On factor-augmented univariate forecasting

Abdou-Aziz Niang[*] and Diègo Legros[†] and Marie-Claude Pichery[‡]

February 1, 2015

First draft

Abstract: The question we examine in this paper is the problem of improvement of univariate forecasting when the variable of interest belongs to a panel with dependant across units. We analyse in possibly nonstationary framework, to what extent forecast based on the augmented univariate process implied by a factor model can show substantial advantages in terms of expected gains, with respect to simple univariate model. Moreover we consider identical autoregressive (AR) roots over the cross sections. Analyses are done theoretically and on the basis of Monte Carlo simulations. Our results show that in the general case where non stationarity is allowed, substantial forecast error reduction of the univariate process can be achieved by simply augmenting each individual time series with its idiosyncratic factor.

Keywords: Panel data, common factors, Forecasting.
JEL Classification: C2, C3, C5.

[*]**L**aboratoire de Recehrche Économique et Sociale, University of Ziguinchor, BP 523 Ziguinchor - SENEGAL, Phone: +221 33 938 85 80 Fax: +221 33 992 50 05, E-mail: aniang@univ-zig.sn

[†]**L**aboratoire d'**É**conomie et de **G**estion, University of Burgundy, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 3520 Fax: +33 (0)3 8039 5443, E-mail: diego.legros@u-bourgogne.fr

[‡]**L**aboratoire d'**É**conomie et de **G**estion, University of Burgundy, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 3520 Fax: +33 (0)3 8039 5443, E-mail: marie-claude.pichery@u-bourgogne.fr

1 Introduction

In recent years, there is an important upsurge of forecasting methods using a large number of predictors. These methods are generally based on factor models that lead to parsimonious econometric structure when working with large dimensional dataset characterized by important cross-sectional correlations. In factor models, each variable is assumed to be the sum of two components: a common component driven by a small number of latent common factors and an idiosyncratic component. Pioneering works that use common and idiosyncratic representation in modeling macroeconomic aggregates subject to strong co-movements were initiated by Burns and Mitchelles (1946) followed by Geweke (1977) and Sargent and Sims (1977). However, in more recent years, important contributions on this field are realized by Stock and Watson (1998, 2002), Forni et al. (2000), Bai and Ng (2002) among others.

The idea which consists of using for forecasting purposes, relevant information set extracted from a large panel of potentially useful variables and summarized via a common component, is due to authors like Stock and Watson (1998, 2002), Forni et al. (2004), Marcellino et al. (2003), Peña and Poncela (2004) etc. If the common factor is directly augmented to the univariate process, this provides what is known as factor-augmented univariate forecast. This procedure has become quite popular since the work initiated by Stock and Watson (1998) and a large body of research in the literature of forecasting has corroborated its relevance (see for example D'Agostino and Giannone 2012, Bai and Ng 2005).

In line with these researches, the main question we address in this paper is the problem of the improvement of univariate forecasting which tends to turn out bad when the variable of interest belongs to a panel of time series which are highly dependant across unit. Focusing on the role of the idiosyncratic factor, we analyse in possibly nonstationary framework, to what extent forecast based on the augmented univariate process implied by a factor model can show substantial advantages in terms of expected gains, with respect to simple univariate model. Following Peña and Poncela (2004), we consider identical autoregressive (AR) roots over the cross sections. Analyses are done theoretically and on the basis of Monte Carlo simulations. Then, on the other hand, forecast of annual GNP growth of some European countries is considered. It is found that in the general case where non stationarity is alllowed, substantial forecast error reduction of the univariate process can be achieved by simply augmenting each individual time series with its idiosyncratic factor.

The rest of the paper is organized as follows. Section 2 sets up the framework of our analysis. Section 3 provides an analysis on the importance of such a procedure in forecasting exercise, and then presents the forecast model used and estimations issues. In Sections 4 and 5 we respectively analyse the forecast performance of the model and present some Monte Carlo results.

2 Framework

Let y_t be a stochastic variable. The corresponding autoregressive process augmented with a single factor can be defined over the sample period $t = 1, ..., T$ as follows

$$y_t = \alpha + \beta f_{t-1} + \rho y_{t-1} + \xi_t \tag{1}$$

where $|\rho| < 1$, α is a constant term and ξ_t the regression error. This representation corresponds to the diffusion index model defined by Stock and Watson (2002) in the case where the number of factors is one. The series f_t, referred to as the unobserved common factor is related to x_t, a large panel of stationary time series. The relation is given by the following factor struture[1]

$$x_{jt} = \mu_j + \gamma_j f_t + z_{jt} \qquad \forall\ j = 1, ..., J \tag{2}$$

where μ_j is an individual fixed effect and z_{jt} is a serially un-correlated idiosyncratic factor. In macroeconomic panels, the common factor can be viewed as an economy-wide shock, affecting all aggregates with heterogeneous intensities (Kabundi and Loots, 2007) . Thus, in such a case the panel of x_{jt} series can include macroeconomic indicators such that industrial production, interest rate, inflation etc. The coefficient γ_j is the factor loading which give a measure of the contribution of the j-th individual to the common shock. The factor-augmented univariate model (1) is relevant only if the autoregressive process of y_t can be explained by the latent common factor which follows the dynamic stationary vector process

$$f_t = \varphi f_{t-1} + \eta_t \tag{3}$$

where $|\varphi| < 1$, $\eta_t \sim iid\left(0, \sigma_\eta^2\right)$ and $E\left(\eta_t \eta_\tau^{'}\right) = 0$ if $t \neq \tau$. In addition $\text{cov}\left(z_{jt}, \eta_\tau\right) = 0$ for all j, t and τ. The vector representation of the single factor model is

$$x_t = \mu + \gamma f_t + z_t \tag{4}$$

with $x_t = \left(x_{1t}, ..., x_{Jt}\right)'$, $\mu = \left(\mu_1, ..., \mu_J\right)'$ and $\gamma = \left(\gamma_1, ..., \gamma_J\right)'$. The common and idiosyncratic factors are assumed to be uncorrelated and have zero mean. Also, there is independence between the factor loadings and the common factor so that $\{\gamma_j\}$, $\{f_t\}$ and $\{z_{jt}\}$ are three independent groups. The assumption that the vector $z_t = \left(z_{1t}, ..., z_{Jt}\right)'$ is mutually uncorrelated implies that the covariance matrix of the idiosyncratic factor (Σ_z) is a diagonal matrix, $\Sigma_z = diag\left(\sigma_{z1}^2, ..., \sigma_{zJ}^2\right)$. This latter assumption is associated with what is called in the literature the strict factor model. The covariance matrix of the observed series is decomposed into two components

$$\Sigma = \gamma V \gamma' + \Sigma_z \tag{5}$$

where V denotes the variance of the common factor.

