EUI Working Papers

RSCAS 2009/32
ROBERT SCHUMAN CENTRE FOR ADVANCED STUDIES

FORECASTING WITH FACTOR-AUGMENTED ERROR CORRECTION MODELS

Igor Masten, Anindya Banerjee, Massimiliano Marcellino

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
ROBERT SCHUMAN CENTRE FOR ADVANCED STUDIES

Forecasting with Factor-augmented Error Correction Models

IGOR MASTEN, ANINDYA BANERJEE, MASSIMILIANO MARCELLINO

EUI Working Paper **RSCAS** 2009/32

ISSN 1028-3625

Printed in July 2009
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu/RSCAS/Publications/
www.eui.eu
cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), directed by Stefano Bartolini since September 2006, is home to a large post-doctoral programme. Created in 1992, it aims to develop inter-disciplinary and comparative research and to promote work on the major issues facing the process of integration and European society.

The Centre hosts major research programmes and projects, and a range of working groups and ad hoc initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration and the expanding membership of the European Union.

Details of this and the other research of the Centre can be found on:
http://www.eui.eu/RSCAS/Research/

Research publications take the form of Working Papers, Policy Papers, Distinguished Lectures and books. Most of these are also available on the RSCAS website:
http://www.eui.eu/RSCAS/Publications/

Forecasting with Factor-augmented Error Correction Models*

Anindya Banerjee† Massimiliano Marcellino‡ Igor Masten§

This version: 10 June 2009

Abstract

As a generalization of the factor-augmented VAR (FAVAR) and of the Error Correction Model (ECM), Banerjee and Marcellino (2009) introduced the Factor-augmented Error Correction Model (FECM). The FECM combines error-correction, cointegration and dynamic factor models, and has several conceptual advantages over standard ECM and FAVAR models. In particular, it uses a larger dataset 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 examine the forecasting performance of the FECM by means of an analytical example, Monte Carlo simulations and several empirical applications. We show that relative to the FAVAR, FECM generally offers a higher forecasting precision and in general marks a very useful step forward for forecasting with large datasets.

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

JEL-Codes: C32, E17

*We are grateful to Helmut Luetkepohl and seminar participants at the EUI, University of Helsinki and the 6th WISE Workshop in Salerno for helpful comments on a previous draft. Responsibility for any errors remains with us.

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

‡European University Institute, Bocconi University and CEPR, Via della Piazzuola 43, 50133 Florence, Italy, e-mail: massimiliano.marcellino@eui.eu

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

1 Introduction

In Banerjee and Marcellino (2009), we introduced the Factor-augmented Error Correction Model (FECM). The main contribution of that paper was to bring together two important recent strands of the econometric literature on modelling co-movements that had a common origin but in their implementation had thus far remained largely apart, namely, cointegration and dynamic factor models. We focused on a theoretical framework that allowed for the introduction of cointegrating or long-run information explicitly into a dynamic factor model and evaluated the role of incorporating long-run information in modelling data, in particular in situations where the dataset available to researchers was potentially very large (as in the empirical illustrations described in Section 5 below.) We argued that the FECM, where the factors extracted from the large dataset are jointly modelled with a limited set of economic variables of interest, represented a manageable way of dealing with the problem posed by large datasets characterized by cointegration, where such cointegration needed in principle to be taken into account. A number of papers have emphasized, see for example Clements and Hendry (1995), the complexity of modelling large systems of equations in which the complete cointegrating space may be difficult to identify. Therefore, proxying for the missing cointegrating information by using factors could turn out to be extremely useful, and we proposed the use of the FECM as a potentially worthwhile approach with a wide range of applicability.

The discussion in Banerjee and Marcellino (2009) concentrated on first establishing a theoretical structure to describe the FECM and then illustrating its efficacy by the use of analytical examples, a simulation study and two empirical applications. Our model-comparisons were based mainly on in-sample measures of model fit, and we studied the improvements provided by FECMs with respect to a standard Error Correction Model (ECM) and Factor-Augmented VARs (FAVAR) such as those considered by Bernanke, Boivin and Eliasz (2005), Favero, Marcellino and Neglia (2005) and Stock and Watson (2005). We viewed the FECM as an improvement both over the ECM, by relaxing the dependence of cointegration analysis on a small set of variables, and over the FAVAR, by allowing for the inclusion of error correction terms into the equations for the key variables under analysis, preventing the errors from being non-invertible MA processes.

The focus of this paper is instead upon the evaluation of the forecasting performance of the FECM in comparison with the ECM and the FAVAR. In our view, establishing forecasting efficacy is an important further key to determining the considerable usefulness of the FECM as an econometric tool. As we show below, the relative rankings of the ECM, the FECM and the FAVAR depend upon 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, both the ECM and the FAVAR are outperformed by the FECM, given that it is a nesting specification.

We start in Section 2 by reviewing the theoretical background of our study, by describing the FECM and comparing it with the ECM and the FAVAR.

Section 3 offers a simple yet comprehensive analytical example to understand the features which are likely to determine the rankings - in terms of forecasting accuracy - of these three models.

Section 4 presents two Monte Carlo designs to illustrate the effectiveness of the different models in providing forecasts. The first design is based on the simple analytical model of Section 3. The second design is more elaborate and mimics one of the estimated models in the empirical examples given in Section 5. We can anticipate that the results of the Monte Carlo show that the strength of error correction alongwith the lengths of the cross-section (N) and time dimension (T) matter greatly in determining the forecast ranking of alternative models. However, in the majority of cases the FECM performs well, and systematically better than the FAVAR.

Section 5 carries the analysis to the practical realm. Forecasting with ECMs and with factor models has attracted considerable attention, see e.g., respectively, Clements and Hendry (1995) and Eickmeier and Ziegler (2008). To provide a thorough comparison of the ECM, FAVAR and FECM, we consider four main applications, and we describe them briefly in turn below.

Stock and Watson (2002b) focused on forecasting a set of four real variables (total industrial production, personal income less transfers, employment on non-agricultural payrolls and real manufacturing trade and sales) and a set of four nominal variables (inflation of producer prices of finished goods, CPI inflation with all items included, CPI inflation less food and the growth of the personal consumption expenditure deflator) for the United States. They compared the performance of factor models, ARs and VARs, typically finding gains from the use of factor models. Since the four variables in each set represent strongly related economic phenomena, it is logical to expect that they are cointegrated. Hence, in this context the FECM represents a natural econometric specification.

As a second application, we focus on a small monetary system consisting of one real, one nominal and one financial variable, in common with standard practice in this literature, see e.g. Rudebusch and Svensson (1998). Favero et al. (2005), among others, considered augmenting this model with factors extracted from a large dataset to assess the effects on estimation and shock transmission. Here we are more interested in forecasting, and in the role of cointegration among the basic variables, and them and the factors. The VAR, FECM and FAVAR models are estimated first for the United States, and then for Germany, the largest country in the euro area, for which much shorter time series are available due to unification.

The third application concerns the term structure of interest rates. A standard model for these variables assumes that they are driven by three factors, the intercept, slope and curvature, see e.g. Dieblod and Li (2006). Hence, there should be a large amount of cointegration among them, in line with the findings by Hall, Anderson and Granger (1992). Therefore, the FECM should be particularly suited in this context.

The fourth and final application deals with exchange rate forecasting. The empirical analysis by Meese and Rogoff (1983) and the theoretical results by Engel and West (2005), among others in this vast literature, point to the difficulties in beating a random walk or simple AR forecast. However, Carriero, Kapetanios and Marcellino (2009b), show that cross-sectional information can be useful, but factor models on their own do not appear to work very well in forecasting. Since this poor performance could be due to the omission of information relating to cointegration, FECMs are the obvious candidates to also try in this framework.

It is helpful to highlight here the key results of this extensive empirical analysis. First, for real variables for the US, the FECM is systematically better than the FAVAR and the ECM. Second, for the nominal US variables, an adaptation, denoted FECMc, to be discussed below, or the ECM are in general the preferred models (depending upon the time coverage and span of the datasets). Third, in the small monetary system for the US, the FECM or FECMc is the dominant model, and the use of long-run information is crucial. Fourth, for the monetary model for Germany, while the FECM provides the best forecast in 6 out of 18 cases, the VAR is marginally the best performer (providing the best forecast in 8 out of 18 cases). This shows that accounting for cointegration and factors may not always be sufficient, although this finding is conditioned heavily on the relatively short estimation and evaluation periods for this example. Fifth, for the term structure of interest rates, the FECM and FECMc provide the best forecasts in a very large number of cases and the gains provided here by these models in relation to their competitors is frequently quite substantial. Finally, for exchange rates, the FECM is again by far the dominant method, with the use of cointegration and factors providing significant gains. Overall, these results emphasize the utility and robustness of FECM methods and shed light on the combined use of factors and cointegrating information.

To conclude, Section 6 provides a detailed summary of the main findings of the paper and suggests directions for additional research in this area.

2 The Factor-augmented Error Correction Model

It is helpful to begin with a brief description of the main theoretical structure underlying the analysis. The discussion in this section is derived from Banerjee and Marcellino (2009) and is useful in setting out the representation of the FECM and its relation to the ECM and the FAVAR.

Consider a set of N $I(1)$ variables x_t which evolve according to the $VAR(p)$ model

$$x_t = \Pi_1 x_{t-1} + ... + \Pi_p x_{t-p} + \epsilon_t, \tag{1}$$

where ϵ_t is $i.i.d.(0, \Omega)$ and the starting values are fixed and set equal to zero for simplicity and without any essential loss of generality. Following Johansen (1995, p.49), the $VAR(p)$

can be reparameterized into the Error Correction Model (ECM)

$$\Delta x_t = \alpha\beta^{'} x_{t-1} + \nu_t, \tag{2}$$

or into the so-called common trend specification

$$x_t = \Psi f_t + u_t. \tag{3}$$

In particular, under these specifications,

$$\Pi = \sum_{s=1}^{p} \Pi_s - I_n = \underset{N\times N-r}{\alpha} \underset{N-r\times N}{\beta^{'}},$$

$$v_t = \Gamma_1 \Delta x_{t-1} + ... + \Gamma_{p-1}\Delta x_{t-p+1} + \epsilon_t, \quad \Gamma_i = -\sum_{s=i+1}^{p} \Pi_s, \quad \Gamma = I - \sum_{i=1}^{p-1} \Gamma_i,$$

$$\underset{N\times r}{\Psi} = \beta_\perp (\alpha_\perp^{'} \Gamma \beta_\perp)^{-1}, \quad \underset{r\times 1}{f_t} = \alpha_\perp^{'} \sum_{s=1}^{t} \epsilon_s, \quad u_t = C(L)\epsilon_t.$$

β' is the $N - r \times N$ matrix of cointegrating vectors with rank $N - r$, where $N - r$ is the number of cointegrating vectors. From this it follows that r is the number of $I(1)$ common stochastic trends (or factors), $0 < r \leq N$, gathered in the $r \times 1$ vector f_t and the matrix $\alpha_\perp^{'}\Gamma\beta_\perp$ is invertible since each variable is $I(1)$. α is the so-called loading matrix, which also has reduced rank $N - r$ and determines how the cointegrating vectors enter into each individual element $x_{i,t}$ of the $N \times 1$ vector x_t.[1] u_t is an $N-$dimensional vector of stationary (and in general, moving average) errors.

We also assume here that there are no common cycles in the sense of Engle and Kozicki (1993), i.e., no linear combinations of the first differences of the variables that are correlated of lower order than each of the variables (in first differences). However, adding such cycles poses no significant theoretical complications and is assumed here only for convenience.[2] Indeed, in the empirical applications in Section 5, we also consider a modification of the FECM, denoted FECMc, consisting of the FECM augmented with common factors extracted from the stationary component of x_t in (3) after the $I(1)$ factors f_t and their corresponding loadings have been estimated. This is because, unlike in a theoretical framework, where these features may be imposed by assumption, it is not possible in empirical examples to rule these out *a priori* . It is therefore of interest to allow for common cycles in the residuals to judge if this makes a difference as far as forecasting performance is concerned.

[1] Note that as $N \to \infty$, and the number of factors r remains fixed, the number of cointegrating relations $N - r \to \infty$.

[2] Common cycles are associated with reduced rank of (some of) the coefficient matrices in $C(L)$, where we remember that the errors in the stochastic trend representation (3) are $u_t = C(L)\epsilon_t$. Therefore, the presence of common cycles is associated with stationary common factors driving x_t, in addition to the I(1) factors.

From equation (3), it is possible to write the model for the first differences of x_t, Δx_t, as

$$\Delta x_t = \Psi \Delta f_t + \Delta u_t, \tag{4}$$

where Δu_t and ν_t can be correlated over time and across variables.

Papers on dynamic factor models (DFM) such as Stock and Watson (2002a,b) and Forni, Hallin,Lippi and Reichlin (2000) have relied on a specification similar to (4) and have focused on the properties of the estimators of the common factors Δf_t, or of the common components $\Psi \Delta f_t$, under certain assumptions on the idiosyncratic errors, when the number of variables N becomes large. A few papers have also analyzed the model in (3) for the divergent N case, most notably Bai and Ng (2004) and Bai (2004).[3]

By contrast, the literature on cointegration has focused on (2), the so-called error correction model (ECM), and studied the properties of tests for the cointegrating rank $(N - r)$ and estimators of the cointegrating vectors $(\beta^{'})$, see e.g. Engle and Granger (1987) or Johansen (1995).

We shall make use of both specifications (3) and (4) when discussing factor models in what follows, in order to explain the correspondence that exists between the two specifications and how this leads to the development of the FECM.

Imposing, without any loss of generality, the identifying condition[4]

$$\underset{N-r\times N}{\beta^{'}} = \left(\underset{N-r\times r}{\beta^{*'}} : \underset{N-r\times N-r}{I} \right),$$

and, from (3), partitioning u_t into

$$u_t = \begin{pmatrix} \underset{r\times 1}{u_{1t}} \\ \underset{N-r\times 1}{u_{2t}} \end{pmatrix},$$

the model for the error correction terms can be written as

$$\beta^{'} x_t = \beta^{'} u_t = \beta^{*'} u_{1t} + u_{2t}. \tag{5}$$

Note that in this model each of the $N - r$ error correction terms is driven by a common component that is a function of only r shocks, u_{1t}, and an idiosyncratic component, u_{2t}. It is possible to change the exact shocks that influence each error correction term by choosing different normalizations, but the decomposition of these terms into a common

[3] Bai and Ng (2004) also allow for the possibility that some elements of the idiosyncratic error u_t are I(1). We will not consider this case and assume instead that the variables under analysis are cointegrated, perhaps after pre-selection. We feel that this is a sensible assumption from an economic point of view, otherwise the variables could drift apart without any bound.

