Forecasting Time Series with a Mixture of Stationary and Non-stationary Factors [1]

Pre-submission Report

Sium Bodha Hannadige

Supervisors:

Professor Jiti Gao

Professor Mervyn Silvapulle

Department of Econometrics and Business Statistics

Monash University

March 2021

[1]This is the thesis title

1 Background and Motivation

Accurate forecasting of key economic variables, such as real economic activity and inflation, is central to making economic policy decisions. Development of methodology for choosing a good forecasting model, estimating its parameters, and constructing reliable forecasts based on the estimated model has been an active area of research. The main objective of this thesis is to develop new improved methodology for forecasting macroeconomic variables, such as GDP and Inflation.

In a method that has attracted considerable attention, an approximate factor model for panel data and a regression model for predicting the macroeconomic variable are used jointly involving two-steps. In the first step, the approximate factor model is used for generating a small number of factors to extract a large proportion of the information contained in a set of panel data on a large number of macroeconomic variables that are also considered to be potential predictors. In the second step, a separate regression model is used for predicting the desired macroeconomic variable, for example GDP or Inflation. In this regression model, the aforementioned small number of generated factors are used as predictors instead of the unmanageably large number of potential predictors in the set of panel data; the regression model may also contain other observed predictors. The resulting model is called *factor augmented regression* model (Bernanke et al. 2005, Stock and Watson 1998a, Stock and Watson 1998a, Stock and Watson 1998b, Stock and Watson 2002). The large number of macroeconomic variables appearing in the set of panel data, typically includes a mixture of stationary and nonstationary variables. Consequently, the collection of factors generated to extract most of the information therein is also typically a mixture of stationary and nonstationary ones (Bai 2004, Eickmeier 2005, Moon and Perron 2007). The objective of this thesis is to develop improved methodology for forecasting a macroeconomic variable when the set of predictors in the forecasting model includes a mixture of stationary and nonstationary factors.

The validity of the aforementioned method based on the *factor-augmented regression* [FAR] model for forecasting has been established when all the variables are stationary (Bai and Ng 2006), and also when they are all nonstationary (Choi 2017). This thesis builds on the aforementioned current literature and develops a method for constructing point forecasts and forecast intervals when the chosen set of factors is a mixture of stationary and nonstationary ones.

There are two separate scenarios that arise in this context. First, let us suppose that the variable that we wish to forecast is $I(0)$. If the panel data set contains a mixture of I(0) and I(1) variables, as is usually the case, then it has been suggested to replace the I(1) variables by their first order differences, which would be $I(0)$, and then apply a method such as that in Bai and Ng [2006] that is applicable when all the variables are stationary. While this method is methodologically valid, it is likely to be inefficient because differencing an I(1) variable is likely to result in loss of information in the level-data that may be important for forecasting. Therefore, there is a need to develop a method for forecasting an $I(0)$ variable when the panel data is a mixture of stationary and nonstationary variables and the method does not difference the nonstationary variables to suit the stationarity assumptions of the method.

To consider the second scenario, let us suppose that the variable that we wish to forecast is nonstationary; throughout this thesis, we use the terms nonstationary and $I(1)$ synonymously. If the panel data set contains a mixture of I(0) and I(1) variables, it has been suggested to delete all the potential predictors that are $I(0)$ and use only the $I(1)$ variables. Deletion of such potential predictors is likely to result in loss of information and statistical efficiency. Therefore, there is a need to develop a method for forecasting an $I(1)$ variable without deleting the $I(0)$ variables in the panel data.

The foregoing areas, discussed two separate scenarios, are underdeveloped and hence there is a need for methodological developments. The main objective of this thesis is to advance the econometric literature in these directions by developing new methodology to improve over the current literature. The thesis contains three main chapters, namely Chapters 3, 4, and 5. Chapter 3 is practically ready for submission to a journal for publication; a part of this year was devoted to tightening the proofs in the Appendix to this chapter. An earlier version of this paper, excluding the Appendix, has also appeared as a Working Paper in the working paper series of the Department of Econometrics and Business Statistics, Monash University; a copy of the revised version of this paper is also attached. Significant progress has been made on Chapters 4 and 5 as outlined later in this report.