[1] Notice that model (2) can be estimated using principal component analysis. In practice, the mean of the time series must be removed prior to principal component estimation.

The literature has shown how diffusion index model can be useful in forecasting exercise by improving the forecast accuracy of the series and outperforming many competing methods. The premise is that factor underlying common movement in a set of macroeconomic variables is a good predictor of the future value of a few key economic aggregates. Using model (1), the h-step ahead forecast proposed by Stock and Watson (2002) is $\hat{y}_{t+h} = \alpha_h + \beta_h f_t + \rho_h y_t$, where the parameters β_h and ρ_h depend on the forecast horizon and the forecast error is assumed to satisfy $E\left(\xi_{t+h} \middle| \{f_\tau, y_\tau\}_{\tau \leq t}\right) = 0$.

In the following methodoligical section, we conduct some analyses on the relevance of the forecast procedure stated here. We also extend the procedure to a possibly nonstationary framework and then use an alternative version of model (1). We are interesting, in Section 4, in its forecast performances[2].

3 Methodology

3.1 Setup

Suppose that y_t is a vector of n stochastic variables with high degrees of cross-section dependences. If n is enought large, the common factor can be extracted using the panel $\{y_{it}\}_{i=1,t=1}^{n,T}$, see for example Boivin and Ng (2005), Bai and Ng (2010). This suggests that in static form, y_{it} also admits a common factor representation and can be written in terms of equation (2)

$$y_{it} = \mu_i + \gamma_i f_t + z_{it}. \tag{6}$$

In a multi-country framework, this representation can be very attractive because it allows to model the dynamics of the macroeconomic variables such as GNP of each economy by controlling global and country specific effects.

To derive the factor-augmented univariate process from model (6) we allow the dynamic of f_t given in (3) to enter into y_{it} directly. After some straightforward developments, we get

$$y_{it} = (1 - \rho)\,\mu_i + \gamma_i\,(\varphi - \rho)\,f_{t-1} + \rho y_{i,t-1} + z_{it} - \rho z_{i,t-1} + \gamma_i \eta_t. \tag{7}$$

Equation (7) shows that when the time series admit a factor model representation the univariate regression is a special case of the one-factor-augmented model where the restriction $\varphi = \rho$ and/or $\gamma_i = 0$ is imposed. Also, we can see that the diffusion index forecast presented in the previous section can be regarded as a case where the common factor and the predicted series y_{it} are not restricted to have the same order of integration but where the restriction that η_t and $(1 - \rho L)\,z_{it}$ are unpredictable is considered, L being a lag operator. On the other hand, the error term of the factor-augmented autoregressive model can be seen as a pure factor structure in which η_t and $z_{it} - \rho z_{i,t-1}$ are respectively the common and idiosyncratic factors.

[2] With respect to the univariate forecasting.

For the purpose of this paper, we set $\varphi = \rho$ and then allow the nonstionarity of the process given in (7). Then, we define the Data Generating Processs (DGP) in line with Moon and Perron (2004) and Moon et al. (2007) PESARAN? by using the following single factor residual model

$$\begin{aligned} y_{it} &= \alpha_i + \rho y_{i,t-1} + \xi_{it} \\ \xi_{it} &= \gamma_i \eta_t + e_{it} \end{aligned} \tag{8}$$

where $\alpha_i = (1-\rho)\,\mu_i$ and the idiosyncratic error e_{it} follows a stationary Moving Average (MA) process such that $e_{it} = z_{it} - \rho z_{i,t-1}$. Notice that each individual has the same contribution (γ_i) to f_t and the factor residual η_t. This factor error structure may reflect different sources of unobserved individual-specific heterogeneity. For example, Cunha and Heckman (2007) argue that in a model of wage determination, γ_i corresponds to an unmeasured skill for the i-th individual, while η_t captures the vector of skill prices which changes intertemporally.

3.2 Forecasting Model

The assumptions on z_{it} imply that e_{it} is independently distributed across units, have mean zero and variance $\sigma^2_{ei} = \sigma^2_{zi}\,(1+\rho^2)$. In addition, with the assumption on η_t, the error ξ_{it} is serially uncorrelated. In fact, the factor residual model is not treated as a simple residual but is exploited for forecasting purpose of the univariate process. Thus, the one-step ahead forecast based on equation (8) yields

$$\hat{y}_{i,t+1} = \alpha_i + \rho y_{it} + \phi z_{it} \tag{9}$$

where $\phi = -\rho$. The associated prediction error is $\varepsilon_{i,t+1} = z_{i,t+1} + \gamma_i \eta_{t+1}$.

We will refer to equation (9) as the idiosyncratic factor-augmented (IFA) univariate forecast and will analyse its forecast accuracy. Notice that Peña and Poncela (2004) also considered the situation where $\varphi = \rho$ and compare prediction accuracy of factor model with respect to simple univariate model. In their paper, forecasts are carried out using a version of equation (6) where the intercept μ_i is nul. Their results show a gain in precision, in terms of the Mean Square Forecast Error (MSFE) in non stationary and stationary cases. However, Boivin and Ng (2005) argue that to achieve forecast error reduction one can simply augment the autoregressive model with the common and/or idiosyncratic factors. We will adopt this strategy by comparing simple and idiosyncratic factor-augmented univariate forecasts for $|\rho| < 1$ and for $\rho = 1$.

If in specifying model (8) we mistakenly ignore the effect of the factor residual, then the resulting univariate series with a constant term (c_i) will follows an ARMA (1,1) process (see for example Meddahi 2002, Peña and Poncela 2004, Banerjee et al. 2008) which can be expressed as follows

$$y_{it} = c_i + \rho y_{i,t-1} + v_{it} - \delta_i v_{i,t-1} \tag{10}$$

where $|\delta_i| < 1$ and $v_{it} \sim i.i.d.\,(0, \sigma_{vi}^2)\ \forall i$. Thus the one-step ahead forecast based on the univariate model (10) is

$$\hat{y}_{i,t+1} = c_i + \rho y_{it} - \delta_i v_{it} \tag{11}$$

with prediction error $v_{i,t+1}$. The forecast errors from model (11) are expected to dominate those from model (9) in absolute value. Indeed, Magnus and Pesaran (1989) show that this type of misspecification in univariate time-series forecasting can generate important statistical problems particularly in near-unit root case.