[4] This is standard practice in this literature, as also implemented by e.g. Clements and Hendry (1995, page 129, lines 1 - 5) and ensures that the transformation from the levels x_t which are $I(1)$ to $I(0)$-space (involving taking the cointegrated combinations and the differences of the $I(1)$ variables) is scale preserving.

component driven by r shocks and an idiosyncratic component remains unchanged. This also corresponds to the stochastic trend representation in (3), where the levels of the variables are driven by r common trends.

Next, suppose, as is commonly the case in empirical studies and forecasting exercises concerning the overall economy, we are interested in only a subset of the variables for which we have information. We therefore proceed by partitioning the N variables in x_t into the N_A of major interest, x_{At}, and the $N_B = N - N_A$ remaining ones, x_{Bt}. A corresponding partition of the common trends model in (3) may be constructed accordingly as

$$\begin{pmatrix} x_{At} \\ x_{Bt} \end{pmatrix} = \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix} f_t + \begin{pmatrix} u_{At} \\ u_{Bt} \end{pmatrix}, \tag{6}$$

where Ψ_A is of dimension $N_A \times r$ and Ψ_B is $N_B \times r$. It is important to note that when the number of variables N increases, the dimension of Ψ_A remains fixed, while the number of rows of Ψ_B increases with the increase in N. Therefore, for (6) to preserve a factor structure asymptotically, driven by r common factors, it is necessary that the rank of Ψ_B remains equal to r. Instead, the rank of Ψ_A can be smaller than r, i.e., x_{At} can be driven by a smaller number of trends, say $r_A \leq r$.

From the specification in (6), it is may be seen that x_{At} and f_t are cointegrated, while the f_t are uncorrelated random walks. Therefore, from the Granger representation theorem, there exists an error correction specification of the form

$$\begin{pmatrix} \Delta x_{At} \\ \Delta f_t \end{pmatrix} = \begin{pmatrix} \gamma_A \\ \gamma_B \end{pmatrix} \delta' \begin{pmatrix} x_{At-1} \\ f_{t-1} \end{pmatrix} + \begin{pmatrix} e_{At} \\ e_t \end{pmatrix}. \tag{7}$$

Since, in practice, the correlation in the errors of (7) is handled by adding additional lags of the differenced dependent variables, the expanded model becomes

$$\begin{pmatrix} \Delta x_{At} \\ \Delta f_t \end{pmatrix} = \begin{pmatrix} \gamma_A \\ \gamma_B \end{pmatrix} \delta' \begin{pmatrix} x_{At-1} \\ f_{t-1} \end{pmatrix} + A_1 \begin{pmatrix} \Delta x_{At-1} \\ \Delta f_{t-1} \end{pmatrix} + ... + A_q \begin{pmatrix} \Delta x_{At-q} \\ \Delta f_{t-q} \end{pmatrix} + \begin{pmatrix} \epsilon_{At} \\ \epsilon_t \end{pmatrix}, \tag{8}$$

where the errors $(\epsilon'_{At}, \epsilon'_t)'$ are $i.i.d.$

The model given by (8) is labelled by Banerjee and Marcellino (2009) as the Factor-augmented Error Correction Model (FECM).

The important feature to note is that there are $N_A + r$ dependent variables in the FECM (8). Since x_{At} is driven by f_t or a subset of them, and the f_t are uncorrelated random walks, there must be N_A cointegrating relationships in (8). Moreover, since Ψ_A is of dimension $N_A \times r$ but can have reduced rank r_A, there are $N_A - r_A$ cointegrating relationships that involve the x_A variables only, say $\delta_A^{'} x_{At-1}$, and the remaining r_A cointegrating relationships involve x_A and the factors f_t.

The cointegrating relationships $\delta_A^{'} x_{At-1}$ would also emerge in a standard ECM for

Δx_{At} only, say

$$\Delta x_{At} = \alpha_A \delta_A^{'} x_{At-1} + v_{At}. \tag{9}$$

However, in addition to these $N_A - r_A$ relationships, in the FECM there are r_A cointegrating relationships that involve x_{At} and f_t, and that proxy for the potentially omitted $N - N_A$ cointegrating relationships in (9) with respect to the equations for Δx_{At} in the full ECM in (2).[5] Moreover, in the FECM there appear lags of Δf_t as regressors in the equations for Δx_{At}, that proxy for the potentially omitted lags of Δx_{Bt} in the standard ECM for Δx_{At} in (9).

The key to understanding the FECM is to see how use is made of the information contained in the unmodelled $N - N_A$ cointegrating relationships which are proxied by the cointegrating relationships between the variables of interest and the factors. Since, with increasing N, this cointegrating information is in principle quite large, its importance in relation to the variables of interest will determine the forecasting performance of the FECM when compared to a standard ECM or a FAVAR (which would not take any cointegrating information into account.)

To continue with this argument further, we see that the FAVAR specification follows easily from (8) by imposing the restrictions $\gamma_A = \gamma_B = 0$ thereby losing all long-run information. The VAR and the standard ECM also emerge as nested cases (by imposing suitable restrictions.) As we show below, this nesting property of the FECM is extremely useful for analyzing its performance. It is true, to be sure, that the theoretical advantages are not necessarily reflected in better forecasts in actual situations, but serve nevertheless as a guide.

To conclude the discussion in this section, we may make two further observations. First, we should note that when the Data Generating Process is the common trends specification in (3), the error process Δu_t in (4) may have a non-invertible moving average component that prevents the approximation of each equation of the model in (4) with an AR model augmented with lags of the factors. Second, and perhaps even more problematic, in (4) Δf_t and Δu_t are in general not orthogonal to each other, and in fact they can be highly correlated. This feature disrupts the factor structure and, from an empirical point of view, can require a large number of factors to summarize the information contained in Δx_t. Even when orthogonality holds, the presence of the first problem still makes the use of FAVAR models problematic.

3 An analytical example

We illustrate analytically the forecasting properties of the FECM relative to the FAVAR and the ECM with a simple but comprehensive example. The example may easily be seen

[5] In the full ECM model (2), there would be up to $N - r_A$ cointegrating relationships in the equations for Δx_{At}, while in (9) there are only $N_A - r_A$ cointegrating relationships, so that there are $N - N_A$ potentially omitted long run relationships in the ECM for Δx_{At} only.

to be a special case of the data generation processes given above, obtained by restricting the dimension of the factor space and of the variables of interest studied.

We suppose that the large information set available for forecasting may be summarized by one ($I(1)$) common factor, f, that the econometrician is particularly interested in forecasting one of the many variables, x_1, and that she can choose any of the three following models. First, a standard ECM for x_1 and x_2, where x_2 is a proxy for f. Second, a FAVAR model where the change in x_1 (Δx_1) is explained by its own lags and by lags of the change in f. And, third, a FECM, where the explanatory variables of the FAVAR are augmented with a term representing the (lagged) deviation from the long run equilibrium of x_1 and f. We want to compare the mean squared forecast error (MSE) for Δx_1 resulting from the three models, under different assumptions on the data generating process (DGP), and show that the FECM can be expected to perform at least as well as the FAVAR in all cases.

To start with, let us consider a system consisting of the two variables x_1 and x_2 and of one factor f. The factor follows a random walk process,

$$f_t = f_{t-1} + \varepsilon_t. \tag{10}$$

The factor loads directly on x_2,

$$x_{2t} = f_t + u_t, \tag{11}$$

while the process for x_1 is given in ECM form as

$$\Delta x_{1t} = \alpha \left(x_{1t-1} - \beta f_{t-1}\right) + \gamma \Delta f_{t-1} + v_t, \quad \alpha < 0. \tag{12}$$

Here the processes ϵ_t and v_t are assumed $i.i.d.(0, I_N)$, while u_t is allowed to have a moving average structure, i.e. $u_t = u_t^* / (1 - \eta L)$, $|\eta| < 1$ and u_t^* is $i.i.d.$ Hence, the DGP is a FECM.

Let us focus on Δx_{1t} and derive the (one-step ahead) MSE when the forecast is based on an ECM for x_1 and x_2 rather than on the FECM. Substituting (11) into (12) gives

$$\Delta x_{1t} = \alpha \left(x_{1t-1} - \beta x_{2t-1}\right) + \gamma \Delta x_{2t-1} + v_t + \alpha\beta u_{t-1} - \gamma \Delta u_{t-1},$$

so that

$$MSE_{ECM} = Var\left(v_t + \alpha\beta u_{t-1} - \gamma \Delta u_{t-1}\right).$$

It then follows that

$$\begin{aligned} MSE_{ECM} - MSE_{FECM} &= Var\left(\alpha\beta u_{t-1} - \gamma \Delta u_{t-1}\right) \\ &= \frac{\left(\alpha\beta - \gamma\right)^2 + \gamma^2}{1 - \eta^2} \sigma_{u^*}^2 > 0. \end{aligned}$$

To assess the role of cointegration, we can evaluate how this MSE difference changes with the strength of the error-correction mechanism. We have that

$$\frac{\partial\left(MSE_{ECM}-MSE_{FECM}\right)}{\partial\alpha}\propto\alpha\beta^{2},$$

where $\propto$ indicates "proportional to". Given that for the system to be error correcting we need $\alpha<0$, the loss of forecasting precision of the ECM relative to the FECM unambiguously increases with the strength of error correction (i.e. when α decreases). Similarly,

$$\frac{\partial\left(MSE_{ECM}-MSE_{FECM}\right)}{\partial\gamma}\propto 4\gamma,$$

so that the larger γ the larger the loss from approximating f with x_2.

The FECM representation of x_1 can also be written as a FAVAR. In fact, since the error-correction term $x_{1t}-\beta f_t$ evolves as

$$\begin{aligned}x_{1t}-\beta f_t &= (\alpha+1)\left(x_{1t-1}-\beta f_{t-1}\right)+\gamma\Delta f_{t-1}+v_t-\beta\varepsilon_t\\ &= \frac{\gamma\Delta f_{t-1}}{1-(\alpha+1)L}+\frac{v_t-\beta\varepsilon_t}{1-(\alpha+1)L},\end{aligned}$$

we can re-write equation (12) as

$$\Delta x_{1t}=\gamma\Delta f_{t-1}+\frac{\alpha\gamma\Delta f_{t-2}}{1-(\alpha+1)L}+v_t+\frac{\alpha\left(v_{t-1}-\beta\varepsilon_{t-1}\right)}{1-(\alpha+1)L} \tag{13}$$

This implies that

$$MSE_{FAVAR}-MSE_{FECM}=\alpha^{2}var\left(\frac{\left(v_{t-1}-\beta\varepsilon_{t-1}\right)}{1-(\alpha+1)L}\right),$$

so that $MSE_{FAVAR}>MSE_{FECM}$ whenever we have cointegration ($\alpha\neq 0$).

If instead $\alpha=0$, so that the DGP becomes a FAVAR rather than a FECM, the FECM and FAVAR become equivalent, and the gains in forecasting precision with respect to the ECM remain positive but shrink to $2\gamma^2/\left(1-\eta^2\right)\sigma^2_{v^*}$.

Finally, we consider the case where the DGP is an ECM instead of a FECM. This returns to the issue highlighted previously of the importance of the cointegrating relationships between the variables of interest and the factors. To illustrate this situation, we consider the same example as above but invert the role of x_2 and f in (10)-(12). Hence, the DGP becomes

$$x_{2t}=x_{2t-1}+\varepsilon_t. \tag{14}$$

$$f_t=x_{2t}+u_t, \tag{15}$$

$$\Delta x_{1t}=\alpha\left(x_{1t-1}-\beta x_{2t-1}\right)+\gamma\Delta x_{2t-1}+v_t. \tag{16}$$

The FECM for Δx_{1t} can be written as

$$\Delta x_{1t} = \alpha \left(x_{1t-1} - \beta f_{t-1} \right) + \gamma \Delta f_{t-1} + v_t + \alpha \beta u_{t-1} - \gamma \Delta u_{t-1}. \tag{17}$$

For the FAVAR, since

$$x_{1t} - \beta f_t = \frac{\gamma \Delta f_{t-1} + (v_t - \beta \varepsilon_t)}{1 - (\alpha + 1) L} + \frac{(\alpha + 1) \beta u_{t-1} - \gamma u_{t-1} - \beta u_t}{1 - (\alpha + 1) L},$$

then

$$\begin{aligned} \Delta x_{1t} &= \gamma \Delta f_{t-1} + \frac{\alpha \gamma \Delta f_{t-2}}{1 - (\alpha + 1) L} + v_t \\ &+ \underbrace{\alpha \beta u_{t-1} - \gamma \Delta u_{t-1}}_{\substack{\text{additional error of} \\ \text{FECM versus ECM}}} + \alpha \underbrace{\frac{v_{t-1} - \beta \varepsilon_{t-1} + (\alpha + 1) \beta u_{t-2} - \gamma u_{t-2} - \beta u_{t-1}}{1 - (\alpha + 1) L}}_{\substack{\text{additional error of} \\ \text{FAVAR versus FECM}}} \end{aligned} \tag{18}$$

Therefore, when the long-run and short-run evolution of x_1 are better explained by an observable variable such as x_2 rather than a common factor f, the ECM generates more accurate forecasts than the FECM. However, even in this case, the MSE of a FECM would be in general lower than that of a FAVAR, with equality only for the case $\alpha = 0$ (no cointegration).

In summary, this simple but comprehensive analytical example shows that from a theoretical point of view, the FECM can be expected to produce more efficient forecasts than the FAVAR in virtually all situations. The rationale, as explained in the previous section, is that the FAVAR is nested in the FECM, in the same way that a VAR in differences is nested in an ECM. However, as also discussed above, the theoretical advantages are not necessarily reflected in better forecasts in actual situations, since the specification of the FECM is more complex than that of the FAVAR, requiring us, for example, to determine the number and the coefficients of the cointegrating vectors. To assess the presence and size of forecasting gains from the FECM in practical situations, we now turn to a Monte Carlo evaluation and then to a set of empirical applications.

4 Monte Carlo experiments

In this section we consider two Monte Carlo experiments. The first experiment takes as the DGP the model (10) - (12) in the analytical example of the previous section. The second experiment considers a FECM DGP with a more complex structure that closely reflects the properties of one of our empirical applications in Section 5, and reflects very clearly the structure of (8).