The novelty of Chapter 3 is that it allows the predictors in the FAR model to be a mixture of stationary and nonstationary variables while the rest of the FAR model structure in the current literature is unchanged. By maintaining a close connection to the extensive literature on FAR models, we were able to make use of the arguments used in the mathematical proofs of the asymptotic results.

In Chapters 4 and 5, we relax some of the assumptions that underlie Chapter 3; these changes to the assumptions are motivated by our prior knowledge and conjectures about the macroeconomic variables involved. In Chapter 4, we extend Chapter 3 by allowing the parameters to be time varying. Such time varying parameters may arise, for example, if the structure of the economy changes during the time period of the data. We estimate time varying parameters by using a kernel method and two systems of orthogonal polynomials. Further, we compare and contrast the performance of this method with its competitors, in particular those in the previous chapter where the parameters are assumed to be not time varying.

An assumption that underlies the FAR methods in Chapters 3 and 4 is that each factor may affect all the variables in the panel. Consequently, the factor model contains a large number of unknown factor loadings that must be estimated. In some cases, it is possible to use our prior knowledge about the relationships among the economic variables to reduce the number of such unknown parameters and thus improve the efficiency of the statistical method; a related point was discussed in Boivin and Ng [2006] and Beck et al. [2009]. Such a scenario may arise in the following context. The macroeconomy could be partitioned into several sectors, and each of a number of factors is specific to only one sector and hence has no effects on the variables in the other sectors, which in turn results in some of the factor loadings being structurally zero. Such a reduction in the number of unknown factor loadings could be expected to improve statistical efficiency. We study this scenario in Chapter 5, and illustrate the benefits of making use of the aforementioned type of prior knowledge about the economy to improve efficiency of the method. We are currently working on Chapters 4 and 5, which are closely related.

In summary, the Chapter 3 improves over the current literature. Chapter 4 improves over Chapter 3, and Chapter 5 improves over Chapters 3 and 4.

2 Overview of the thesis

The thesis contains six chapters. The contents of each chapter are described below.

Chapter 1: Introduction

This introductory chapter for the entire thesis provides a broad overview of the topic and discussions about the contents of each chapter to follow. The first part of the chapter provides a gentle introduction to the factor augmented regression model [FAR]. A specific empirical example is studied to introduce the topic and illustrate the use of FAR model for forecasting the macroeconomic variables, GDP and Inflation. This method uses a set of panel data on a large number of macroeconomic variables. The data set that we use is the well-known set from FREQ-QD; it contains a set of panel data on a large number of macroeconomic variables. After deleting variables with missing observations, we obtain a set of quarterly panel data for $N = 100$ macroeconomic variables over $T = 240$ quarters. The panel data contain a mixture of stationary and nonstationary macroeconomic variables. In this introductory chapter, we apply the techniques in the current literature as a gentle introduction and illustrative purposes.

In the first part of the chapter, we illustrate the FAR method for forecasting the $I(0)$ and $I(1)$ variables. First, we study forecasting the $I(0)$ variables, GDP growth rate and inflation, using the stationary FAR method in Bai and Ng [2006]. To apply this method, we take the first difference of all the $I(1)$ variables. Similarly, we also illustrate the FAR method in Choi [2017] for forecasting an $I(1)$ variable; in this example, we illustrate it for forecasting GPD and Industrial Production. In this part of the illustrative example for forecasting an $I(1)$ variable, we delete all the $I(0)$ variables and use only the nonstationary ones. We discuss the possibility of loss of important information and the resulting loss of efficiency in forecasting; for forecasting an $I(0)$ variable, this relates to taking first difference of the $I(1)$ variables, and for forecasting an $I(1)$ variable it relates to deletion of $I(0)$ variables. While this example uses only the existing methodology, the discussions are extended to provide a motivation for the new methods developed in the thesis. In doing so, we highlight the complexities when there is a mixture of stationary and nonstationary variables, and the potential gains in efficiency if the mixture of stationary and nonstationary variables are used.