Remark 1. *For each i, the variance of the error v_{it} in the ARMA process (10) is given by*

$$\sigma_{vi}^2 = \sigma_{zi}^2 + \gamma_i^2 \sigma_\eta^2 (1 - \rho\delta_i)^{-1}. \tag{12}$$

Proof See Appendix. □

Building on this remark, we can establish the following relation by introducing the equation (24) given in the Appendix into (12)

$$\frac{\rho}{\delta_i} = 1 + \frac{\gamma_i^2 \sigma_\eta^2}{\sigma_{zi}^2\,(1 - \rho\delta_i)} > 1. \tag{13}$$

This last relation implies that $|\rho| > |\delta_i|$ and that in addition, these two parameters have the same sign. Moreover, the larger is γ_i^2, the greater ρ/δ_i. So, the gap between the values of the moving average parameter of the i-th individual and the pooled AR parameter depends to the sensitivity of the individual to the common information and that this difference exists as long as γ_i will be non-zero. If there is no restriction and that $\gamma_i = 0$ simply because the effect of the common factor is null, then the factor representtion of the residual ξ_{it} no more holds and the corresponding series follows a simple univariate autoregressive process. To ensure that the representation in (8) and (10) is valid for each i, one can assume as it is usual in the literature that the factor has nontrivial contribution to the dynamic of each time series, say $\gamma_i \neq 0\ \forall i = 1, ..., n$.

3.3 Estimations Issues

Above, we treated the parameters of the models as well as the latent element (z_{it}) as known, but in an empirical perspective they need to be estimated. At each step of the forecast, the estimation of the models IFA and ARMA (1,1) can respectively be summarized as follows.

- For the IFA model, we first estimate the parameters of the fixed-effect dynamic panel model with pooled autoregressive root. Then, principal components is used on the covariance matrix of the observed data demeaned beforehand and the estimator of the idiosyncratic elements z_{it} are collected. Indeed it is well known that when the cross-sectional dimension is large, principal component gives the estimators of the factors which are as good as if the true latent factors were observed (Doz et al. , 2012 and Breitung et al., 2006). Thus, the forecast error possibly implied by the estimation of the unobserved factors should be negligible.

- For the ARMA(1,1) model, we also begin by estimating the dynamic panel model and collect the estimated residuals, $\hat{\xi}_{it}$. Then, these residual are used to fit an ARMA(1,1) model and get the estimator of the MA(1) parameter.

In the next section, we derive and compare the mean square error of prediction associated to specifications (9) and (11) for further ahead forecast. The aim being to check theoretically how much forecast improvement the factor-augmented univariate model can provide. Furthermore, attention is also given to the difference between the nonstationary and the stationary cases.

4 Forecast Accuracy

Now suppose that one is interested by an h-steps ahead forecast of a vector of time series with information up to time t. This can be done by extending the prediction of the observed series defined in (9) to horizon h. Thus, proceding to a sequence of one-step ahead forecasts we obtain at $t+h$,

$$\hat{y}_{i,t+h} = \sum_{j=0}^{h-1} \rho^j \alpha_i + \rho^h y_{it} + \rho^{h-1} \phi z_{it}. \tag{14}$$

It is clear from equation (14) that idiosyncratic factor can play an important role on the precision of the forecast. However, in some cases it is overlooked and treated as irrelevant noise. For example, in a framework of integrated economies idiosyncratic movements are generally assumed to include future random shocks and measurement error and cnosidered as unforecastable. In doing so, modelers expect to capture a more reliable signal for policy makers and to prevent them from reacting to country-specific movements (see Breitung and Eickmeier, 2006). Another argument is that, although possibly shared by many units, idiosyncratic causes of variation can have their effect concentrated on a finite number of units and tending to zero as the individual dimension tends to infinity (see Altissimo et al., 2001). As we will see later, in the framework stated here, as well as common factors, idiosyncratic elements can also improve significantly forecast accuracy. We will see that, if we ignore their effects then the further ahead we forecast the less precise the simple univariate model becomes with respect to the augmented model.

Remark 2. *Let $\mathcal{MSFE}_i^{(1)}$ be the prediction mean square error based on (9). For an horizon h it can be established that*

$$\mathcal{MSFE}_i^{(1)} = \sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)(1 - \rho\delta_i)}{\delta_i} \sigma_{zi}^2. \tag{15}$$

Proof See Appendix. □

For the ARMA model (11), the h-steps ahead forecast with information up to time t is

$$\hat{y}_{i,t+h} = \sum_{j=0}^{h-1} \rho^j c_i + \rho^h y_{it} - \rho^{h-1} \delta_i v_{it} \tag{16}$$

As stressed by Meddahi (2002), this ARMA representation of the forecast model can obviously be useful in multi-step forecast since it corresponds to the analytical steady-state of the latent factor. For example, he note that the expected value of a latent variable like volatility can be easily obtained using the ARMA representation of its corresponding observed variable. However, for a more correct modeling of data with dynamic common factor structure, a factor model is needed to capture at least the persistence of the unobserved component.

Remark 3. *Let $\mathcal{MSFE}_i^{(2)}$ be the prediction mean square error of the univariate ARMA model (11). For an horizon h, we have*

$$\mathcal{MSFE}_i^{(2)} = \frac{(2\rho - \delta_i)}{\rho} \sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho \delta_i} \sigma_{zi}^2. \tag{17}$$

Proof See Appendix. □

From Remarks 2 and 3, we derive the result given in the above Propositon in which the length of the forecast horizon may take any value, small as well as large. This result implies that on the basis of the IFA model, we can obtain better forecast accuracy with respect to the ARMA model even by allowing nonstationarity of the time series, say by letting $\rho = 1$.