4.1 A simple design

In accordance with the analytical example, we consider two types of DGP, a FECM and an ECM, since we are interested in the ranking of FAVAR and FECM in the two cases. For simplicity, we assume that the error process u_t does not contain a moving-average component. Hence, the FECM DGP is

$$\begin{bmatrix} \Delta x_{1t} \\ \Delta x_{2t} \\ \Delta f_t \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -\beta \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ f_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & \gamma & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{1t-1} \\ \Delta x_{2t-1} \\ \Delta f_{t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ \varepsilon_t - u_t \\ \varepsilon_t \end{bmatrix}, \tag{A1}$$

while in the case of the ECM DGP it is

$$\begin{bmatrix} \Delta x_{1t} \\ \Delta x_{2t} \\ \Delta f_t \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -\beta & 0 \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ f_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & \gamma & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x_{1t-1} \\ \Delta x_{2t-1} \\ \Delta f_{t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ \varepsilon_t \\ \varepsilon_t - u_t \end{bmatrix}, \tag{A2}$$

The parameters of the benchmark DGP are $\alpha = -0.5$, $\beta = 1.0$ and $\gamma = 0.6$. These are then changed to assess respectively the effects of the increased importance of the lagged differences of factors ($\gamma = 0.9$) and of the increased or decreased importance of the error-correction terms ($\alpha = -0.75$ or $\alpha = -0.25$).

The previous theoretical derivations suggest that we should observe gains in forecasting precision from using the FECM rather than the FAVAR for all DGPs, with larger gains when γ and α (in absolute terms) are larger in the case of a FECM DGP, and when α is larger with an ECM DGP. The ranking of the FECM to the ECM should instead depend on the type of DGP. In addition to the ECM, FAVAR and FECM, which are the main subjects of comparison, we also include three common empirical specifications in the comparison exercise, namely a simple autoregression (AR), a factor-augmented AR model (FAR) and a VAR consisting of the bivariate system given by $[\Delta x_{1t}, \Delta x_{2t}]'$. In all the models that allow for cointegration, a rank equal to one is imposed. In all the models the dynamics are determined by the Bayesian Information criterion (BIC), starting with six lags for each explanatory variable. The factors are assumed to be known in the estimated models and are included in levels in the FECM and in differences in the FAVAR and the FAR.[6]

We use (A1) and (A2) to generate 5000 random samples, each of 200 time series observations ($T = 200$), with the final 50 observations retained for out-of-sample forecasting. For the simple DGP we focus on the forecasting accuracy for x_1, which is the error-correcting variable in system (10) - (12). The $h-$step ahead forecasts are given by looking

[6] Typically factor estimation matters very little for forecasting, even when the sample size is relatively small, see e.g. the simulation experiments in Banerjee et al. (2008). In the next experiment we will also consider the case of estimated rather than known cointegration rank.

at $\hat{x}^h_{1,\tau+h} - x_{1,\tau},\ \tau = T-h-50,...,T-h$ and are constructed as

$$\hat{x}^h_{1,\tau+h} = x_{1,\tau} + \sum_{i=1}^{h} \Delta\hat{x}_{1,\tau+i},\ \tau = T-h-50,...,T-h. \tag{19}$$

The MSE is given by

$$MSE_h = \frac{1}{50}\sum_{j=1}^{50}\left(x^h_{1,T-50+j} - \hat{x}^h_{1,T-50+j}\right)^2 \tag{20}$$

and the MSEs from the competing models are benchmarked with respect to the MSE of the AR model.

We consider six different forecast horizons, $h = 1, 3, 6, 12, 18, 24$. In contrast to our use of iterated h-step ahead forecasts (dynamic forecasts), Stock and Watson (1998 and 2002a,b) adopt direct h-step ahead forecasts, but Marcellino, Stock and Watson (2006) find that iterated forecasts are often better, except in the presence of substantial misspecification.[7]

The results are reported in Table 1 . Starting with $h = 1$, the values are in line with the theoretical predictions. In particular, the FECM is virtually always better than the FAVAR. The MSE gains increase with α and γ and are also present for an ECM DGP. The ECM is worse than the FECM (and the FAVAR) with a FECM DGP, but becomes the best with an ECM DGP. However, interestingly, in this case the relative loss from the use of a FECM is rather small, although this result may be due to the relatively small dimension of the DGP considered here. Concerning the other models, the AR is systematically dominated since there is substantial interaction across the variables in both DGPs; the VAR is systematically worse than the ECM (cointegration matters); and the FAR is systematically better than the AR (the factor matters).

When the forecast horizon increases, the pattern described above remains qualitatively valid and the FECM consistently dominates all other models, but the MSE differences shrink substantially. In particular, already for $h = 3$ the FAVAR and ECM generate similar MSEs with a FECM DGP, and when $h = 24$ the MSEs from all models, including the AR, are very similar. This notable finding also emerges in earlier studies on the role of cointegration for forecasting, see e.g. Clements and Hendry (1995), and is due to the stationarity of the variables under analysis, which implies that the optimal h-step ahead forecast converges to the unconditional mean of the variable when the forecast horizon increases.

In summary, the Monte Carlo results confirm the theoretical findings for sample sizes common in empirical applications. The FECM appears to dominate the FAVAR in all cases, even when the FECM is not the DGP but cointegration matters. However, the

[7] 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: **Monte Carlo results: Out-of-sample forecasts of x1 from A1 and A2 DGPs**

					MSE relative to MSE of AR model				
h	**DGP**	α	β	γ	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**
	FECM	-0.50	1.00	0.60	0.54	0.81	0.54	0.65	0.48
	FECM	-0.50	1.00	0.90	0.43	0.76	0.44	0.62	0.36
1	FECM	-0.75	1.00	0.60	0.40	0.79	0.40	0.63	0.32
	FECM	-0.25	1.00	0.60	0.63	0.89	0.63	0.77	0.68
	ECM	-0.50	1.00	0.60	0.84	0.54	0.55	0.43	0.63
	FECM	-0.50	1.00	0.60	0.81	0.92	0.81	0.77	0.67
	FECM	-0.50	1.00	0.90	0.84	0.94	0.84	0.85	0.74
3	FECM	-0.75	1.00	0.60	0.83	0.94	0.83	0.80	0.69
	FECM	-0.25	1.00	0.60	0.84	0.94	0.84	0.76	0.69
	ECM	-0.50	1.00	0.60	0.93	0.83	0.83	0.67	0.77
	FECM	-0.50	1.00	0.60	0.88	0.94	0.88	0.82	0.76
	FECM	-0.50	1.00	0.90	0.90	0.95	0.90	0.90	0.81
6	FECM	-0.75	1.00	0.60	0.91	0.96	0.91	0.86	0.79
	FECM	-0.25	1.00	0.60	0.90	0.96	0.90	0.80	0.75
	ECM	-0.50	1.00	0.60	0.95	0.89	0.89	0.75	0.83
	FECM	-0.50	1.00	0.60	0.94	0.97	0.94	0.88	0.82
	FECM	-0.50	1.00	0.90	0.94	0.97	0.94	0.94	0.88
12	FECM	-0.75	1.00	0.60	0.95	0.98	0.95	0.90	0.84
	FECM	-0.25	1.00	0.60	0.95	0.98	0.95	0.87	0.83
	ECM	-0.50	1.00	0.60	0.97	0.93	0.93	0.87	0.92
	FECM	-0.50	1.00	0.60	0.94	0.97	0.94	0.92	0.88
	FECM	-0.50	1.00	0.90	0.95	0.98	0.95	0.93	0.90
18	FECM	-0.75	1.00	0.60	0.97	0.98	0.97	0.93	0.89
	FECM	-0.25	1.00	0.60	0.95	0.98	0.95	0.90	0.84
	ECM	-0.50	1.00	0.60	0.98	0.96	0.95	0.89	0.92
	FECM	-0.50	1.00	0.60	0.96	0.98	0.96	0.94	0.91
	FECM	-0.50	1.00	0.90	0.96	0.98	0.96	0.95	0.91
24	FECM	-0.75	1.00	0.60	0.98	0.98	0.98	0.92	0.89
	FECM	-0.25	1.00	0.60	0.97	0.98	0.97	0.91	0.87
	ECM	-0.50	1.00	0.60	0.98	0.96	0.96	0.90	0.93

Notes: 5000 Monte Carlo replications. T=200, last 50 observations retained for forecasting. Cointegration rank in ECM and FECM set to 1. Lag selection using BIC criterion.

simulations also indicate that the gains can shrink rapidly with the forecast horizon.

4.2 A more elaborate design

The second Monte Carlo experiment considers a more complex data generating process, which mimics the features observed in one of the empirical examples reported in Section 5, based on a large set of variables for the US. In particular, we estimate over the period 1985-2003, a FECM for four real variables (total industrial production, personal income less transfers, employment on non-agricultural payrolls and real manufacturing trade and sales) and four $I(1)$ factors extracted from the 104 $I(1)$ variables (out of 132 series) used in Stock and Watson (2005). The rank of the system is set to 4, in accordance with the estimates and in line with theoretical expectations. For simplicity, we set the number of lagged differences to 1, even though empirically this may not be sufficient.

As in the previous section, in the case of this DGP also, we want to assess how the relative forecasting precision of the FECM is affected by the importance of the error-correction mechanism. To this end, in addition to the basic design, we also consider experiments where we multiply the loading coefficient matrix α in the FECM by a constant c that takes on values 1, 0.75, 0.50 and 0.25, where by lowering c - relative to $c = 1$, which is the estimated model - we reduce the share of variability in the data induced by the variability of the error-correction term.

Overall, the DGP is

$$\begin{bmatrix} \Delta x_{At} \\ \Delta f_t \end{bmatrix} = \alpha_0 + c \begin{bmatrix} \gamma_A \\ \gamma_B \end{bmatrix} \delta' \begin{bmatrix} x_{At-1} \\ f_{t-1} \end{bmatrix} + A_1 \begin{bmatrix} \Delta x_{At-1} \\ \Delta f_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{At} \\ \epsilon_t \end{bmatrix}, \qquad (21)$$

with $c = \{0.25, 0.50, 0.75, 1.00\}$. The parameter values α_0, γ_A, γ_B, δ, and A_1 are taken to be equal to the estimated values from the system of real variables described above. The error process of the system is drawn from a multivariate normal distribution with variance-covariance matrix estimated from the data. The sample size and the length of the out-of-sample forecast period are constructed so as to match the empirical example, i.e. data sample 1985:1 - 2003:12 and forecast period 1996:1 - 2003:12. As in the case of the simple DGP the factors are assumed to be known.

We consider 10000 replications. For each replication, the lag length and the cointegration rank for the ECM and the FECM are determined recursively for each updating of the estimation sample as we move through the forecasting period. Determination of lag length is based on BIC for the results presented in Tables 2 and 3, but we have also checked robustness by using the Hannan-Quinn (HQ) criterion. The results appear robust to the use of different information criteria (details available upon request). As for the cointegration test, we have considered two approaches: the Johansen trace test (Johansen, 1995) and the Cheng and Phillips (2008) semi-parametric test based on standard information criteria. Both methods gave very similar results (details available upon request), but due to the lower computational burden and also ease of implementation in practice, we gave preference to the Cheng and Phillips method. As for determination of the lag length, the BIC information criterion was used.[8]

For the sake of brevity, we report in the main text only the results for $c = 1$ (Table 2) and $c = 0.25$ (Table 3). The details of the intermediate cases of $c = 0.75$ and 0.5 are deferred to the Appendix. The MSE calculations for each of the four variables are analogous to (19) and (20). Starting with Table 2 and $h = 1$, the FECM is indeed better than the FAVAR for all four variables. The FECM is also better than the ECM for all four variables, with comparable gains. The relative ranking of the other models is not clear-cut: VAR is the best for the fourth variable and the second best in terms of MSE for the first variable, while the ECM is the second best for the second and third variables. This is an interesting finding since it highlights the fact that the role of cointegration and of the factors can be rather unclear when misspecified models are compared.

When the forecast horizon h increases, four main findings emerge. First, the dominance

[8] Simulation results in Cheng and Phillips (2008) show that use of BIC tends to underestimate rank when true rank is not very low, while it performs best when true cointegration rank is very low (0 or 1). Given that BIC model selection is generally prefered for model selection for forecasting, we chose to use it also for testing for cointegration rank. However, our results (available upon request) are robust also to the use of HQ.

Table 2: **Monte Carlo results - DGP corresponding to FECM with real variables, c = 1.00**

		RMSE	MSE relative to MSE of AR model				
h	**Var**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**
	1	0.005	1.13	0.93	0.99	0.98	0.87
1	2	0.007	1.02	0.92	0.95	0.94	0.82
	3	0.001	1.10	1.09	1.34	1.05	0.86
	4	0.009	1.03	0.98	1.02	1.10	1.01
	1	0.011	1.25	0.88	0.96	1.00	0.72
3	2	0.012	1.02	0.82	0.85	0.76	0.54
	3	0.003	1.34	1.22	1.45	1.08	0.59
	4	0.014	1.01	0.91	0.94	1.04	0.81
	1	0.020	1.24	0.90	0.96	0.95	0.64
6	2	0.019	1.02	0.76	0.81	0.66	0.47
	3	0.007	1.39	1.27	1.41	0.98	0.57
	4	0.020	1.01	0.90	0.91	1.03	0.76
	1	0.037	1.17	0.92	0.94	1.00	0.69
12	2	0.031	1.02	0.79	0.82	0.64	0.45
	3	0.014	1.40	1.33	1.38	1.10	0.67
	4	0.030	1.00	0.90	0.90	1.07	0.76
	1	0.054	1.15	0.96	0.97	0.99	0.74
18	2	0.042	1.01	0.82	0.85	0.60	0.46
	3	0.023	1.39	1.33	1.36	1.10	0.77
	4	0.040	1.00	0.91	0.91	1.16	0.81
	1	0.070	1.07	0.96	0.97	1.20	0.91
24	2	0.052	1.01	0.85	0.87	0.79	0.58
	3	0.032	1.23	1.20	1.20	1.10	0.83
	4	0.048	1.01	0.93	0.93	1.36	0.90
		AR	2.03	1.20	2.64	1.07	
	Lags	FAR	0.70	0.76	0.99	0.81	
		VAR	FAVAR	ECM	FECM		
		0.99	0.71	0.17	0.08		
			ECM			FECM	
Cointegration		mean	min	max	mean	min	max
rank		1.51	0.98	2.38	3.09	2.46	3.51

Notes: 10000 replications. The DGP corresponds to the FECM estimated on 4 real US variables and 4 factors with cointegration rank 4 and 1 lagged difference. Sample sizes and out-of-sample forecast period are constructed so as to fit the empirical example, i.e. data sample 1985:1 - 2003:12 and forecast period 1996:1 - 2003:12. Cheng and Phillips (2008) cointegration rank test and lag selection based on BIC information criterion.

of the FECM over other models becomes more pronounced. Second, in contrast with the simple DGP of the first experiment, the MSE gains of the FECM with respect to the AR in general increase as long as $h < 24$, and start decreasing only for $h = 24$. Third, the FAVAR remains systematically worse than the FECM for all variables and horizons, but it also becomes worse than the ECM in most cases. This suggests that for this DGP cointegration does matter, possibly more than the inclusion of the factors. Finally, the ECM performs quite well with respect to the other models; it is the second-best choice for most variables and forecast horizons.