The second part of this chapter introduces an extension of the mixture-FAR method to accommodate time-varying coefficients in the prediction model. Since the panel data set for macroeconomic variables spans a long time period, it is possible that there might have been significant structural changes in the economy and the corresponding changes to parameters in the model. Such changes may well be a consequence of policy changes, institutional switching, and technological progress, among others. We discuss the possible inconsistency of the estimators and unreliability of the forecasts due to such changes if the current FAR methods are used. Using the same empirical example, motivation to the semi-parametric FAR model and the potential gains compared to its parametric counterpart, are also discussed.

The third part of this chapter discusses multi-level factor structure in the factor model, and the semi-parametric mixture FAR model with multi-level factor structure in the FAR model. We discuss potential gains in the forecasting performance of the new method compared to the one-level FAR models.

Chapter 2: Literature review

This chapter provides a review of the relevant literature on FAR models for forecasting. Using the literature review as the foundation, we discuss the importance of (a) developing methodology for using a mixture of nonstationary and stationary factors in the FAR model for forecasting nonstationary time series, (b) allowing the parameters corresponding to the factors in the regression model to be time-varying, and (c) introducing the multi-level factor structure to the family of factor augmented regression models.

Chapter 3: Forecasting time series with a mixture of nonstationary and stationary factors as predictors

Let $X_{[T\times N]} = \{X_{it} : t = 1, \ldots, T; i = 1, \ldots, N\}$ denote a set of observable panel data, and let $Y_{[T\times 1]} = \{Y_t : t = 1, \ldots, T\}$ denote a univariate series that we wish to predict at time $T + h$ for some $h > 0$. Then, the main structure of the model is the following:

$$\text{Factor model: } X_{it} = \lambda_i' F_t + e_{it} \qquad (i = 1, ..., N; t = 1, ..., T), \tag{1}$$

$$\text{FAR model: } Y_{t+h} = \theta' F_t + \omega' W_t + \epsilon_{t+h} \quad (t = 1, ..., T), \tag{2}$$

where F_t is an $r \times 1$ vector of unobservable factors, $\{e_{it}, \epsilon_t\}$ are idiosyncratic errors, λ_i is an $r \times 1$ vector of factor loadings, $\{\theta, \omega\}$ are unknown parameters, and W_t is a vector of observable variables that may also include time lags of the dependent variable.

Since the factors in the regression model are unobservable, we follow a two-step method. First, the latent factors $\{F_t : t = 1, \ldots, T\}$ are estimated by using principal components analysis [PCA]. Let $\tilde{F}_t$ denote the estimated factor at time t. Then, the parameter (θ, ω) is estimated using the model,

$$Y_{t+h} = \theta' \tilde{F}_t + \omega' W_t + error, \quad (t = 1, ..., T). \tag{3}$$

The difference between (2) and (3) is that the latter has $\tilde{F}_t$ for the F_t in the former; consequently, the error terms are also different. Although (2) and (3) are different, we do not distinguish between them unless it is necessary to do so; only the latter is used in the numerical computations. In what follows, we write (θ, ω) for $(\theta', \omega')'$. Let $(\hat{\theta}, \hat{\omega})$ be the OLS estimator of (θ, ω). Then, the forecast for Y_{T+h} is

$$\hat{Y}_{T+h} = \hat{\theta}' \tilde{F}_T + \hat{\omega}' W_T.$$

The novelty of this chapter is that it allows the set of factors $\{F_t \in \mathbb{R}^r : t = 1, \ldots, T\}$ to be a mixture of stationary and nonstationary variables.. Therefore, we refer to (2) and (3) as *mixture-FAR* model. The main methodological contributions of this chapter are the following: (a) establishes consistency of the estimated factors $\tilde{F}_t$, up to a rotation of the latent factors F_t; (b) establishes the consistency and some asymptotic properties of the estimators $(\hat{\theta}, \hat{\omega})$; and

(c) constructs an asymptotically valid prediction interval for Y_{T+h} and a confidence interval for its expected value, denoted $Y_{T+h|T}$.