Proposition 1. *Let $\Delta_i = \mathcal{MSFE}_i^{(2)} - \mathcal{MSFE}_i^{(1)}$ be the measure of the forecast performance of the augmented univariate model (9), with respect to the simple univariate model (11). For the h-steps ahead forecast, we have:*

In the non stationary case ($\rho = 1$),

$$\Delta_i = (1 - \delta_i)\, \sigma_{zi}^2 > 0. \tag{18}$$

In the stationary case ($|\rho| < 1$),

$$\Delta_i = \rho^{2h-1} (\rho - \delta_i)\, \sigma_{zi}^2 > 0. \tag{19}$$

Proof See Appendix. □

The fact that we have $\Delta_i > 0 \; \forall i$ indicates that the passage from model (16) to (14) reduces the Mean Square Forecast Error significantly for each individual time series and thus, for the whole panel. It is worthwhile noting that the MSFE of model (14) is relatively far less than that of model (16), irrespective to the order of integration ($I(0)$ or $I(1)$) of the individual time series. But in the stationary case, when the horizon of prediction approaches infinity, the difference between both models vanishes. Also notice that as we can expect, the importance of the share of variance of the idiosyncratic factor also can be determinant. A larger value of σ^2_{zi} implies a greater forecast precision of the IFA model relatively to the simple univariate model.

5 Monte Carlo Study

To investigate the performance of the augmented-model, we conduct a set of simulations. We explore the extent of the forecast error in relation to the corresponding AR and ARMA models. Throughout, we first proceed by generating a fixed-effect dynamic panel model with a residual common factor structure

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \xi_{it}$$
$$\xi_{it} = \gamma_i \eta_t + e_{it},$$

where α_i , γ_i and η_t are $i.i.d$ $N(0,1)$ for all i and t. The idiosyncratic part of the residual common factor is generated according to $e_{it} = z_{it} - \rho z_{i,t-1}$ with $z_{it} \sim N(0,1)$. As we saw in Section ?, the static form of such dynamic process corresponds to a dynamic common factor model where z_{it} represents the idiosyncratic element. Using these simulated data, three models are compared:

(1) the naive forecast given by a simple AR(1) process

$$\hat{y}_{i,t+1} = \alpha_i^{AR} + \rho^{AR} y_{i,t},$$

(2) the ARMA(1,1) forecast

$$\hat{y}_{i,t+1} = \alpha_i^{ARMA} + \rho^{ARMA} y_{i,t} - \delta_i v_{i,t},$$

(3) the IFA forecast

$$\hat{y}_{i,t+1} = \alpha_i^{IFA} + \rho^{IFA} y_{i,t} + \phi z_{it}.$$

The MSFE associated with each of these specifications is examined according to the size of the cross-sectional dimension n, the value of the pooled autoregressive parameter ρ and the forecast horizons h. The total number of time observations generated at each replication is $T = t + h$. The value of t is fixed to 101 which in turn is used to compute forecasts for period $t + h$. Thus, the three models was estimated with data up to time t, then we generated one-step ahead forecasts. We reestimated the models adding one observation at the time and made new forecasts. Furthermore, unlike classical factor-augmented forecasting which is particularly interested in the common information for predicting the variable of interest, here we need to isolate this common factor considered as a nuisance for the prediction of the dynamic individual time series. However, in practice, the common factor as well as the idiosyncratic ones are unobserved and forecasters will extract estimates of these from the panel of observed variables. In the simulations, we used principal component analysis to extract the latent variables.

Each simulation run is carried out with 1,000 replications and new series are generated for each draw. Considering a vector of forecast horizons, H with H_j the jth element of H, forecast performance measured by the Mean Squared Forecast Error is computed for each H_j, as

$$\widehat{\mathcal{MSFE}} = \frac{1}{H_j} \sum_{h=1}^{H_j} (y_{i,t+h} - \hat{y}_{i,t+h})^2, \qquad i = 1, ..., n \ . \tag{20}$$

In fact, our analysis is based on the average square root of $\widehat{\mathcal{MSFE}}$ across the individual dimension and the total number of simulation draws. In Table 1, we report results for $n = 10, 20, 50$, $\rho = 0.2, 0.4, 0.6, 0.8, 1.0$ and $H = 1, 5, 10$ for each of the forecasting models.

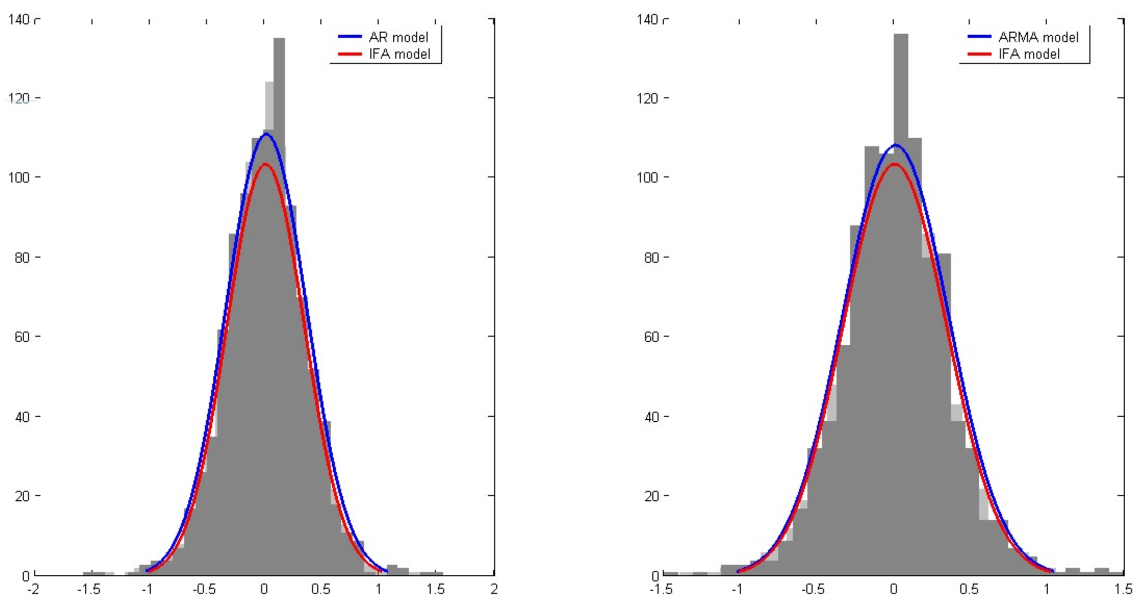

Figure 1: Distribution fit of forecast errors

As can be seen, the ARMA model outperforms the AR model which in turn is dominated by the IFA model. Furthermore, it appears that when the value of the