The results on the role of the strength of the error correction mechanism, which is much weaker in Table 3 where we use $c = 0.25$, are perhaps even more interesting. When $h = 1$, the FECM becomes worse than AR for all four variables, even if it is the specification that corresponds to the DGP. Moreover, the gains with respect to the FAVAR and to the ECM basically disappear, and the performance of the three models is very similar, and similar to that of the AR, FAR and VAR. One reason for this result may be the fact that the Cheng and Phillips (2008) test for rank based on BIC heavily underestimates the rank.

Table 3: **Monte Carlo results - DGP corresponding to FECM with real variables, c = 0.25**

		RMSE	MSE relative to MSE of AR model				
h	**Var**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**
	1	0.005	1.00	0.99	1.00	1.00	1.01
1	2	0.006	1.03	1.04	1.05	1.06	1.06
	3	0.001	0.94	1.07	1.11	1.05	1.12
	4	0.009	1.03	1.09	1.12	1.13	1.14
	1	0.009	1.01	0.99	1.01	0.98	1.00
3	2	0.010	1.00	1.00	1.00	1.00	1.01
	3	0.003	1.14	1.19	1.26	1.09	1.21
	4	0.012	1.01	1.02	1.03	1.05	1.04
	1	0.015	1.01	0.99	1.00	0.98	0.99
6	2	0.015	1.00	0.99	1.00	0.99	1.00
	3	0.005	1.14	1.16	1.20	1.04	1.14
	4	0.017	1.01	1.02	1.02	1.10	1.03
	1	0.022	1.01	0.99	1.00	0.99	1.00
12	2	0.020	1.01	1.00	1.00	1.02	1.01
	3	0.008	1.13	1.14	1.17	1.05	1.12
	4	0.023	1.01	1.01	1.02	1.09	1.03
	1	0.032	1.01	0.99	1.00	1.04	1.00
18	2	0.028	1.00	0.99	1.00	0.97	1.00
	3	0.012	1.10	1.11	1.12	1.10	1.10
	4	0.030	1.01	1.00	1.01	1.08	1.02
	1	0.043	1.01	1.00	1.01	1.11	0.99
24	2	0.032	1.00	0.99	1.00	1.09	1.01
	3	0.016	1.11	1.11	1.12	1.08	1.06
	4	0.038	1.01	1.00	1.00	1.23	1.02
		AR	0.21	0.88	1.36	1.02	
	Lags	FAR	0.50	0.52	0.77	0.74	
		VAR	FAVAR	ECM	FECM		
		0.30	0.10	0.09	0.00		
			ECM			FECM	
Cointegration		mean	min	max	mean	min	max
rank		0.56	0.09	1.42	0.29	0.02	0.78

Notes: see Table 3.

However, a robustness check with respect to the use of the HQ criterion leaves this finding virtually unchanged despte the fact that with HQ the cointegration rank is on average correctly set to four. The issue is that in this context of mild error correction, parsimony pays: dropping, by mistake, the error-correction terms and the lagged factors can even be beneficial! When h increases the FECM returns to beating the FAVAR systematically, but not the ECM, and the AR model remains a tough competitor.

We have also checked whether these results may be influenced by the size of the estimation sample. Indeed, by increasing the length of the time series of generated data from 228 to 600 in the Monte Carlo, the FECM returns to being the best model at all horizons. But consistent with the fact that the share of data variability induced by the error correction term is considerably smaller than in the case of original DGP, the observed gains are also considerably smaller.

In summary, the more complex Monte Carlo design indicates that in empirically relevant situations the strength of the error correction mechanism matters in determining the ranking of the alternative forecasting models. While the FECM remains better than the FAVAR in most cases, simpler models such as the ECM or even AR can become tough competitors when the explanatory power of the error correction terms and/or of the factors is reduced, and the sample size is not very large. Thus, while having a suitably large N dimension is beneficial for the computation of the factors, a relatively short T dimension will imply that the cointegrating information may be poorly incorporated in the FECM. Thus if cointegration is important, but the factors less so, a large N environment (which facilitates the use of factors) will not necessarily represent an advantage for the FECM. In such circumstances, as we show below, the ECM may be the dominant method.

5 Empirical applications

In order to provide convincing evidence of the usefulness of the FECM as a forecasting tool, we consider a number of empirical examples that differ in terms of the type of economic application, countries and time periods. In these examples we extract factors from four different datasets.

As discussed in the introduction, the first dataset is a large panel of monthly US macroeconomic variables from Stock and Watson (2005) that includes 132 monthly time series, over the period 1959:1 to 2003:12. For the estimation of the $I(1)$ factors to be used in the FECMs, we have considered two options. First, we have retained only the 104 series that are considered as $I(1)$ by Stock and Watson. Second, we have cumulated the remaining 28 $I(0)$ series and added them to the $I(1)$ dataset before extracting the $I(1)$ factors. Since our main findings are robust to the use of either option, we report results based only on the former. The data series as well as the transformations implemented are listed in Table 17 in the Appendix.

Based on this dataset we consider forecasting three different systems of variables. The first two follow the choice of variables in Stock and Watson (2002b), i.e. we forecast four real variables and four inflation rates. The third system is in spirit closer to the standard practice of a small-scale macroeconomic modelling as it includes indicators of real output, inflation rate and the nominal interest rate.

The second dataset is taken from Marcellino and Schumacher (2008). It contains 90 monthly series for the German economy over the sample period 1991:1-2007:12. As in the case of the US dataset, the time series cover broadly the following groups of data: prices, labour market data, financial data (interest rates, stock market indices), industry statistics and construction statistics. The source of the time series is the Bundesbank database. The details of this dataset are given in Table 16 in the Appendix. With the factors extracted from this dataset we estimate a system analogous to the US three-variable system of mixed variables, which includes measures of real output, inflation rate and the short-term nominal interest rate.

The use of the third dataset is motivated by the analysis of the yield curve where it is commonly assumed that the dynamics of this curve are driven by a small number of factors, typically referred to as the level, slope and the curvature factors. In other words, theoretically we expect to find a lot of cointegration among the yields at different maturities. We therefore extract the factors from a panel consisting of nominal yields only and consider forecasting interest rates at different maturities. The dataset used is taken from Carriero et al. (2009b) who use the US Treasury zero coupon yield curve estimates by Gürkaynak, Levin & Swanson (2009). The data on 18 different maturities - from 1 month to 10 years - are monthly, ranging from 1980:1 to 2007:12.

In our final example we consider forecasting three major bilateral exchange rates (the euro, the British pound and the Japanese yen against the US dollar) with or without using information on a large set of other exchange rates. This example is of interest since Carriero et al. (2009b) find that cross-sectional information may be relevant for forecasting exchange rates. Economic theory provides less guidance here on the number of common trends and the amount of cointegration we should expect in the data and the exercise is therefore a challenging application for a model like FECM. The data are taken from , Carriero et al (2009b) and comprise the monthly averages of the exchange rates vis-a-vis the dollar for 43 currencies for the period 1995:1 - 2008:4. Details of this data are given in Table 18 in the Appendix.

Prior to computation of the factors and estimation of the competing forecasting models, the raw data were transformed in the following way. First, natural logarithms were taken for all time series except interest rates. In addition, the logarithms of price series were differenced, which implies that inflation rates were treated as $I(1)$. To achieve stationarity for the extraction of the $I(0)$ factors used in the FAVAR analysis, all series (including inflation rates) were differenced once. If not adjusted already at the source, the time series were tested for presence of seasonal components and adjusted accordingly with the $X-11$ filter prior to the forecast simulations. Extreme outlier correction was achieved using a modification of the procedure proposed by Watson (2003). Large outliers are defined as observations that differ from the sample median by more than six times the sample interquartile range (Watson, 2003, p. 93). As in Stock and Watson (2005), the identified outlying observations were set to the median value of the preceding five observations.

For the computation of $I(1)$ factors included in the FECM all variables are treated as $I(1)$ with non-zero mean. The $I(1)$ factors are estimated with the method of Bai (2004) (see details below on the number of factors extracted from each dataset). For the $I(0)$ factors included in the FAVAR and FAR, we first transform the data to stationarity and then use the principal component based estimator of Stock and Watson (2002a).

Three further issues related to the factors deserve comment. First, the estimated factors are consistent only for the space spanned by the true factors but not necessarily for the true factors themselves. However, this is not a problem in a forecasting context, since if the true factors have forecasting power a rotation of these factors preserves this property. In addition, if the original factors are $I(1)$, not cointegrated amongst themselves, but cointegrated with the variables of interest, these features are also preserved by a rotation.

Second, the use of estimated factors rather than true factors does not create a generated regressor problem as long as the longitudinal dimension grows faster than the temporal dimension, the precise condition is $T^{1/2}/N$ is $o(1)$, see Bai and Ng (2006). Intuitively, the principal component based estimator estimates the factors as weighted averages of N contemporaneous variables. Thus, when N is large enough with respect to the temporal dimension T, the convergence of the estimator is sufficiently fast to avoid the generated regressor problems.

Third, we find a mismatch in the number of $I(1)$ and $I(0)$ factors which suggests that the variables in levels could be driven by (one or more) $I(0)$ factors in addition to the $I(1)$ factors, but the former are "hidden" by the $I(1)$ factors. While the $I(1)$ factors are related to the common trends, the $I(0)$ factor generates common cycles. To assess the possible presence of $I(0)$ factors, we have computed the (stationary) residuals of a regression of the $I(1)$ variables on the $I(1)$ estimated factors, and then computed principal components of the residuals. In some cases it turns out that the first component explains a significant proportion of the total variability of the residuals (for example about 22% in the case of the US data), providing support for the existence of an additional $I(0)$ factor for the variables in levels. The equation for this additional $I(0)$ factor is then added as an additional equation in the FECM, and we label the resulting model as FECMc, where "c" stands for common cycles.

The number of $I(1)$ and $I(0)$ factors is kept fixed over the forecasting period, but their estimates are recursively updated. Each forecasting recursion also includes model selection. As in the second Monte Carlo experiment, both the cointegration rank and the lag length are based on using the BIC. As a robustness check we have experimented with the use of the Johansen trace test to determine the cointegration rank and with HQ for cointegration rank and/or lag length determination, but the results (available upon request), are qualitatively similar.

Forecasting is performed using the same set of models we have considered in the previous section. Hence, we construct AR, VAR and ECMs that are all based on the observable variables, and FAR, FAVAR and FECM specifications that augment, respectively, the AR, VAR and ECMs with factors extracted from the larger set of available variables, in order to assess the forecasting role of the additional information.

The levels of the real variables (measures of output) are treated as $I(1)$ with 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 to obtain the forecast of the level $h-$steps ahead. This is also the case for the nominal exchange rates. For the inflation rates and interest rates, the dynamic forecasts of the differences of the variables $h-$steps ahead are cumulated to obtain the forecast of the level of the specific inflation rate or interest rate $h-$steps ahead.

The results of the forecast comparisons are presented in two ways. First, for each empirical example, we first list the MSEs of the competing models relative to the MSE of the AR at different horizons for each variable under analysis. These tables also report information on cointegration rank selection and the number of lags in each model. However, in order to present the information in a more condensed fashion we provide a summary table at the end of this section. Specifically, the upper panel of Table 13 reports the occurrence of the best performance of the competing models across horizons and variables. In addition, the lower panel of Table 13 reports summary statistics that we use in assessing the overall importance of cointegration and factors for forecasting. The role of potential extra information embedded in the factors can be evaluated by comparing the relative performance of the FAVAR relative to the VAR, and the FECM relative to the ECM. Conversely, information on the importance of cointegration can be obtained by comparing the ECM and the VAR, and the FECM and the FAVAR. Observing that the FECM significantly improves over both the ECM and the FAVAR can be seen as an indication that it may not be sufficient to consider separately either cointegration or factors, but rather the information that $I(1)$ factors have about the long run or equilibrium dynamics of the data. The sub-sections which follow contain details of each of the empirical applications.

5.1 Forecasting US nominal and real variables

As discussed previously, in the first empirical application we consider forecasting two sets of US macroeconomic monthly series in line with the choice of Stock and Watson (2002a,b). In particular, the set of real variables is given by: total industrial production (IP), personal income less transfers (PI), employment on non-agricultural payrolls (Empl), and real manufacturing trade and sales (ManTr). The set of nominal variables, on the other hand, is given by: inflation of producer prices of finished goods (PPI), CPI inflation, all items (CPIall), inflation of CPI less food (CPI no food), and growth of personal consumption expenditure deflator (PCEdefl).

Concerning the choice of sample period, we proceed in the following manner. Precise estimation of the cointegration relationships and their loadings, and the need for a long evaluation sample, would suggest use of the longest available sample. Instead, the possible presence of structural breaks that could have affected both the long run and the short run dynamics, such as the Great Moderation, suggests that focusing on a shorter but more homogeneous sample could be better. Since it is *a priori* unclear which option is best, we consider two periods. First, we focus on the post-1985 data. The forecast period in this case is 1996:1 - 2003:12. Second, we start estimation in 1959:1 and, for comparability with Stock and Watson (2002b), in this case the forecast period spans from 1970:1 to 1998:12.

The number of factors included in the FECM is set to four, since four factors explain 96% of data variability in the 1985 - 2003 sample. We have also tried the IPC2 criterion from Bai (2004) to determine the number of factors, and it signalled no common trends in the entire dataset but four factors on the subset of real data. Since the information criteria are sometimes sensitive to the sample size and the properties of the idiosyncratic errors, and given that in our context overestimating the number of factors is less problematic than underestimating it, we proceeded with the analysis using four factors.

As explained above, we assess the possible presence of an additional $I(0)$ factor in the FECM. To this end, we have computed the (stationary) residuals of a regression of the $I(1)$ variables on the four $I(1)$ estimated factors, and then computed the principal components of the residuals. The first component explains a significant proportion of the total variability of residuals (for example about 22% in the case of US data), providing support for the existence of an additional $I(0)$ factor for the variables in levels. In comparison with the FECM, our FECMc model contains one additional $I(0)$ factor.

For the $I(0)$ factors included in the FAVAR and FAR, we use the principal component based estimator of Stock and Watson (2002a) and set their number to five, in line with the choice for the FECMc above, since five factors are able to explain 90% of the overall variability in the stationary data. Moreover, the Bai and Ng (2002) $PC2$ criterion also suggests five factors.

