.

In terms of structural analysis, we have investigated the transmission of monetary shocks, comparing the responses of several variables with those from the FAVAR based analysis of Bernanke et al. (2005). The shape of the impulse responses is overall similar across the FECM and FAVAR models for most variables. Quantitatively, however, the responses may differ significantly due to the error-correction terms. For example, relevant differences are observed for monetary aggregates, the yen-dollar exchange rate, and measures of consumption. Omission of the error-correction terms in the FAVAR model can thus have an important impact on the empirical results.

Overall, our empirical results provide further compelling evidence that the FECM provides an important extension of classical ECM and FAVAR models both for forecasting and structural modelling. This finding, combined with the ease of estimation and use of the FECM model, suggests that it could be quite useful for empirical analyses.

References

[1] Adolfson, M, Laseen, S, Linde, J. and M. Villani (2007). Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics,* 72(2), 481-511.

[2] Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, **122**, 137-183.

[3] Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica*, **70**, 191-221.

[4] Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica*, **72**, 1127-1177.

[5] Banerjee, A. and M. Marcellino (2009). Factor-augmented error correction models, in J.L. Castle and N. Shephard, N. (eds.), *The Methodology and Practice of Econometrics – A Festschrift for David Hendry.* Oxford: Oxford University Press, 227-254.

[6] Banerjee, A., M. Marcellino and I. Masten (2014a). Forecasting with Factor-augmented Error Correction Models. *International Journal of Forecasting,* **30(3),** 589-612.

[7] Banerjee, A., M. Marcellino and I. Masten (2014b). Structural FECM: Cointegration in large-scale structural FAVAR models, CEPR Discussion paper No. 9858.

[8] Bernanke, B.S., J. Boivin and P. Eliasz (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, **120**, 387-422.

[9] Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132, 169-194.

[10] Cheng, X., & Phillips, P. C. B. (2009). Semiparametric cointegrating rank selection. *Econometrics Journal*, *12*(s1), S83-S104.

[11] Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291-311.

[12] Eickmeier, S. (2009), Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model. *Journal of Applied Econometrics*, **24(6)**, 933-959.

[13] Fagan, G., Henry, J. and Mestre, R. (2001), "An area-wide model for the Euro area", ECB Working Paper No. 42, European Central Bank

[14] Gengenbach, C., J-P. Urbain and J. Westerlund (2008). Panel error correction testing with global stochastic trends. METEOR Research Memorandum 51.

[15] Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, Oxford and New York.

[16] King, R. G., Plosser, C I., Stock, J. H. and M. W. Watson (2005). Stochastic Trends and Economic Fluctuations. *American Economic Review*, **81(4),** 819-40.

[17] Lütkepohl, H. (2014). Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey. DIW Discussion paper 1351.

[18] Smets, Frank, and Rafael Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, **97(3)**: 586-606.

[19] Stock, J.H. and M.W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, **20**, 147-162.

[20] Stock, J.H. and M.W. Watson (2005). Implications of dynamic factor models for VAR analysis. NBER Working Paper 11467.

[21] Warne, A. (1993), "A Common Trends Model: Identification, Estimation and Inference". University of Stockholm, IIES Seminar Paper No. 555.