Table 1: Simulation results

Model	AR(1)			ARMA(1,1)			IFA		
ρ/h	1	5	10	1	5	10	1	5	10
					($n = 10$)				
0.2	1.3558	1.4116	1.4222	1.3604	1.4125	1.4226	1.3520	1.4109	1.4218
0.4	1.3745	1.4557	1.4733	1.3730	1.4551	1.4730	1.3592	1.4526	1.4718
0.6	1.4204	1.5442	1.5753	1.4078	1.5409	1.5736	1.3798	1.5355	1.5711
0.8	1.5002	1.7260	1.8202	1.4589	1.7157	1.8155	1.4022	1.7007	1.8083
1.0	1.6872	2.1982	2.7304	1.5430	2.0922	2.6578	3.1579	3.5147	3.9121
					($n = 20$)				
0.2	1.3624	1.4144	1.4223	1.3665	1.4151	1.4227	1.3587	1.4137	1.4220
0.4	1.3776	1.4490	1.4696	1.3761	1.4484	1.4694	1.3613	1.4458	1.4681
0.6	1.4400	1.5530	1.5791	1.4275	1.5495	1.5773	1.3962	1.5437	1.5746
0.8	1.5102	1.7339	1.8373	1.4691	1.7232	1.8324	1.4059	1.7069	1.8247
1.0	1.7016	2.1907	2.7295	1.5533	2.0808	2.6542	3.3083	3.6608	4.0570
					($n = 50$)				
0.2	1.3697	1.4139	1.4246	1.3741	1.4147	1.4250	1.3655	1.4131	1.4242
0.4	1.3871	1.4573	1.4724	1.3857	1.4566	1.4720	1.3706	1.4541	1.4708
0.6	1.4304	1.5451	1.5792	1.4161	1.5413	1.5774	1.3852	1.5355	1.5746
0.8	1.5160	1.7372	1.8350	1.4750	1.7267	1.8302	1.4099	1.7097	1.8220
1.0	1.7049	2.1819	2.7065	1.5561	2.0714	2.6328	3.2171	3.5235	3.9240

Notes: The values in the table correspond to the square root of $\widehat{\mathcal{MSFE}}$ for AR(1), ARMA(1,1) and IFA models computed using equation (20). The number of Monte Carlo repetitions is 1,000 and in each draw the data generated are used to estimate the three models.

autoregressive parameter approaches zero, this leads to an increase in the $\widehat{\mathcal{MSFE}}$ of the ARMA forecast relative to that of the simple AR. However, IFA model yields more accurate prediction both for near zero and near unit values of ρ. The benefits of IFA model can be seen clearly in the distributions of the forecast error plotted in Figure 1 which is obtained using simulated data. We set n=20, t=51 and $\rho = 0.8$ and used 1,000 replications of average values of one-step ahead forecast errors across the whole panel. The forecast errors distribution for the AR(1) and ARMA(1,1) models have slightly heavier tail in both right and left sides with a relatively more important tail on the right. This implies that both sides of these distributions produce high values at a greater rate than it would be expected from the IFA forecasting model. Finally, notice that the non-stationary case gives very mitigated results. Indeed, as stresed by Bai and Ng (2004), consistent estimation of the space spanned by the unobserved common factor and idiosyncratic factors is in fact not possible when the observed series are integrated.

6 Conclusion

In this paper we examine the problem of improvement of univariate forecasting in panel models with dependant across units. In an nonstationary framework, we derive forecasts based on an augmented univariate process implied by a factor model and show. Using Monte Carlo Experiments, we show that substantial forecast error reductions can be achieved by augmenting each individual time series with its idiosyncratic factor.

References

[1] Burns A. M., Mitchell W. C., (1946), Measuring Business Cycles, *National Bureau of Economic Research* New York.

[2] Geweke J., (1977), The dynamic factor analysis of economic time series, In: Aigner, D.J., Goldberger, A.S.(Eds.), Latent Variables in Socio-Economic Models. North-Holland, Amsterdam.

[3] Sargent T.J., Sims C.A., (1977), Business cycle modelling without pretending to have too much a priori economic theory, In: Sims, C.A. (Ed.), New Methods in Business Research. Federal Reserve Bank of Minneapolis, Minneapolis.

[4] Stock J. H., Watson M. H., (1998), Diffusion Indexes, *National Bureau of Economic Research*, no. 6702.

[5] Stock J. H., Watson M. W., (2002), Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association*, 97, 147-162.

[6] Forni M., Hallin M., Lippi M., Reichlin L., (2000), The Generalized Dynamic Factor Model: Identification and Estimation, *The Review of Economics and Statistics*, 82, 540-554.

[7] Bai J., Ng S., (2002), Determining the Number of Factors in Approximate Factor Models, *Econometrica*, 70,191-221.

[8] Forni M., Hallin M., Lippi M., Reichlina L., (2004), The generalized dynamic factor model consistency and rates, *Journal of Econometrics*, 119, 231-255.

[9] Banerjee A., Marcellino M., Masten I., (2008), Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change, *CEPR Discussion Papers* n° 6706.

[10] Pena D., Poncela P., (2004), Forecasting with nonstationary dynamic factor models, *Journal of Econometrics*, 119, 291-321.

[11] D'Agostino A., Giannone D., (2012), Comparing Alternative Predictors Based on Large-Panel Factor Models, Oxford Bulletin of Economics and Statistics, 74, 306-326.

[12] Boivin J., Ng S., (2005), Understanding and Comparing Factor-Based Forecasts, International Journal of Central Banking, vol. 1(3).

[13] Bai J., Ng S., (2010), Panel Unit Root Tests With Cross-Section Dependence: A Further Investigation, *Econometric Theory*, 26, 1088-1114.

[14] Moon H.R., Perron B., (2004), Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics*, 122, 81-126.

[15] Moon H.R., Perron B., Phillips P. C. B., (2007), Incidental Trends and the Power of Panel Unit Root Tests, *Journal of Econometrics*, 141, 416-459.

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

[17] Magnus J.R., Pesaran B., (1989), The exact multi-period mean-square forecast error for the first-order autoregressive model with an intercept, *Journal of Econometrics*, 42, 157-179.

[18] Doz C., Giannone D., Reichlin L., (2012), A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models, *The Review of Economics and Statistics*, 94, 1014-1024.

[19] Breitung J., Eickmeier S., (2006), Dynamic factor models, *AStA Advances in Statistical Analysis*, 90, 27-42.

[20] Altissimo F., Cristadoro R., Forni M., Lippi M., Veronese G., (2001), The construction of coincident and leading indicators for the euro area business cycle, *Banca d'Italia*, WP n 434.