5.1.1 Forecasting real variables

Tables 4 to 7 report the MSEs, computed analogously to (19) and (20), of the FAR, VAR, FAVAR, ECM, FECM and FECMc relative to that of the AR model for forecasting the four real and four nominal variables over the two sub-periods.

Table 4 reports the results for forecasting the four real variables over the sample 1996 - 2003, with estimation starting in 1985. When $h = 1$, only few models are better than the AR. The FECM is the best model for industrial production and employment but performs worse than the FAVAR and the ECM for personal income less transfers and real manufacturing trade and sales. This pattern suggests that cointegration matters, but parsimony is also important, so much so that the AR is diffcult to beat.

When h increases, the picture changes. Now the FECM is better than the AR in 12 out of 20 cases, and it produces the lowest MSE in 4 cases. However, combined also with the results of the FECMc, the overall score of best performance increases to 14. The FAVAR and the ECM perform best only in 1 case each. The gains of the FECM relative to the benchmark AR increase with the forecast horizon, levelling off after $h = 12$ and slightly diminishing at the longest, two-year horizon. For some of the variables, such as industrial production and employment, the gains relative to the AR exceed 30%. Other models do not offer comparable gains.

These results show that for the real variables the inclusion of both additional information and adjustment to disequilibrium significantly contribute to forecasting precision, except at the shortest horizon. It is not easy to disentangle the relative contribution of the two elements. Table 13 provides some aid in this respect. The fact that the ECM outperforms the VAR, and the FECM the FAVAR in more than half of the cases suggests that cointegration matters, in line with theory and the simulation results of the previous section. But the fact that the FAVAR outperforms the VAR only twice, while the corresponding score of the FECM relative to the ECM is 18 out of 24, suggests that it is the

Table 4: **Forecasting US real variables, evaluation period 1996 - 2003**

h	Log of	RMSE of AR	FAR	VAR	FAVAR	ECM	FECM	FECMc
			MSE relative to MSE of AR model					
	PI	0.004	1.15	0.82	0.84	1.21	1.35	2.12
1	ManTr	0.008	1.16	1.02	1.10	1.22	1.31	1.69
	IP	0.005	1.09	0.98	1.05	1.03	0.92	1.04
	Empl	0.001	1.18	1.52	1.59	1.22	0.87	1.11
	PI	0.009	1.04	0.79	0.79	1.37	1.26	1.17
3	ManTr	0.011	1.02	0.97	0.99	1.25	1.15	1.58
	IP	0.011	0.96	0.93	0.94	0.95	0.77	0.76
	Empl	0.002	1.15	2.12	2.16	1.34	0.61	0.83
	PI	0.015	1.03	0.83	0.84	1.07	1.03	0.90
6	ManTr	0.016	1.02	0.98	1.00	1.35	1.07	1.66
	IP	0.020	0.95	0.92	0.93	0.83	0.73	0.68
	Empl	0.005	1.20	2.36	2.37	1.29	0.65	1.00
	PI	0.027	1.01	0.88	0.89	0.81	0.90	0.73
12	ManTr	0.024	1.00	0.97	0.97	1.17	1.05	1.48
	IP	0.036	0.98	0.95	0.95	0.84	0.79	0.73
	Empl	0.012	1.16	1.90	1.90	1.19	0.86	0.86
	PI	0.037	1.01	0.91	0.93	0.74	0.88	0.71
18	ManTr	0.031	1.00	0.97	0.97	1.09	1.06	1.31
	IP	0.050	1.00	0.97	0.98	0.87	0.81	0.83
	Empl	0.019	1.15	1.66	1.67	1.16	0.99	0.77
	PI	0.047	1.01	0.93	0.96	0.72	0.92	0.81
24	ManTr	0.038	1.00	0.98	0.99	1.06	1.08	1.03
	IP	0.064	1.01	0.98	0.99	0.91	0.83	0.91
	Empl	0.027	1.12	1.44	1.46	1.14	1.00	0.71
		AR	1.00	2.55	0.68	3.00		
	Lags	FAR	1.83	1.84	1.83	2.00		
		VAR	FAVAR	ECM	FECM	FECMc		
		1.00	0.59	0.00	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			1.93	1.00	3.00	3.18	2.00	4.00

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: 1985:1 - 2003:12, forecasting: 1996:1 - 2003:12.
Variables: IP - Industrial production, PI - Personal income less transfers, Empl - Employees on non-aggr. payrolls, ManTr - Real manufacturing trade and sales

combination of cointegration and a large information set that really matters both at short and long forecast horizons.

In Table 5 we investigate the longer forecasting sample 1970 - 1998, with estimation starting in 1959, as considered by Stock and Watson (2002b).[9] In essence, these results confirm the evidence of the FECM or the FECMc as the best forecasting model. The only notable difference with respect to the shorter evaluation period is in the relation between the FAVAR and the VAR. The FAVAR now outperforms the VAR 16 times instead of only twice, in line with Stock and Watson (2002b) although their results were based on direct rather than iterated forecasts. This difference across samples indicates the diminishing importance of factors for forecasting in the recent period, a finding also documented by D'Agostino, Giannone and Surico (2007). The FECM or FECMc remain the best models in 15 out of 24 cases. The FAVAR is best in only 4 out of 24 cases and the ECM never

[9] On a common estimation and evaluation sample we can confirm that the method of direct h-step-ahead forecasts and our iterative h-step-ahead forecasts produce similar benchmark results. Namely, the root mean sqared errors of the AR models reported by Stock and Watson (2002b) for personal income, industrial production, manufacturing trade and sales and non-agricultural employment at 12-month horizon are 0.027, 0.049, 0.045 and 0.017 respectively. Our corresponding RMSEs are 0.026, 0.049, 0.045 and 0.020.

Table 5: **Forecasting US real variables, evaluation period 1970 - 1998**

		RMSE		**MSE relative to MSE of AR model**				
h	**Log of**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECMc**
	PI	0.007	1.02	0.94	0.92	0.93	0.90	0.93
1	ManTr	0.011	1.04	0.98	0.95	1.10	1.03	1.00
	IP	0.007	0.99	1.08	0.95	1.11	1.24	1.15
	Empl	0.002	1.09	1.33	1.20	1.40	1.34	1.40
	PI	0.011	1.01	0.91	0.87	0.94	0.85	0.91
3	ManTr	0.018	1.01	1.01	0.96	1.21	0.97	0.93
	IP	0.017	0.96	1.04	0.94	1.10	1.17	1.09
	Empl	0.005	1.12	1.51	1.40	1.64	1.52	1.57
	PI	0.016	1.00	0.94	0.92	1.02	0.86	0.95
6	ManTr	0.029	1.01	1.01	0.98	1.17	0.89	0.87
	IP	0.029	0.97	1.00	0.96	1.08	1.08	1.02
	Empl	0.010	1.10	1.34	1.32	1.49	1.36	1.37
	PI	0.026	1.00	0.96	0.96	1.04	0.87	0.93
12	ManTr	0.045	1.01	0.99	0.98	1.07	0.74	0.75
	IP	0.049	0.99	1.00	0.99	1.03	0.96	0.94
	Empl	0.020	1.02	1.11	1.12	1.25	1.10	1.11
	PI	0.036	1.01	0.98	0.98	1.09	0.89	0.96
18	ManTr	0.058	1.00	1.00	0.99	1.06	0.71	0.73
	IP	0.065	1.00	1.00	1.00	1.08	0.93	0.96
	Empl	0.029	0.96	0.99	1.00	1.15	0.97	0.99
	PI	0.042	1.01	0.99	0.99	1.07	0.90	0.96
24	ManTr	0.069	1.01	1.00	1.01	0.99	0.64	0.66
	IP	0.076	1.01	0.99	1.00	1.07	0.90	0.95
	Empl	0.037	0.91	0.91	0.92	1.04	0.88	0.91
		AR	0.99	0.66	1.81	3.15		
	Lags	FAR	1.85	1.84	1.85	1.85		
		VAR	FAVAR	ECM	FECM	FECMc		
		1.33	0.93	0.81	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			3.66	2.00	4.00	3.87	1.00	4.00

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

produces the lowest MSE (see Table 13).

5.1.2 Forecasting nominal variables

The results for forecasting nominal variables are reported in Tables 6 and 7 for, respectively, the more recent and longer evaluation sample. Focusing first on the sample 1985 - 2003, we clearly observe a much weaker performance of the FECM (and the FECMc) relative to its performance in forecasting the real variables. The FECM is never the best model. Also relative to the FAVAR the performance of the FECM is relatively weak, outperforming it only 7 times.

Turning our attention to forecasting nominal variables over the period 1970 - 1998 (Table 7), we find that the FECM performs considerably better. In particular, the FECM is the best model on average 15 out of 24 times, while combined with the FECMc the score increases to 18 (see also Table 13). The performance of the FECM relative to the FAVAR and the ECM also changes dramatically. It almost always outperforms the FAVAR and is better than the ECM in two-thirds of the cases.

The differences in the findings across the two samples suggest that the decrease of

Table 6: **Forecasting US nominal variables, evaluation period 1996 - 2003**

h	Inflation of	RMSE of AR	MSE relative to MSE of AR model FAR	VAR	FAVAR	ECM	FECM	FECMc
	PPI	0.005	1.22	1.31	1.50	0.99	1.36	1.38
1	CPI all	0.002	1.10	1.09	1.36	1.04	1.30	1.33
	CPI no food	0.002	0.99	1.02	1.25	0.94	1.18	1.21
	PCEdefl	0.002	1.01	0.93	1.38	0.92	0.93	0.96
	PPI	0.005	1.23	1.18	1.40	0.94	1.19	1.22
3	CPI all	0.002	1.09	1.10	1.30	1.03	1.31	1.36
	CPI no food	0.002	1.08	1.08	1.20	0.98	1.19	1.23
	PCEdefl	0.002	1.19	1.10	1.65	1.16	1.37	1.41
	PPI	0.005	1.19	1.17	1.39	0.89	1.51	1.55
6	CPI all	0.002	1.17	1.17	1.44	1.03	1.77	1.85
	CPI no food	0.002	1.01	1.03	1.28	0.90	1.52	1.58
	PCEdefl	0.002	0.99	0.95	1.03	1.08	1.21	1.25
	PPI	0.006	0.98	1.06	1.20	0.76	1.66	1.73
12	CPI all	0.002	1.14	1.12	1.31	0.89	1.94	2.05
	CPI no food	0.003	1.07	1.09	1.26	0.85	1.74	1.83
	PCEdefl	0.002	1.13	1.06	1.23	1.02	1.68	1.76
	PPI	0.006	1.10	1.09	1.09	0.76	1.77	1.84
18	CPI all	0.002	1.13	1.09	1.29	0.94	2.29	2.41
	CPI no food	0.003	1.04	1.04	1.18	0.88	2.05	2.16
	PCEdefl	0.002	1.07	1.03	1.27	0.95	1.82	1.90
	PPI	0.006	1.07	1.02	1.11	0.75	1.94	2.02
24	CPI all	0.002	1.19	1.17	1.41	0.93	2.48	2.55
	CPI no food	0.003	1.08	1.06	1.29	0.88	2.14	2.18
	PCEdefl	0.002	1.17	1.12	1.39	1.04	2.11	2.18
		AR	5.47	5.99	4.75	5.78		
	Lags	FAR	1.86	1.85	1.84	1.93		
		VAR	FAVAR	ECM	FECM	FECMc		
		2.07	1.03	0.00	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			4.00	4.00	4.00	4.01	4.00	8.00

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: 1985:1 - 2003:12, forecasting: 1996:1 - 2003:12.
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)

importance of factors for forecasting for the more recent period, which we have already observed to some extent for real variables, seems to be stronger for the case of nominal variables.

5.2 A monetary FECM for the US

There is by now a large literature on the use of small VAR models to assess and forecast the effects of monetary policy, see e.g. Rudebusch and Svensson (1998). Favero et al. (2005), inter alia, have proposed augmenting these models with factors extracted from large datasets. In concordance with this approach, we now assess the performance of a FECM which includes as economic variables total industrial production (IP), CPI excluding food (CPI no food) and a three-month interest rate (3m T-bill).

The results are reported in Tables 8 and 9 for, respectively, the more recent and longer evaluation sample, where the factors are extracted from the same dataset as in the previous sub-section.

Focusing first on the sample 1985 - 2003, we see in Table 8 the superior performance of the FECM (and FECMc) for forecasting the real variable (IP) and the nominal variable

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

		RMSE	**MSE relative to MSE of AR model**					
h	**Inflation of**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECMc**
	PPI	0.005	1.05	1.05	1.04	0.90	0.90	0.90
1	CPI all	0.002	1.04	1.01	1.05	0.95	0.86	0.86
	CPI no food	0.002	0.98	0.94	0.99	0.91	0.93	0.91
	PCEdefl	0.002	1.04	0.97	1.04	1.04	0.92	0.92
	PPI	0.005	1.13	1.12	1.16	0.89	0.93	0.96
3	CPI all	0.003	1.10	1.08	1.14	1.06	0.82	0.83
	CPI no food	0.003	1.03	1.01	1.06	0.98	0.90	0.91
	PCEdefl	0.002	1.14	1.12	1.19	1.39	1.18	1.20
	PPI	0.005	1.15	1.13	1.22	1.03	0.97	1.00
6	CPI all	0.003	1.19	1.17	1.24	1.35	1.01	1.02
	CPI no food	0.003	1.04	1.02	1.09	1.13	0.97	0.98
	PCEdefl	0.002	1.12	1.10	1.15	1.67	1.25	1.29
	PPI	0.005	1.12	1.11	1.18	0.93	0.91	0.95
12	CPI all	0.003	1.07	1.06	1.09	1.16	0.84	0.86
	CPI no food	0.003	1.03	1.01	1.05	1.00	0.86	0.88
	PCEdefl	0.002	1.06	1.05	1.07	1.41	0.95	0.98
	PPI	0.006	1.08	1.06	1.14	0.95	0.96	1.02
18	CPI all	0.003	1.05	1.04	1.08	1.07	0.88	0.91
	CPI no food	0.004	1.03	1.02	1.06	0.99	0.95	0.97
	PCEdefl	0.003	1.05	1.04	1.08	1.22	0.97	1.02
	PPI	0.006	1.12	1.12	1.18	0.76	0.84	0.91
24	CPI all	0.004	1.11	1.10	1.14	0.85	0.82	0.85
	CPI no food	0.004	1.05	1.03	1.07	0.79	0.84	0.87
	PCEdefl	0.003	1.09	1.07	1.11	1.03	0.85	0.90
		AR	5.10	4.70	4.38	5.12		
	Lags	FAR	1.99	1.87	1.89	1.85		
		VAR	FAVAR	ECM	FECM	FECMc		
		2.53	1.35	0.00	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			4.00	4.00	4.00	4.00	4.00	7.00

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)

(CPI no food) for all horizons up to $h = 24$. For these two variables, the FECM or FECMc is the best-performing model in 11 cases out of 12 (it is equal-best in one case with the VAR, i.e. for IP when $h = 1$). The ECM, while being dominated by the FECM, is nevertheless clearly better than the FAR, VAR and FAVAR for both the real and nominal variable. Taken together, these results emphasize the importance of both factors and cointegrating information in forecasting in this system.