In the second part of the chapter, we conduct a simulation study to evaluate the finite sample properties of the mixture-FAR model. Furthermore, we compare and contrast the forecasting performance of the method with its competitors in the recent literature. We observed that the overall performance of the method developed in this paper was better than that of its competitors for forecasting a nonstationary variable when there is a mixture of stationary and nonstationary factors. The details are provided in the attached draft copy of the paper.

In the last part of this chapter, we consider an empirical example. For this example, we use macroeconomic quarterly data from the USA for the period 1959:Q1 - 2018:Q4. We considered forecasting the two I(1) variables, *GDP* and *industrial production* [IP], and the two I(0) variables, *GDP growth rate* and *inflation*. In the empirical example, we observed that the mixture-FAR model performed better than the nonstationary FAR and AR(4) for forecasting the two nonstationary variables. However, for forecasting inflation, AR(4) and the stationary FAR performed better than the mixture-FAR. This result was not unexpected since the stationary-FAR is likely to perform better than the nonstationary-FAR for forecasting $I(0)$ variables. Forecast evaluation for inflation is repeated with an additional predictor, *economic policy uncertainty index* [EPUI], and observed that the mixture-FAR was better than AR(4) in terms of forecast performance. Overall, for forecasting a nonstationary variable, there is strong evidence that the method developed in this chapter is likely to perform better than its competitors in the literature. In summary, the mixture-FAR method developed in this chapter makes a significant contribution to advance the econometric literature.

Chapter 4: Forecasting a time series using a semi-parametric mixture FAR model

This chapter extends the methods developed in Chapter 3 to the case when the unknown parameters may be time varying. We refer to this as *semi-parametric mixture FAR* method; the term 'semi-parametric' relates to the coefficients being nonconstant and changes nonparametrically over time. The first part of the chapter contains the methodological developments on the semi-parametric mixture-FAR model. The next two parts of the chapter contain a simulation study and an empirical application to illustrate the new method. In the numerical studies, we compare and contrast the forecasting performance of the proposed semi-parametric mixture-FAR model with its parametric counterpart in the previous chapter.

Note: Research for this chapter is in progress. We are in the process of comparing three potential methods using numerical studies. Once these provide sufficiently strong indication of the relative performance of the methods, we will choose one method and develop the theory for the chosen method.

Methodology

Let $X_{T\times N}$ and $Y_{T\times 1}$ be as in the previous chapter. The objective is also the same as in the previous chapter, namely to predict Y_{T+h}. We start with the model specifications (1)-(3). Let $F_t = (E_t', G_t')'$, where E_t and G_t are, respectively, the stationary and the nonstationary components of F_t. The semi-parametric FAR model studied in this chapter is

$$Y_{t+h} = \alpha_t' E_t + \beta_t' G_t + \omega' W_t + \xi_{t+h} = \theta_t' F_t + \omega' W_t + \epsilon_{t+h}, \qquad (t = 1, \ldots, T; h > 0) \tag{4}$$

where W_t is an $m \times 1$ vector of observable regressors, ω is an $m \times 1$ vector of constant parameters, and $\theta_t = (\alpha_t', \beta_t')'$ is an $r \times 1$ vector of time-varying parameters of the form

$$\alpha_t = \alpha(\tau_t), \quad \beta_t = \beta(\tau_t), \qquad (\tau_t = t/T; t = 1, \ldots, T) \tag{5}$$

with $\alpha(.)$ and $\beta(.)$ being unknown smoothing functions. We refer to the model defined by (4) and (5) as *semi-parametric mixture-FAR model.* The main difference between the mixture-FAR model in Chapter 3 and the aforementioned semi-parametric mixture-FAR model of this chapter is that the latter assumes that the coefficients α and β are time varying. The methods of estimation adopted in this chapter are suitable modifications of those in Chapter 3.