Appendix A

Proof of Remark 1. *This remark can be shown using the same arguments as in the proof of part 1 of the Peña and Poncela's (2004) auxiliary lemma. First notice that from (8) and (10) we have*

$$\begin{aligned} y_{it} - \rho y_{i,t-1} &= c_i + v_{it} - \delta_i v_{i,t-1} \\ &= (1-\rho)\,\mu_i + z_{it} - \rho z_{i,t-1} + \gamma_i \eta_t. \end{aligned} \tag{21}$$

This implies that

$$\mathrm{Var}\left(c_i + v_{it} - \delta_i v_{i,t-1}\right) = \mathrm{Var}\left((1-\rho)\,\mu_i + z_{it} - \rho z_{i,t-1} + \gamma_i \eta_t\right). \tag{22}$$

With the independence of η_t to the idiosyncratic factor z_{it} for all lags and the absence of serial correlation in z_{it}, the right-hand part of equality (21) yields

$$\begin{aligned} \mathrm{Var}\left((1-\rho)\,\mu_i + z_{it} - \rho z_{i,t-1} + \gamma_i \eta_t\right) &= \mathrm{Var}\left(\gamma_i \eta_t\right) + \mathrm{Var}\left(z_{it} - \rho z_{i,t-1}\right) \\ &= \gamma_i^2 \sigma_\eta^2 + \sigma_{zi}^2 + \rho^2 \sigma_{zi}^2. \end{aligned}$$

Then, with $\mathrm{Var}\left(c_i + v_{it} - \delta_i v_{i,t-1}\right) = \sigma_{vi}^2 + \delta_i^2 \sigma_{vi}^2$ *equation (22) gives*

$$\gamma_i^2 \sigma_\eta^2 + \sigma_{zi}^2 + \rho^2 \sigma_{zi}^2 = \sigma_{vi}^2 + \delta_i^2 \sigma_{vi}^2. \tag{23}$$

Now using relation (21) and equating first-order autocovariances, we get

$$\begin{aligned}
&\operatorname{cov}\left(y_{it}-\rho y_{i,t-1}, y_{it-1}-\rho y_{i,t-2}\right) \\
&=\mathbb{E}\left[\left(c_i+v_{it}-\delta_i v_{i,t-1}\right)\left(c_i+v_{i,t-1}-\delta_i v_{i,t-2}\right)\right]-c_i^2 \\
&=\mathbb{E}\left[\left((1-\rho)\mu_i+z_{it}-\rho z_{i,t-1}+\gamma_i\eta_t\right)\left((1-\rho)\mu_i+z_{i,t-1}-\rho z_{i,t-2}+\gamma_i\eta_{t-1}\right)\right] \\
&\quad-(1-\rho)^2\mu_i^2 .
\end{aligned}$$

Thus,

$$c_i^2-\delta_i\mathbb{E}\left[v_{i,t-1}^2\right]-c_i^2=(1-\rho)^2\mu_i^2-\rho\mathbb{E}\left[z_{i,t-1}^2\right]-(1-\rho)^2\mu_i^2$$

and we obtain

$$\rho/\delta_i=\sigma_{vi}^2/\sigma_{zi}^2. \tag{24}$$

Finally, plugging equation (24) into equation (23), it is straightforward to deduce the result. □

Proof of Remark 2. *For each i, the MSFE is given by*

$$\begin{aligned}
\mathcal{MSFE}_i&=\mathbb{E}(y_{i,t+h}-\hat{y}_{i,t+h})^2 \\
&=\operatorname{Var}\left(y_{i,t+h}-\hat{y}_{i,t+h}\right)+\left[\mathbb{E}\left(y_{i,t+h}-\hat{y}_{i,t+h}\right)\right]^2.
\end{aligned} \tag{25}$$

Notice that at $t+h$, the true value of the forecast given in equation (14) can be written as

$$y_{i,t+h}=\sum_{j=0}^{h-1}\rho^j\alpha_i+\rho^h y_{it}+\rho^{h-1}\phi z_{it}+\varepsilon_{i,t+h}$$

where $\varepsilon_{i,t+h}=\sum_{j=0}^{h-1}\rho^j\gamma_i\eta_{t+h-j}+z_{i,t+h}$. The forecast error is therefore

$$y_{i,t+h}-\hat{y}_{i,t+h}=\sum_{j=0}^{h-1}\rho^j\gamma_i\eta_{t+h-j}+z_{i,t+h}.$$

Thus,

$$\mathcal{MSFE}_i^{(1)}=\sigma_{zi}^2+\sum_{j=0}^{h-1}\rho^{2j}\gamma_i^2\sigma_\eta^2.$$

Equation (13) yields

$$\gamma_i^2\sigma_\eta^2=\frac{\left(\rho-\delta_i\right)\left(1-\rho\delta_i\right)}{\delta_i}\sigma_{zi}^2. \tag{26}$$

Thus, using this last equation we obtain the expression of $\mathcal{MSFE}_i^{(1)}$ given in equation (15)

$$\mathcal{MSFE}_i^{(1)}=\sigma_{zi}^2+\sum_{j=0}^{h-1}\rho^{2j}\frac{\left(\rho-\delta_i\right)\left(1-\rho\delta_i\right)}{\delta_i}\sigma_{zi}^2. \quad \square$$

Proof of Remark 3. *For the univariate forecast (16), the true value at $t+h$ is*

$$y_{i,t+h} = \sum_{j=0}^{h-1} \rho^h c_i + \rho^h y_{i,t} - \rho^{h-1}\delta_i v_{it} + \sum_{j=1}^{h-1} \rho^{j-1}\left(\rho v_{i,t+h-j} - \delta_i v_{i,t+h-j}\right) + v_{i,t+h}. \qquad (27)$$

Thus, we have

$$y_{i,t+h} - \hat{y}_{i,t+h} = \sum_{j=1}^{h-1} \rho^{j-1}\left(\rho v_{i,t+h-j} - \delta_i v_{i,t+h-j}\right) + v_{i,t+h}$$

and the corresponding MSFE is

$$\mathcal{MSFE}_i^{(2)} = \sigma_{vi}^2 + \sum_{j=1}^{h-1} \rho^{2(j-1)}(\rho - \delta_i)^2 \sigma_{vi}^2.$$