For the financial variable (3m T-bill), FECM, ECM and FECMc never provide the best-performing model, while FAVAR is equal to or narrowly better than the VAR, and delivers the best forecasting model, in 5 out of 6 cases . For $h = 1$, the VAR is the best model. In this example, the use of long-run information in forecasting the financial variable is thereby seen to be limited, although factors remain important.

For the period 1970 - 1998 (Table 9), the FECM or FECMc are the best models in 9 out of 18 cases. VAR does best in 6 out of 18 cases, although all these 6 cases are for the 3m T-Bill rate. Therefore in 9 out of 12 cases where a real or nominal variable is involved, both factors and long-run information are relevant. Within this category (real or nominal) the ECM does best in 2 out of 12 cases (for IP at horizons 12 and 18) while

Table 8: **US monetary FECM, evaluation sample 1996 - 2003**

		RMSE		**MSE relative to MSE of AR model**				
h	**Var**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECMc**
	IP	0.005	1.09	0.97	1.03	1.07	0.97	0.94
1	CPI no food	0.002	0.99	1.26	1.34	0.93	0.89	0.90
	3m T-Bill	0.180	1.03	0.89	0.96	1.17	0.96	1.09
	IP	0.011	0.96	0.96	0.96	1.08	0.86	0.80
3	CPI no food	0.002	1.08	1.17	1.43	0.93	0.91	0.91
	3m T-Bill	0.394	1.05	0.88	0.88	1.30	0.96	1.15
	IP	0.020	0.95	0.98	0.97	1.02	0.80	0.71
6	CPI no food	0.002	1.01	1.26	1.38	0.88	0.85	0.86
	3m T-Bill	0.675	1.06	0.96	0.95	1.35	1.07	1.34
	IP	0.036	0.98	1.00	0.99	1.05	0.83	0.73
12	CPI no food	0.003	1.07	1.34	1.33	0.92	0.89	0.94
	3m T-Bill	1.232	0.99	0.94	0.93	1.32	1.41	1.65
	IP	0.050	1.00	1.02	1.01	1.15	0.83	0.71
18	CPI no food	0.003	1.04	1.27	1.25	0.96	0.95	1.01
	3m T-Bill	1.610	0.97	0.94	0.94	1.24	1.52	1.81
	IP	0.064	1.01	1.02	1.01	1.23	0.83	0.76
24	CPI no food	0.003	1.08	1.37	1.45	0.99	0.95	1.00
	3m T-Bill	1.929	0.99	0.96	0.95	1.09	1.49	1.75
		AR	0.68	4.75	2.92			
	Lags	FAR	1.83	1.84	1.83			
		VAR	FAVAR	ECM	FECM	FECMc		
		1.19	0.86	0.45	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			1.69	1.00	3.00	2.98	2.00	3.00

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: 1985:1 - 2003:12, forecasting: 1996:1 - 2003:12. *Variables:* IP - log of industrial production index, CPI no food - inflation of consumer prices without food, 3m T-Bill - 3-month T-Bill yield.

in the remaining case (IP at horizon 24) the FAR provides the best model. In common with the shorter sample, the usefulness of long-run information in forecasting the financial variables is limited. In addition, for this longer sample, we find that factors are not useful for the 3m T-bill rate, with the VAR dominating the FAVAR (albeit narrowly).

5.3 A monetary FECM for Germany

We now consider a monetary FECM as in the previous example but using data for Germany, the largest economy in the euro area, for which a smaller sample is available due to the reunification. The economic variables under analysis are: total industrial production (IP), Inflation of consumer price index excluding food (CPI no food), and the 3 month money market rate (3m IntRate).

The FECM system in this case includes 2 $I(1)$ factors, which account for 76% and 11% of overall data variability respectively. Into the FECMc we have included only one additional factor. The number of factors included in the FAVAR is set to four. In this case the first principal component is not so dominant in explaining the variability of the data as it captures 30% of the variation. The second component follows closely with 28%, while the third and fourth account for 12% and 6% respectively. The monthly data spans over the 1991 - 2007 period, and we set the forecast evaluation sample to 2002:1 - 2007:12.

Table 10 reports the MSEs, computed analogously to (19) and (20), of the FAR, VAR,

Table 9: **US monetary FECM, evaluation sample 1970-1998**

		RMSE			**MSE relative to MSE of AR model**			
h	**Var**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECMc**
	IP	0.007	0.99	1.00	1.01	1.04	1.01	0.93
1	CPI no food	0.002	0.98	0.94	0.96	0.99	0.91	0.90
	3m T-Bill	0.583	0.96	0.89	0.93	1.00	0.95	0.92
	IP	0.017	0.96	1.00	1.00	0.97	0.90	0.88
3	CPI no food	0.003	1.03	1.00	1.02	1.05	0.91	0.90
	3m T-Bill	1.230	0.93	0.90	0.93	1.09	0.90	0.91
	IP	0.029	0.97	1.00	0.99	0.93	0.88	0.88
6	CPI no food	0.003	1.04	1.07	1.09	1.17	0.96	0.99
	3m T-Bill	1.674	0.90	0.89	0.95	1.15	0.93	0.96
	IP	0.049	0.99	1.01	1.00	0.88	1.03	1.01
12	CPI no food	0.003	1.03	1.02	1.03	1.00	0.87	0.88
	3m T-Bill	2.127	0.96	0.94	0.99	1.15	1.01	1.13
	IP	0.065	1.00	1.02	1.02	0.94	1.19	1.19
18	CPI no food	0.004	1.03	1.02	1.04	1.04	0.91	0.92
	3m T-Bill	2.688	0.98	0.96	0.97	0.99	1.03	1.16
	IP	0.076	1.01	1.03	1.03	1.06	1.34	1.35
24	CPI no food	0.004	1.05	1.08	1.08	0.96	0.86	0.85
	3m T-Bill	3.085	1.00	0.98	1.00	0.94	1.13	1.22
		AR	1.81	4.38	3.94			
	Lags	FAR	1.85	1.89	1.84			
		VAR	FAVAR	ECM	FECM	FECMc		
		1.61	1.59	1.30	0.31	0.31		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			2.37	1.00	3.00	3.00	3.00	6.00

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 - log of industrial production index, CPI no food - inflation of consumer prices without food, 3m T-Bill - 3-month T-Bill yield.

FAVAR, ECM, FECM and FECMc relative to that of the AR model. The FECM does best in 6 out of the 18 cases. This relatively poor performance is mostly determined by the fact that it is never the best method for industrial production. This result is in line with the rather poor performance of factor models for forecasting GDP growth in Germany, see Marcellino and Schumacher (2008).

For inflation and the interest rate, the FECM performs best in half the cases, with gains in forecasting precision relative to the benchmark AR model in some cases exceeding 50%. The ECM is the best performing model in only one case.

The model with the highest occurrence of best performance is the VAR, which is always the best for industrial production. It is also interesting to note that the FAVAR never produces the best forecast on average. The fact that the FECM outperforms the ECM in 10 out of 18 cases indicates the importance of factors in the analysis, and demonstrates that factors in the cointegration space proxy successfully for the cointegration relations that are otherwise missing in the small ECM. But comparison with the other models also shows that it is crucial how this information is included in the model. Although very indicative, we are aware that these findings may be heavily conditioned by the relative shortness of the sample (in the T dimension), leading to relatively short estimation and evaluation periods. For example, this could explain why the the FECM was not able to outperform the VAR for the real variable.

Table 10: **German monetary FECM, evaluation period 2002 - 2007**

h	**Var**	**RMSE of AR**	**FAR**	**MSE relative to MSE of AR model** VAR	**FAVAR**	**ECM**	**FECM**	**FECMc**
	IP	0.011	1.05	0.94	1.05	1.15	1.51	1.05
1	CPI no food	0.001	1.28	1.27	1.84	1.12	1.09	1.26
	3m IntRate	0.075	1.25	1.19	1.25	1.25	0.79	1.70
	IP	0.014	0.99	0.83	0.96	1.13	2.76	1.14
3	CPI no food	0.001	0.96	0.98	1.14	1.15	1.12	1.11
	3m IntRate	0.215	1.18	1.02	1.16	0.97	0.41	1.36
	IP	0.023	1.01	0.86	0.99	1.21	3.12	1.70
6	CPI no food	0.001	0.96	0.94	1.07	0.94	0.95	1.06
	3m IntRatel	0.463	1.16	1.04	1.14	1.07	0.50	1.29
	IP	0.039	1.00	0.92	1.00	1.22	2.59	2.30
12	CPI no food	0.001	1.05	1.12	1.25	0.69	0.73	0.79
	3m IntRate	0.918	1.06	1.00	1.04	1.54	0.73	1.84
	IP	0.053	1.00	0.95	1.00	1.19	2.02	2.18
18	CPI no food	0.002	1.00	1.08	1.19	0.67	0.63	0.65
	3m IntRate	1.330	1.00	0.98	0.99	2.19	1.03	2.12
	IP	0.066	1.00	0.97	1.00	1.22	1.62	2.01
24	CPI no food	0.002	1.00	1.08	1.25	0.77	0.68	0.76
	3m IntRate	1.639	1.00	1.00	0.99	2.42	1.10	2.44
		AR	1.00	5.22	1.26			
	Lags	FAR	1.18	1.95	1.01			
		VAR	FAVAR	ECM	FECM	FECMc		
		1.50	0.98	0.00	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			3.00	3.00	3.00	4.34	3.00	5.00

Notes: The FECM contains 2 I(1) factors, while an additional I(0) factor is added to the FECMc. The FAVAR includes 4 I(0) factors. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1991:1 - 2007:12, forecasting: 2002:1 - 2007:12. *Variables:* IP - log of industrial production index, CPI no food - inflation of consumer pricesx without food, 3m T-Bill - 3-month T-Bill yield.

5.4 Forecasting the term structure of government bond yields

Forecasting the term structure of interest rates has received considerable attention in the literature, and several methods have been proposed, see e.g. Carriero et al. (2009a) for a recent overview. In this subsection we construct a FECM based on a monthly dataset of maturities ranging from 1 to 120 months, taken from Gurkaynak et al. (2009). For the sake of brevity we focus on forecasting the 3-month, 2-year and 10-year interest rates for the US.

This example is also motivated by the theoretical consideration that since yields are linked by the term structure, we would expect to find only a handful of common trends driving them. The literature studying the yield curve often refers to the three factors driving the yield curve as the level factor, slope factor and the curvature factor. In our application, when considering extraction of $I(1)$ factors from the interest rates in levels, we find that 99% of overall data variability is captured by a single factor. However, to maintain comparability with the three-factor model, we introduce two additional stationary factors in the FECMc.

For the $I(0)$ factors included in the FAVAR and FAR, we also set their number to three. While here too the first principal component explains 98% of the variability in the data, we retain three factors for comparability with the FECMc.

In common with our approach in the previous examples, we also construct AR, VAR

Table 11: **Forecasting interest rates at different maturities, evaluation period 2000 - 2007**

h	Yield	RMSE of AR	MSE relative to MSE of AR model FAR	VAR	FAVAR	ECM	FECM	FECMc
1	3-month	0.214	0.95	1.16	1.18	0.84	0.82	0.78
	2-year	0.284	1.04	1.01	1.01	1.05	0.99	1.03
	10-year	0.261	1.00	1.00	1.00	1.03	1.01	1.05
3	3-month	0.495	1.00	1.16	1.19	0.81	0.73	0.68
	2-year	0.535	1.02	1.02	1.03	1.11	1.02	1.09
	10-year	0.408	1.00	1.00	1.00	1.05	1.04	1.16
6	3-month	0.896	1.02	1.12	1.12	0.84	0.74	0.71
	2-year	0.827	1.03	1.01	1.01	1.12	0.98	1.05
	10-year	0.507	1.00	1.00	1.00	1.11	1.07	1.32
12	3-month	1.651	1.01	1.06	1.06	0.96	0.85	0.76
	2-year	1.396	1.02	1.01	1.02	1.10	0.97	1.00
	10-year	0.729	1.00	1.00	1.00	1.02	0.95	1.33
18	3-month	2.251	1.01	1.03	1.03	1.08	1.01	0.84
	2-year	1.922	1.01	0.99	0.99	1.08	1.00	0.94
	10-year	0.879	1.00	1.00	1.00	0.97	0.89	1.28
24	3-month	2.702	1.00	1.02	1.02	1.19	1.15	0.87
	2-year	2.306	1.01	0.99	0.99	1.13	1.10	0.91
	10-year	0.946	1.00	1.00	1.00	1.02	0.96	1.38
		AR	1.49	1.00	0.00			
	Lags	FAR	0.94	0.76	0.00			
		VAR	FAVAR	ECM	FECM	FECMc		
		0.12	0.00	0.00	0.00	0.00		
				ECM			FECM	
	Cointegration rank		mean	min	max	mean	min	max
			2.00	2.00	2.00	3.00	3.00	3.00

Notes: The FECM contains one I(1) factor, while two I(0) factors are added to FECMc. The FAVAR contains three factors. Cheng and Phillips (2008) cointegration test and lag selection based on BIC. Data: 1985:1 - 2007:12, forecasting: 2000:1 - 2007:12
Variables: levels of yields at 3-month, 2-year and 10-year horizons.

and ECMs that are all based on the observable variables only. Estimation of the models begins in 1985 to avoid potential problems with model instability in the first half of the 1980s. The sample for forecast evaluation is set to 2000:1 - 2007:12.

Table 11 shows the substantial efficacy of the FECM and FECMc approach, since these models provide the best forecasts in 14 out of 18 cases. For the remaining 4, AR is best (or joint-best) and three of these rates are the 10-year yields at $h = 1, 3$ and 6. Some of the gains provided by FECM or FECMc are indeed quite substantial in relation to the competing models. In addition, the fact that the FECM always outperforms the ECM clearly indicates the importance of inclusion of information embedded in the factors for forecasting the yield curve. Similarly, the fact that the FECM outperforms the FAVAR 12 out of 18 times indicates that taking explicit account of the information contained in the factors for the long run significantly increases the forecasting precision of the yield curve.