Let $\tilde{F}_t$ denote the estimated factors as in the previous chapter. Therefore, we may rewrite (4) as,

$$Y_{t+h} = \theta_t' \tilde{F}_t + \omega' W_t + \epsilon_{t+h}, \qquad (t = 1, \ldots T). \tag{6}$$

Estimation procedure

We estimate the semi-parametric FAR model using the following three estimation methods:

a) non-parametric kernel smoothing estimation,

b) Hermite polynomial estimation, and

c) trigonometric polynomial estimation.

a) Nonparametric kernel estimation

For estimating the semi-parametric mixture-FAR model using a kernel, we adopt the methods in Gao and Hawthorne [2006], Li et al. [2011], and Chen et al. [2012]. To this end, first we introduce a weight function $V_{Tk}(t)$ of the form,

$$V_{Tk}(t) = \frac{K\left(\dfrac{\tau_t - \tau_k}{h_w}\right)}{\sum_{u=1}^{T} K\left(\dfrac{\tau_t - \tau_u}{h_w}\right)}, \tag{7}$$

where $K(.)$ is the kernel smoothing function, h_w is the bandwidth, and $\tau_t = t/T$ $(t = 1, \ldots, T)$. Throughout this paper we use the Gaussian kernel function,

$$K(u) = \frac{1}{\sqrt{2\pi}} exp^{-u^2/2}. \tag{8}$$

The bandwidth h_w in (7) is assumed to satisfy $h_w \to 0$ and $Th_w \to \infty$ as $T \to \infty$. Based on the extensive literature on estimation using kernel functions, we would expect that the performance of our method also to be sensitive to the choice of the bandwidth but not to the choice of the kernel function. Motivated by the foregoing ideas, we adopted the following three-step method for estimating the model.

Step 1: *Nonparametric estimation*

For a given ω, we estimate the time-varying coefficients $\theta(\tau_t)$ by minimizing the loss function,

$$\sum_{k=1}^{T-h} \left(Y_{k+h} - \omega' W_k - \theta(\tau_k)' \tilde{F}_k \right)^2 K\left(\frac{\tau_t - \tau_k}{h_w} \right). \tag{9}$$

Therefore, the estimated $\theta(\tau_t)$ is given by,

$$\begin{aligned} \tilde{\theta}(\tau_t) = & \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) Y_{k+h} \\ & - \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \omega' W_k, \end{aligned}$$

and

$$\tilde{\theta}(\tau_t)' \tilde{F}_t = \sum_{k=1}^{T-h} V_{Tt}(k) \left(Y_{k+h} - \omega W_k \right) = \sum_{k=1}^{T-h} V_{Tt}(k) Y_{k+h} - \sum_{k=1}^{T-h} V_{Tt}(k) \omega' W_k.$$

Step 2: *Parametric estimation*

Replace θ_t in (6) by the estimated coefficient $\tilde{\theta}(\tau_t)$ and obtain,

$$\begin{aligned} Y_{t+h} &= \tilde{\theta}(\tau_t)' \tilde{F}_t + \omega' W_t + \epsilon_{t+h} \\ &= \tilde{F}_t' \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) Y_{k+h} \\ &\quad - \tilde{F}_t' \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \omega' W_k + \omega' W_t + \epsilon_{t+h}. \end{aligned}$$

Let

$$\tilde{Y}_{t+h} = Y_{t+h} - \tilde{F}_t' \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) Y_{k+h}, \tag{10}$$

$$\tilde{W}_t = W_t - \tilde{F}_t' \left(\sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K\left(\frac{\tau_k - \tau_t}{h_w} \right) W_k. \tag{11}$$

Then, $\tilde{Y}_{t+h} = \omega' \tilde{W}_t + \epsilon_{t+h}$, and the corresponding least-square estimate of ω is

$$\hat{\omega} = \left(\sum_{t=1}^{T-h} \tilde{W}_t' \tilde{W}_t \right)^{-1} \left(\sum_{t=1}^{T-h} \tilde{W}_t' \tilde{Y}_{t+h} \right)^{-1}. \tag{12}$$