Using equation (24), we obtain

$$\begin{aligned}
\mathcal{MSFE}_i^{(2)} &= \frac{\rho}{\delta_i}\sigma_{zi}^2 + \sum_{j=1}^{h-1} \rho^{2(j-1)}(\rho - \delta_i)^2 \frac{\rho}{\delta_i}\sigma_{zi}^2 \\
&= \frac{\rho}{\delta_i}\sigma_{zi}^2 + \sum_{j=1}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2 \\
&= \frac{\rho}{\delta_i}\sigma_{zi}^2 - \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2 \\
&= \frac{\rho^2}{\rho\delta_i}\sigma_{zi}^2 - \frac{\rho^2 - 2\rho\delta_i + \delta_i^2}{\rho\delta_i}\sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2 \\
&= \frac{(2\rho - \delta_i)}{\rho}\sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2
\end{aligned}$$

which corresponds to the result. □

Proof of Proposition 1. *The result of this proposition is a consequence of Remarks 2 and 3. To beggin the proof, consider the following expression of $\mathcal{MSFE}_i^{(2)}$,*

$$\mathcal{MSFE}_i^{(2)} = \sigma_{zi}^2 + \frac{(\rho - \delta_i)}{\rho}\sigma_{zi}^2 + \sum_{j=0}^{h-1} \rho^{2j} \frac{(\rho - \delta_i)^2}{\rho\delta_i}\sigma_{zi}^2.$$

Then we have,

$$\begin{aligned}\Delta_i &= \mathcal{MSFE}_i^{(2)} - \mathcal{MSFE}_i^{(1)} \\ &= \frac{(\rho-\delta_i)}{\rho}\sigma_{zi}^2 + \sum_{j=0}^{h-1}\rho^{2j}\left(\frac{(\rho-\delta_i)^2}{\rho\delta_i} - \frac{(\rho-\delta_i)(1-\rho\delta_i)}{\delta_i}\right)\sigma_{zi}^2 \\ &= \frac{(\rho-\delta_i)}{\rho}\sigma_{zi}^2 + \sum_{j=0}^{h-1}\rho^{2j}\left(\frac{(\rho-\delta_i)(\rho^2-1)}{\rho}\right)\sigma_{zi}^2.\end{aligned}$$

For the nonstationary case ($\rho = 1$) *the result is immediate,*

$$\Delta_i = (1-\delta_i)\,\sigma_{zi}^2 > 0.$$

For the stationary case ($|\rho| < 1$) *we have*

$$\begin{aligned}\Delta_i &= \frac{(\rho-\delta_i)}{\rho}\sigma_{zi}^2 + \left(\sigma_{zi}^2\frac{(\rho-\delta_i)(\rho^2-1)}{\rho}\right)\left(\frac{1-\rho^{2h}}{1-\rho^2}\right) \\ &= \frac{(\rho-\delta_i)}{\rho}\sigma_{zi}^2 + \frac{(\rho-\delta_i)(\rho^2-1)}{\rho(1-\rho^2)}\sigma_{zi}^2 \\ &- \left(\sigma_{zi}^2\frac{(\rho-\delta_i)(\rho^2-1)}{\rho}\right)\frac{\rho^{2h}}{1-\rho^2} \\ &= \frac{(\rho-\delta_i)(1-\rho^2)}{\rho(1-\rho^2)}\sigma_{zi}^2 - \frac{(\rho-\delta_i)(1-\rho^2)}{\rho(1-\rho^2)}\sigma_{zi}^2 \\ &+ \left(\sigma_{zi}^2\frac{(\rho-\delta_i)(\rho^2-1)}{\rho}\right)\frac{\rho^{2h}}{\rho^2-1} \\ &= \rho^{2h-1}(\rho-\delta_i)\,\sigma_{zi}^2.\end{aligned}$$

We saw from relation (13) that for each i, δ_i *have the same sign that the pooled autoregressive parameter* ρ *and* $|\rho| > |\delta_i|$. *It follows that* $\rho^{2h-1}(\rho-\delta_i)\,\sigma_{zi}^2$ *will be always positive. Thus, we finally have*

$$\Delta_i = \rho^{2h-1}(\rho-\delta_i)\,\sigma_{zi}^2 > 0. \ \square$$

Appendix B

Derivation of equation (14). *Equation (14) can be obtained by applying recursively* h *times the true value of* $\hat{y}$ *given in equation (9) . We have*

$$y_{i,t+1} = \rho y_{it} + z_{i,t+1} - \rho z_{it} + \gamma_i\eta_{t+1} + \alpha_i$$

$$\begin{aligned}y_{i,t+2} &= \rho y_{i,t+1} + \underbrace{z_{i,t+2} - \rho z_{i,t+1} + \gamma_i\eta_{t+2} + \alpha_i}_{=(I)} \\ &= \rho(\rho y_{it} + z_{i,t+1} - \rho z_{it} + \gamma_i\eta_{t+1} + \alpha_i) + (I) \\ &= \rho^2 y_{it} + \rho z_{i,t+1} - \rho^2 z_{it} + \rho\gamma_i\eta_{t+1} + \rho\alpha_i + (I)\end{aligned}$$

$$
\begin{aligned}
y_{i,t+3} &= \rho y_{i,t+2} + \underbrace{z_{i,t+3} - \rho z_{i,t+2} + \gamma_i \eta_{t+3} + \alpha_i}_{=(II)} \\
&= \rho\left(\rho^2 y_{it} + \rho z_{i,t+1} - \rho^2 z_{it} + \rho\gamma_i\eta_{t+1} + \rho\alpha_i + (I)\right) + (II) \\
&= \rho^3 y_{it} + \rho^2 z_{i,t+1} - \rho^3 z_{it} + \rho^2\gamma_i\eta_{t+1} + \rho^2\alpha_i + \rho\,(I) + (II)
\end{aligned}
$$

$$
\begin{aligned}
y_{i,t+4} &= \rho y_{i,t+3} + \underbrace{z_{i,t+4} - \rho z_{i,t+3} + \gamma_i \eta_{t+4} + \alpha_i}_{=(III)} \\
&= \rho\left(\rho^3 y_{it} + \rho^2 z_{i,t+1} - \rho^3 z_{it} + \rho^2\gamma_i\eta_{t+1} + \rho^2\alpha_i + \rho\,(I) + (II)\right) + (III) \\
&= \rho^4 y_{it} + \rho^3 z_{i,t+1} - \rho^4 z_{it} + \rho^3\gamma_i\eta_{t+1} + \rho^3\alpha_i + \rho^2\,(I) + \rho\,(II) + (III)
\end{aligned}
$$