5.5 Forecasting exchange rates

Our final empirical example focuses on forecasting nominal exchange rates. It is well known that beating a random walk, or more generally an AR model, in forecasting exchange rates is a tough challenge, see for example Engel and West (2005) for a theoretical explanation. However, Carriero et al. (2009b) have shown that a cross-section of exchange rates can contain useful information. We now reconsider this issue within the framework

Table 12: **Forecasting nominal exchange rates against USD, evaluation period 2002 - 2008**

h	**Currency**	**RMSE of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**	**FECMc**
			MSE relative to MSE of AR model					
1	EURO	0.025	1.05	1.00	1.03	1.03	0.98	1.10
	JAPY	0.025	1.13	1.00	1.08	1.02	1.00	1.11
	GBP	0.023	1.06	1.00	1.04	1.04	0.99	1.11
3	EURO	0.049	1.04	1.00	1.04	1.03	0.95	1.16
	JAPY	0.045	1.04	1.00	1.03	1.05	0.99	1.15
	GBP	0.037	1.05	1.00	1.03	1.09	0.95	1.31
6	EURO	0.077	1.00	1.00	1.00	0.97	0.93	1.05
	JAPY	0.064	1.00	1.00	1.00	1.06	0.99	1.03
	GBP	0.055	1.00	1.00	0.99	1.02	0.91	1.29
12	EURO	0.127	1.00	1.00	1.00	0.93	0.90	1.11
	JAPY	0.083	1.00	1.00	1.00	1.12	1.00	0.88
	GBP	0.082	1.00	1.00	1.00	0.97	0.87	1.42
18	EURO	0.175	1.00	1.00	1.00	0.82	0.88	1.01
	JAPY	0.110	1.01	1.00	1.01	1.01	1.04	1.14
	GBP	0.107	1.00	1.00	1.00	0.83	0.90	1.34
24	EURO	0.226	1.00	1.00	1.00	0.81	0.88	0.95
	JAPY	0.138	1.01	1.00	1.00	0.92	1.05	1.13
	GBP	0.131	1.00	1.00	1.00	0.75	0.88	1.29
		AR	0.00	0.00	0.00			
	Lags	FAR	0.58	0.67	0.58			
		VAR	FAVAR	ECM	FECM	FECMc		
		0.00	0.58	0.00	0.00	0.00		
				ECM			FECM	
Cointegration rank			mean	min	max	mean	min	max
			1.00	1.00	1.00	1.00	1.00	1.00

Notes: Both FECM and FAVAR contain one factor. Cheng and Phillips (2008) cointegration test and lag selection based on BIC information criterion. Data: 1995:1 - 2008:4, forecasting: 2002:1 - 2008:4

of our FECM approach.

We focus on three key bilateral exchange rates: euro exchange rate to dollar (EUR), Japanese yen exchange rate to the dollar (YEN), and pound sterling exchange rate to the dollar (GBP). The data sample in this application is the shortest of all the examples, consisting of monthly observations from 1995:1 - 2008:4. The period over which we evaluate the relative forecasting performance of the models is 2002:1 - 2008:4.

As was the case for the government bond yield example, only one factor is needed to explain a very large share of overall data variability. In the $I(1)$ case this share is 98%, while it is 88% in the $I(0)$ case. For this reason we set the number of factors both in the FECM and FAVAR to one.

Table 12 reports the MSEs relevant for the comparison of the models. FECM (or FECMc) is again by far the dominant method, providing the lowest MSEs (relative to AR) in 12 out of 18 cases, (in one case tied with the AR and the VAR) with gains of up to 13% over the AR which would be considered fairly large within the context of exchange rate forecasts. The ECM is the best model on 5 occasions with AR (tied with VAR) accounting for the remaining case. The ECM does best at the longer forecast horizons of 18 and 24, while the FAVAR never performs the best on average. The reasoning about the importance of cointegration and factors is very similar to the other examples where the FECM provided significant gains in forecasting precision.

6 Conclusions

The FECM, introduced by Banerjee and Marcellino (2009), offers two important advantages for modelling in a VAR context. First, inclusion of factors proxies for missing cointegration information in a standard ECM, and hence relaxes the dependence of ECMs on a small number of variables of interest. This dependence is in principle also relaxed by FAVAR models estimated on stationary data. The FECM, however, allows for the error-correction term in the equations for key variables under analysis, which prevents errors from being non-invertible moving average processes (and therefore difficult to approximate by long-order VARs), and avoids omitted variables bias. This paper confirms that both these features of the FECM also affect forecasting performance. From a theoretical point of view, since the FECM nests the FAVAR (and the ECM), it can be expected to provide better forecasts unless either the error correction terms or the factors are barely significant, or their associated coefficients are imprecisely estimated due to small sample size.

By means of extensive Monte Carlo simulations we demonstrate that the FECM consistently improves on other common models when error correction is present in the data and where inclusion of factors significantly increase the information content of the models. For the simpler DGP discussed in Section 4.1, the Monte Carlo results confirm the theoretical findings for sample sizes common in empirical applications. The FECM appears to dominate the FAVAR in all cases, even when the FECM is not the DGP but cointegration matters. However, the simulations also indicate that the gains shrink rapidly with the forecast horizon. For the more elaborate DGP, in Section 4.2, the results show that in empirically relevant situations the strength of the error correction mechanism again matters in determining the ranking of the alternative forecasting models. While the FECM remains better than the FAVAR in most of the cases, simpler models such as an ECM or even an AR can become tough competitors when the explanatory power of the error correction terms and/or of the factors is reduced or the sample size is not large.

It is clear in considering these simulation results that several issues are important here, including the role of considerable amounts of additional information incorporated *via* the factors, of cointegration and the strength of adjustment to disequilibrium, and the length of the forecasting horizons. Assessing the relative roles of cointegration and of the factors, and disentangling their effects, is not straightforward when models misspecified to some degree are compared. This is also the reason why the relative rankings of the models are not always clear-cut, and why the forecasting performance of the FECM should be also evaluated in a large set of empirical applications.

We have considered four main economic applications: forecasting a set of key real and nominal macroeconomic variables, evaluating extened versions of small scale monetary models, forecasting the term structure of interest rates, and assessing the merits of alternative exchange rate forecasts. In all cases we have considered univariate and small

Table 13: **Summary of empirical results**

		Ocurrence of best perfomance					
Model	**Out of**	**FECM**	**FECMc**	**FAVAR**	**ECM**	**VAR**	**FAR**
US real 85-03	24	6	8	1	1	7	0
US real 60-98	24	12	3	4	0	0	1
US nominal 85 - 03	24	0	0	0	18	1	0
US nominal 60 - 98	24	15	3	0	6	0	0
US 3-var 85-03	18	5	7	4	0	2	0
US 3-var 60-98	18	5	5	0	3	4	0
Germany 3-var	18	6	0	1	1	8	1
Interest rates	18	6	8	2	0	1	2
Exchange rates	18	11	1	0	5	1	0

		Importance of:			
		Cointegration		Factors	
	Out of	**FECM< FAVAR**	**ECM< VAR**	**FECM< ECM**	**FAVAR< VAR**
US real 85-03	24	14	13	18	2
US real 60-98	24	16	2	21	16
US nominal 85 - 03	24	7	22	0	0
US nominal 60 - 98	24	23	15	18	1
US 3-var 85-03	18	13	6	15	10
US 3-var 60-98	18	11	7	13	4
Germany 3-var	18	10	5	10	1
Interest rates	18	12	5	18	4
Exchange rates	18	15	8	12	7

multivariate models, with and without cointegration, and with or without factors. The factors summarize the information in large sets of variables, for different countries and periods of time. Based on Section 5 and Table 13, the following summary of the empirical results may be offered.

For forecasting the real variables for the Unites States, the FECM (or FECMc) is systematically better than the FAVAR and the ECM over both the samples considered. This is not necessarily true for the nominal variables, where the results are more sample-dependent. While the 1960 - 1998 sample reinforces the message of dominance of FECM methods, the more recent 1985 - 2003 dataset shows the ECM to be the dominant model, with FECM still beating FAVAR. As noted above, this finding is related to the decrease of importance of factors in forecasting for recent periods, also noted by D'Agostino et al. (2007). The overall picture however, taking both real and nominal variables into account over the two periods, remains very favourable for the use of FECM methods.

The results of the forecasting exercise based on the monetary model of the US offers unmitigated support for the use of FECMs in forecasting IP and CPI inflation. Moreover, for these variables, the ECM itself, while not providing the best model, dominates the models that do not make use of long-run information. Therefore, the usefulness of factors and cointegration, the underpinnings of the FECM approach, is again confirmed. The results for the interest rate variable however do not show much promise for the use of FECMs. This finding depends on the choice of the information set, and it is in fact reversed in the term structure example.

The monetary system using German data offers some interesting insight into working with FECMs in rather short samples. As noted in Section 5.3, in this example the model with the highest occurrence of best performance is the VAR. Here, while both factors and cointegration are important (as reflected in the dominance of the FECM over ECM and the FECM over FAVAR respectively), it appears that accounting for these features in the data may not always be sufficient. In other words, cointegration or factors per se may not increase the forecasting precision of models. It is only when information in the factors bears upon the long-run properties of the data that forecasting is benefited by including such information. As discussed in the theoretical analysis, and particularly with reference to the Monte Carlo exercise in Section 4.2, simpler models than the ECM or the FECM can become tough competitors when the explanatory power of the error correction terms and/or of the factors is reduced or the sample size is not large. The issue of sample size is one that has substantial relevance in the context of the German dataset.

The empirical example on the term structure of government bond yields allows us to return to the issue of forecasting interest rate variables. For this dataset, the results on the use of FECM methods are extremely promising and the gains in forecasting precision are significant. Unlike the monetary system for Germany, the importance of the inclusion of equilibrium information contained in the factors is clear. Taking the results of the monetary system for the US into account, a coherent picture also emerges of the crucial role of the information set and the sample used to construct the forecasts. The trade-offs evident from the theory and the simulations are present with vibrant force in the empirical implementations.

The final example on forecasting exchange rates again shows the FECM as the best model by far. The reasons are similar to the other cases where the FECM performed well and reinforce the findings gained from the previous examples.

The results of the paper also show several interesting nuances and tradeoffs to be investigated further, for example related to the role of structural breaks or to the temporal versus cross-sectional coverage of the dataset. In addition, since forecasts are the basic ingredient in the computation of impulse response functions, the performance of structural factor augmented error correction models also deserves investigation.

To conclude, the theory, simulation and empirical results taken together give us excellent grounds for optimism concerning the usefulness of long-run information captured through the factors and the efficacy of factor-augmented error correction models.

References

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

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

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

[4] Bai, J, and S. Ng (2006). Confidence intervals for diffusion index forecasts with a large number of predictors and inference for factor-augmented regressions. *Econometrica,* **74**, 1133-1150.

[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 (2008). Forecasting macroeconomic variables using diffusion indexes in short samples with structural change, in M. E. Wohar and D. E. Rapach, (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty.* Amsterdam: Elsevier, 149-194.

[7] 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.

[8] Carriero, A., G. Kapetanios and M. Marcellino (2009a). A shrinkage approach to yield curve forecasting. Mimeo.

[9] Carriero, A., G. Kapetanios and M. Marcellino, (2009b). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting.* Forthcoming.

[10] Cheng, X. and P. C. B. Phillips (2008). Semiparametric cointegrating rank selection. Cowles Foundation Discussion Paper No. 1658.

[11] Clements, M.P. and D.F. Hendry (1995). Forecasting in cointegrated systems. *Journal of Applied Econometrics*, **10**, 127-146.

[12] D'Agostino, A. , D. Giannone and P. Surico (2007). (Un)predictability and macroeconomic stability," CEPR Discussion Papers 6594, C.E.P.R. Discussion Papers.

[13] Diebold, F. X., and Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, **130**, 337–364.

[14] Eickmeier, S. and C. Ziegler (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27, 237-265.

[15] Engel, C.M. and K.D. West (2005). Exchange rate and fundamentals. *Journal of Political Economy*, **113**, 485-517.

[16] Engle, R.F. and C.W. Granger (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, **55**, 257-276.

[17] Engle, R.F. and S. Kozicki (1993). Testing for common features. *Journal of Business and Economic Statistics*, **11**, 369-390.

[18] Favero, C., M. Marcellino and F. Neglia (2005). Principal components at work: the empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, **20**, 603-620.

[19] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000). The generalized dynamic-factor model: Identification and Estimation. *Review of Economics and Statistics*, **82**, 540-554.

[20] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005). The generalized dynamic factor model: One sided estimation and forecasting. *Journal of the American Statistical Association*, **100**, 830-840.

[21] Gürkaynak, R.S., A. Levin and E. Swanson (2009). Does inflation targeting anchor long-run inflation expectations? Evidence from long-term bond yields in the U.S., U.K., and Sweden. *Journal of the European Economics Association*. Forthcoming.

[22] Hall, D.A., H.M. Anderson and C.W.J. Granger (1992). A cointegration analysis of Treasury Bill yields. *Review of Economics and Statistics*, 75, 116-126.

[23] Hendry, D.F. (2006). Robustifying forecasts from equilibrium-correction systems. *Journal of Econometrics*, **135**, 399-426.

[24] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231-254.

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

[26] Marcellino, M., J. H. Stock and M.W. Watson (2006). A comparison of direct and iterated AR methods for forecasting macroeconomic series h-steps ahead. *Journal of Econometrics*, **135**, 499-526.

[27] Marcellino, M. and C. Schumacher (2008). Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP. CEPR Discussion Papers 6708.

[28] Meese, R., and K. Rogoff (1983). Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? *Journal of International Economics*, **14**, 3-24.

[29] Rudebusch, G.D. & L. E. O. Svensson (1998). Policy rules for inflation targeting. Proceedings, Federal Reserve Bank of San Francisco, March.

[30] Stock, J.H. and M.W. Watson (1998). Testing for common trends. *Journal of the American Statistical Association*, **83**, 1097-1107.

[31] Stock, J.H. and M.W. Watson (2002a). Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association*, **97**, 1167-1179.

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

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

[34] Stock, J.H. and M.W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking,* **39,** 3 - 33.

Appendix A: Additional results of Monte Carlo experiments

Table 14: **Monte Carlo results - DGP corresponding to FECM with real variables, c = 0.75**

		RMSE	MSE relative to MSE of AR model				
h	**Var**	**of AR**	**FAR**	**VAR**	**FAVAR**	**ECM**	**FECM**
	1	0.005	1.09	0.94	1.01	0.97	0.90
1	2	0.007	1.02	0.95	1.00	0.97	0.93
	3	0.001	1.14	1.15	1.45	1.04	0.92
	4	0.009	1.04	1.00	1.06	1.12	1.08
	1	0.011	1.19	0.91	1.05	0.97	0.77
3	2	0.011	1.01	0.88	0.93	0.83	0.71
	3	0.003	1.47	1.29	1.73	1.05	0.71
	4	0.013	1.01	0.95	0.99	1.08	0.90
	1	0.019	1.20	0.93	1.02	0.97	0.69
6	2	0.018	1.02	0.86	0.91	0.77	0.61
	3	0.006	1.54	1.40	1.68	1.06	0.64
	4	0.019	1.01	0.92	0.95	1.13	0.83
	1	0.034	1.12	0.95	1.01	0.99	0.76
12	2	0.029	1.00	0.86	0.90	0.66	0.54
	3	0.012	1.50	1.41	1.53	1.04	0.74
	4	0.028	1.00	0.93	0.95	1.10	0.84
	1	0.046	1.09	0.98	1.02	0.97	0.77
18	2	0.037	1.00	0.90	0.94	0.71	0.59
	3	0.019	1.38	1.33	1.41	0.97	0.77
	4	0.036	1.00	0.97	0.97	1.13	0.95
	1	0.064	1.10	0.98	1.01	1.12	0.84
24	2	0.049	1.01	0.90	0.92	0.82	0.59
	3	0.028	1.35	1.32	1.36	1.09	0.83
	4	0.045	1.01	0.95	0.96	1.22	0.93
		AR	1.62	1.03	2.76	1.02	
	Lags	FAR	0.55	0.62	0.93	0.72	
		VAR	FAVAR	ECM	FECM		
		0.94	0.47	0.11	0.08		
			ECM			FECM	
Cointegration		mean	min	max	mean	min	max
rank		1.43	0.81	2.26	2.55	1.40	3.27

Notes: See Table 2.

Table 15: **Monte Carlo results - DGP corresponding to FECM with real variables, c = 0.50**

h	Var	RMSE of AR	MSE relative to MSE of AR model FAR	VAR	FAVAR	ECM	FECM
	1	0.005	1.04	0.97	1.03	0.98	0.96
1	2	0.007	1.02	0.98	1.03	1.01	1.03
	3	0.001	1.12	1.15	1.42	1.03	1.15
	4	0.009	1.03	1.03	1.10	1.14	1.12
	1	0.010	1.09	0.92	1.02	0.94	0.89
3	2	0.011	1.01	0.94	0.98	0.93	0.90
	3	0.003	1.46	1.37	1.74	1.16	1.16
	4	0.013	1.01	0.98	1.01	1.10	1.01
	1	0.016	1.07	0.95	1.01	0.97	0.89
6	2	0.015	1.01	0.94	0.97	0.91	0.87
	3	0.005	1.50	1.37	1.55	1.10	1.10
	4	0.017	1.01	0.98	1.00	1.14	1.00
	1	0.029	1.06	0.99	1.03	0.99	0.89
12	2	0.025	1.00	0.95	0.98	0.84	0.81
	3	0.010	1.35	1.31	1.40	1.00	1.00
	4	0.026	1.01	0.99	1.00	1.14	0.96
	1	0.041	1.05	0.99	1.02	1.05	0.90
18	2	0.033	1.00	0.95	0.98	0.85	0.78
	3	0.016	1.38	1.35	1.43	1.13	1.03
	4	0.033	1.00	0.98	0.99	1.13	1.00
	1	0.052	1.03	0.99	1.01	1.08	0.92
24	2	0.041	1.00	0.96	0.98	0.91	0.81
	3	0.021	1.22	1.21	1.25	1.09	0.97
	4	0.038	1.01	0.99	1.00	1.26	0.98
		AR	0.94	0.91	2.51	1.03	
	Lags	FAR	0.45	0.53	0.83	0.71	
		VAR	FAVAR	ECM	FECM		
		0.71	0.23	0.12	0.05		
			ECM			FECM	
Cointegration		mean	min	max	mean	min	max
rank		1.18	0.54	1.98	1.19	0.35	2.21

Notes: See Table 2.

Appendix B: Lists of data

Table 16: **German dataset**

Short descr.	Tcode	Short desc.	Tcode
Prices		exp. turn. electr. eng.	5
PPI	5	dom. turn. electr. eng.	5
PPI w/o energy	5	exp. turn. veh. eng.	5
CPI	5	dom. turn. veh. eng.	5
CPI w/o energy	5	dom. orders interm. goods	5
exp. prices	5	exp. orders interm. goods	5
imp. prices	5	dom. orders cap. goods	5
oil price Brent	5	exp. orders cap. goods	5
Labour market		dom. orders cons. goods	5
unemployed	5	exp. orders cons. goods	5
unemp. rate	1	dom. orders mech. eng.	5
empl. and self-empl.	5	exp. orders mech. eng.	5
empl., short-term	5	dom. orders electr. eng.	5
prod. per emp.	5	exp. orders electr. eng.	5
prod. per hour	5	dom. orders veh. eng.	5
wages per empl.	5	exp. orders veh. eng.	5
wages per hour	5	ind. prod.	5
vacancies	5	**Construction**	
Financials		constr. ord. building	5
mon. mar. rate, overnight	1	constr. ord. civ. eng.	5
mon. mar. rate, 1 month	1	constr. ord. resid. building	5
mon. mar. rate, 3 month	1	constr. ord. non-res. building	5
bond yields, 1-2 years	1	hours build. constr.	5
bond yields, 5-6 years	1	hours civ. eng.	5
bond yields, 9-10 years	1	hours resid. build.	5
CDAX share price index	5	hours ind. build.	5
DAX share index	5	hours pub. build.	5
REX bond index	5	turnover build. constr.	5
exch. rate USD/DM	5	turnover civ. eng.	5
Comp. Ind.	5	turnover resid. build.	5
M1	5	turnover ind. build.	5
M2	5	turnover pub. build.	5
M3	5	prod. in construction	5
Manufacturing activity		**Miscellaneous**	
prod. interm. goods	5	CA: exports	5
prod. cap. goods	5	CA: imports	5
prod. cons. goods	5	CA: serv. imp.	5
prod. mech. eng.	5	CA: serv. exp.	5
prod. electr. eng.	5	CA: transf. in	5
prod. veh. eng.	5	CA: transf. out	5
exp. turn. interm. goods	5	HWWA raw mat. prices	5
dom. turn. interm. goods	5	HWWA raw mat. prices w/o energy	5
exp. turn. cap. goods	5	HWWA raw mat.prices indu. mat.	5
dom. turn. cap. goods	5	HWWA raw mat.prices: energy	5
exp. turn. cons. goods	5	new car registrations	5
dom. turn. cons. goods	5	new private car registrations	5
exp. turn. mech. eng.	5	retail sales turnover	5
dom. turn. mech. eng.	5		

Source: Bundesbank. Sample: 1991:1-2007:12
Transformation codes: 1 no transformation; 2 first difference; 3 second difference; 4 logarithm; 5 first difference of logarithm; 6 second difference of logarithm.

Table 17: **US dataset**

Code	Short desc.	Tcode	Code	Short desc.	Tcode
a0m052	PI	4	HSBR	BP: total	4
A0M051	PI less transfers	4	HSBNE	BP: NE	4
A0M224R	Consumption	4	HSBMW	BP: MW	4
A0M057	M and T sales	4	HSBSOU	BP: South	4
A0M059	Retail sales	4	HSBWST	BP: West	4
IPS10	IP: total	4	PMI	PMI	1
IPS11	IP: products	4	PMNO	NAPM new ordrs	1
IPS299	IP: final prod	4	PMDEL	NAPM vendor del	1
IPS12	IP: cons gds	4	PMNV	NAPM Invent	1
IPS13	IP: cons dble	4	A0M008	Orders: cons gds	4
IPS18	iIP:cons nondble	4	A0M007	Orders: dble gds	4
IPS25	IP:bus eqpt	4	A0M027	Orders: cap gds	4
IPS32	IP: matls	4	A1M092	Unf orders: dble	4
IPS34	IP: dble mats	4	A0M070	M and T invent	4
IPS38	IP:nondble mats	4	A0M077	M and T invent/sales	1
IPS43	IP: mfg	4	FM1	M1	5
IPS307	IP: res util	4	FM2	M2	5
IPS306	IP: fuels	4	FM3	M3	5
PMP	NAPM prodn	1	FM2DQ	M2 (real)	4
A0m082	Cap util	1	FMFBA	MB	5
LHEL	Help wanted indx	1	FMRRA	Reserves tot	5
LHELX	Help wanted/emp	1	FMRNBA	Reserves nonbor	5
LHEM	Emp CPS total	4	FCLNQ	C and I loans	5
LHNAG	Emp CPS nonag	4	FCLBMC	C and I loans	1
LHUR	U: all	1	CCINRV	Cons credit	5
LHU680	U: mean duration	1	A0M095	Inst cred/PI	1
LHU5	U $<$ 5 wks	4	FYFF	FedFunds	1
LHU14	U 5-14 wks	4	FYGM3	3 mo T-bill	1
LHU15	U 15+ wks	4	FYGT1	1 yr T-bond	1
LHU26	U 15-26 wks	4	FYGT10	10 yr T-bond	1
LHU27	U 27+ wks	4	PWFSA	PPI: fin gds	5
A0M005	UI claims	4	PWFCSA	PPI: cons gds	5
CES002	Emp: total	4	PWIMSA	PPI: int materials	5
CES003	Emp: gds prod	4	PWCMSA	PPI: crude materials	5
CES006	Emp: mining	4	PSCCOM	Commod: spot price	5
CES011	Emp: const	4	PSM99Q	Sens materials price	5
CES015	Emp: mfg	4	PMCP	NAPM com price	1
CES017	Emp: dble gds	4	PUNEW	CPI-U: all	5
CES033	Emp: nondbles	4	PU83	CPI-U: apparel	5
CES046	Emp: services	4	PU84	CPI-U: transp	5
CES048	Emp: TTU	4	PU85	CPI-U: medical	5
CES049	Emp: wholesale	4	PUC	CPI-U: comm.	5
CES053	Emp: retail	4	PUCD	CPI-U: dbles	5
CES088	Emp: FIRE	4	PUS	CPI-U: services	5
CES140	Emp: Govt	4	PUXF	CPI-U: ex food	5
A0M048	Emp-hrs nonag	4	PUXHS	CPI-U: ex shelter	5
CES151	Avg hrs	1	PUXM	CPI-U: ex med	5
CES155	Overtime: mfg	1	GMDC	PCE defl	5
aom001	Avg hrs: mfg	1	GMDCD	PCE defl: dlbes	5
PMEMP	NAPM empl	1	GMDCN	PCE defl: nondble	5
HSFR	HStarts: Total	4	GMDCS	PCE defl: services	5
HSNE	HStarts: NE	4	CES275	AHE: goods	5
HSMW	HStarts: MW	4	CES277	AHE: const	5
HSSOU	HStarts: South	4	CES278	AHE: mfg	5
HSWST	HStarts: West	4	HHSNTN	Consumer expect	1

Notes: Dataset extracted from Stock and Watson (2005). Sample: 1959:1-2003:12
Transformation codes: 1 no transformation; 2 first difference; 3 second difference; 4 logarithm; 5 first difference of logarithm; 6 second difference of logarithm.

Table 18: **Exchange-rate dataset**

Name	Code	Name	Code
1 AUSTRALIAN Dollar TO US Dollar	AUST	23 POLISH ZLOTY TO US Dollar	POLI
2 BRAZILIAN REAL TO US Dollar	BRAZ	24 SINGAPORE Dollar TO US Dollar	SING
3 CANADIAN Dollar TO US Dollar	CANA	25 SLOVAK KORUNA TO US Dollar	SLOV
4 CHILEAN PESO TO US Dollar	CHIL	26 SOUTH KOREAN WON TO US Dollar	SOUT
5 COLOMBIAN PESO TO US Dollar	COLO	27 SRI LANKAN RUPEE TO US Dollar	SRI
6 CZECH KORUNA TO US Dollar	CZEC	28 SWEDISH KRONA TO US Dollar	SWED
7 DANISH KRONE TO US Dollar	DANI	29 SWISS FRANC TO US Dollar	SWIS
8 EURO TO US Dollar	EURO	30 TAIWAN new Dollar TO US Dollar	TAIW
9 FINNISH MARKKA TO US Dollar	FINN	31 THAI BAHT TO US Dollar	THAI
10 UK č to USDollar	GBP	32 TURKISH LIRA TO US Dollar	TURK
11 HUNGARIAN FORINT TO US Dollar	HUNG	33 URUGUAYAN PESO FIN. TO US Dollar	URUG
12 INDIAN RUPEE TO US Dollar	INDI	34 TAIWAN NEW Dollar TO US Dollar	TAIW
13 IRISH PUNT TO US Dollar	IRIS	35 BRUNEI Dollar TO US Dollar	BRUN
14 ISRAELI SHEKEL TO US Dollar	ISRA	36 HONG KONG Dollar TO US Dollar	HONG
15 JAPANESE YEN TO US Dollar	JAPA	37 INDONESIAN RUPIAH TO US Dollar	INDO
16 MALTESE LIRA TO US Dollar	MALT	38 SOUTH KOREAN WON TO US Dollar	SOUT
17 MEXICAN PESO TO US Dollar	MEXI	39 KUWAITI DINAR TO US Dollar	KUWA
18 NEW ZEALAND Dollar TO US Dollar	NEWZ	40 LEBANESE č TO US Dollar	LEBA
19 NORWEGIAN KRONE TO US Dollar	NORW	41 NEW GUINEA KINA TO US Dollar	NEWG
20 PAKISTAN RUPEE TO US Dollar	PAKI	42NIGERIAN NAIRA TO US Dollar	NIGE
21 PERUVIAN NUEVO SOL TO US Dollar	PERU	43 SAUDI RIYAL TO US Dollar	SAUD
22 PHILIPPINE PESO TO US Dollar	PHIL		

Sources: WMR/Reuters, Global Trade Information Services
and the New York FED. Sample: 1995:1-2008:4.
Transformation codes: All series were logged and treated as I(1).

Authors contacts:

Igor Masten
European University Institute,
University of Ljubljana,
Kardeljeva pl. 17
1000 Ljubljana
Slovenia
Email: igor.masten@ef.uni-lj.si

Anindya Banerjee
Department of Economics
University of Birmingham
Edgbaston, Birmingham, B15 2TT
United Kingdom
Email: a.banerjee@bham.ac.uk

Massimiliano Marcellino
European University Institute,
Bocconi University and CEPR
Via della Piazzuola 43
I-50133 Florence
Email: massimiliano.marcellino@eui.eu