Thus, for horizon 4 we have

$$
\begin{aligned}
y_{i,t+4} &= \rho^4 y_{it} + \left(\rho^3 z_{i,t+1} - \rho^4 z_{it} + \rho^3\gamma_i\eta_{t+1} + \rho^3\alpha_i\right) + \left(\rho^2 z_{i,t+2} - \rho^3 z_{i,t+1} + \rho^2\gamma_i\eta_{t+2} + \rho^2\alpha_i\right) \\
&\quad + \left(\rho z_{i,t+3} - \rho^2 z_{i,t+2} + \rho\gamma_i\eta_{t+3} + \rho\alpha_i\right) + \left(z_{i,t+4} - \rho z_{i,t+3} + \gamma_i\eta_{t+4} + \alpha_i\right) \\
&= \sum_{j=0}^{4-1}\rho^j\alpha_i + \rho^4 y_{it} + \sum_{j=0}^{4-1}\rho^j\gamma_i\eta_{t+4-j} + \left(\rho^3 z_{i,t+1} - \rho^4 z_{it}\right) \\
&\quad + \left(\rho^2 z_{i,t+2} - \rho^3 z_{i,t+1}\right) + \left(\rho z_{i,t+3} - \rho^2 z_{i,t+2}\right) + \left(z_{i,t+4} - \rho z_{i,t+3}\right) \\
&= \sum_{j=0}^{4-1}\rho^j\alpha_i + \rho^4 y_{it} - \rho^4 z_{it} + \sum_{j=0}^{4-1}\rho^j\gamma_i\eta_{t+4-j} + z_{i,t+4}
\end{aligned}
$$

So, using the relation $\phi = -\rho$, we have at horizon h

$$
y_{i,t+h} = \sum_{j=0}^{h-1}\rho^j\alpha_i + \rho^h y_{it} + \rho^{h-1}\phi z_{it} + \sum_{j=0}^{h-1}\rho^j\gamma_i\eta_{t+h-j} + z_{i,t+h}
$$

We can then deduce the h-steps ahead prediction (Equation (14))

$$
\hat{y}_{i,t+h} = \sum_{j=0}^{h-1}\rho^j\alpha_i + \rho^h y_{it} + \rho^{h-1}\phi z_{it}
$$

Derivation of equation (16). *Proceeding in the same way as above, and based this time on the true value of the forecast given in equation (11), we get*

$$
y_{i,t+1} = \rho y_{it} - \delta_i v_{it} + v_{i,t+1} + c_i
$$

$$
\begin{aligned}
y_{i,t+2} &= \rho y_{i,t+1} \underbrace{-\delta_i v_{i,t+1} + v_{i,t+2} + c_i}_{=A} \\
&= \rho\left(\rho y_{it} - \delta_i v_{it} + v_{i,t+1} + c_i\right) + A \\
&= \rho^2 y_{it} - \rho\delta_i v_{it} + \rho v_{i,t+1} + \rho c_i + A
\end{aligned}
$$

$$
\begin{aligned}
y_{i,t+3} &= \rho y_{i,t+2} \underbrace{-\delta_i v_{i,t+2} + v_{i,t+3} + c_i}_{=B} \\
&= \rho\left(\rho^2 y_{it} - \rho\delta_i v_{it} + \rho v_{i,t+1} + \rho c_i + A\right) + B \\
&= \rho^3 y_{it} - \rho^2\delta_i v_{it} + \rho^2 v_{i,t+1} + \rho^2 c_i + \rho A + B
\end{aligned}
$$

$$
\begin{aligned}
y_{i,t+4} &= \rho y_{i,t+3} \underbrace{-\delta_i v_{i,t+3} + v_{i,t+4} + c_i}_{=C} \\
&= \rho\left(\rho^3 y_{it} - \rho^2\delta_i v_{it} + \rho^2 v_{i,t+1} + \rho^2 c_i + \rho A + B\right) + C \\
&= \rho^4 y_{it} - \rho^3\delta_i v_{it} + \rho^3 v_{i,t+1} + \rho^3 c_i + \rho^2 A + \rho B + C.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
y_{i,t+4} &= \rho^4 y_{it} - \rho^3\delta_i v_{it} + \rho^3 v_{i,t+1} + \rho^3 c_i + \rho^2\left(-\delta_i v_{i,t+1} + v_{i,t+2} + c_i\right) \\
&+ \rho\left(-\delta_i v_{i,t+2} + v_{i,t+3} + c_i\right) + \left(-\delta_i v_{i,t+3} + v_{i,t+4} + c_i\right) \\
&= \rho^4 y_{it} + \left(-\rho^3\delta_i v_{it} + \rho^3 v_{i,t+1} + \rho^3 c_i\right) + \left(-\rho^2\delta_i v_{i,t+1} + \rho^2 v_{i,t+2} + \rho^2 c_i\right) \\
&+ \left(-\rho\delta_i v_{i,t+2} + \rho v_{i,t+3} + \rho c_i\right) + \left(-\delta_i v_{i,t+3} + v_{i,t+4} + c_i\right) \\
&= \left(\rho^3 c_i + \rho^2 c_i + \rho c_i + c_i\right) + \rho^4 y_{it} - \rho^3\delta_i v_{it} + \left(\rho^3 v_{i,t+1} - \rho^2\delta_i v_{i,t+1}\right) \\
&+ \left(\rho^2 v_{i,t+2} - \rho\delta_i v_{i,t+2}\right) + \left(\rho v_{i,t+3} - \delta_i v_{i,t+3}\right) + v_{i,t+4} \\
&= \sum_{j=0}^{4-1}\rho^j c_i + \rho^4 y_{it} - \rho^{4-1}\delta_i v_{it} + \sum_{j=1}^{4-1}\rho^{j-1}\left(\rho v_{i,t+4-j} - \delta_i v_{i,t+4-j}\right) + v_{i,t+4}.
\end{aligned}
$$

Finally, at horizon h *the true value of* y_{it} *is*

$$
y_{i,t+h} = \sum_{j=0}^{h-1}\rho^j c_i + \rho^h y_{it} - \rho^{h-1}\delta_i v_{it} + \sum_{j=1}^{h-1}\rho^{j-1}\left(\rho v_{i,t+h-j} - \delta_i v_{i,t+h-j}\right) + v_{i,t+h}.
$$

Its predicted value is then given by

$$
\hat{y}_{i,t+h} = \sum_{j=0}^{h-1}\rho^j c_i + \rho^h y_{it} - \rho^{h-1}\delta_i v_{it}
$$

which corresponds to equation (16).