Step 3: *Semi-parametric estimation*

Substitute $\hat{\omega}$ for ω in the expression for the infeasible estimate $\tilde{\theta}(\tau_t)$, and obtain the feasible estimate $\hat{\theta}(\tau_t)$ as,

$$\hat{\theta}(\tau_t) = \left(\sum_{k=1}^{T-h} \tilde{F}_k K \left(\frac{\tau_k - \tau_t}{h_w} \right) \tilde{F}_k' \right)^{-1} \sum_{k=1}^{T-h} \tilde{F}_k K \left(\frac{\tau_k - \tau_t}{h_w} \right) (Y_{k+h} - \hat{\omega} W_k) . \tag{13}$$

b) Estimation using Hermite polynomials

Let $L^2[\mathbb{R}, exp(-x^2/2)]$ denote the Hilbert space of continuous functions on $\mathbb{R}$ with the inner product defined by $< g, h >= \int g(x) f(x) \exp(-x^2/2) \; dx$. Let $H_i(\cdot)$ $(i = 0, 1, \ldots)$ denote the sequence of Hermite polynomials defined by

$$H_i(x) = (-1)^i exp(x^2/2) \frac{d^i}{dx^i} exp(-x^2/2), \quad (i = 0, 1, \ldots,).$$

Then, the first few polynomials are $H_0(x) = 1,\; H_1(x) = x,\; H_2(x) = x^2 - 1$, and $H_3(x) = x^3 - 3x$. These polynomials are orthogonal in the sense $\int H_i(x) H_j(x) exp(-x^2/2) dx = \sqrt{\pi} 2^i i! \delta_{ij}$, where δ_{ij} is the Kronecker delta. Let $h_i(x) = \left(\sqrt{2\pi} i! \right)^{-1/2} H_i(x)$. Then, $\{h_i(x), i = 0, 1, \ldots\}$ is an orthonormal basis.

Let $g(x) \in L^2[\mathbb{R}, exp(-x^2/2)]$. Then, we have the orthogonal series expansion

$$\begin{aligned} g(x) &= \sum_{i=0}^{\infty} c_i h_i(x), \qquad c_i = \langle g(x), h_i(x) \rangle \\ &= \sum_{i=0}^{k-1} c_i h_i(x) + \sum_{i=k}^{\infty} c_i h_i(x) \\ &= g_k(x) + \gamma_k(x). \end{aligned}$$

Therefore,

$$g(x) = Z_k(x)' C + \gamma_k(x),$$

where $C_{[k \times 1]} = (c_0, c_1, \ldots, c_{k-1})'$ and $Z_k(x) = (h_0(x), h_1(x), \ldots, h_{k-1}(x))'$.

Now, we use these to estimate the coefficients in the semi-parametric FAR model (6). Let $x = \tau_t = t/T$ and $g(x) = \theta(\tau_t)$. Then, we may rewrite model (6) as,

$$Y_{t+h} = \left(\tilde{F}_t' \otimes Z_k(\tau_t)' \right) vec(C) + \omega W_t + \epsilon_{t+h}, \qquad (t = 1, \ldots, T) \tag{14}$$

where $Z_k(\tau_t) = (h_0(\tau_t), h_1(\tau_t), \ldots, h_{k-1}(\tau_t))'$ and $C = (c_0, c_1, \ldots, c_{k-1})'$.

Let $U_t = \begin{pmatrix} \tilde{F}_t{}' \otimes Z_k(\tau_t)' & W_t \end{pmatrix}$ $(t = 1, \ldots, T)$, and $\delta = \begin{pmatrix} (vecC)' & \omega' \end{pmatrix}'$. Then, we have

$$Y_{t+h} = U_t'\delta + \epsilon_{t+h}, \qquad (t = 1, \ldots, T; h > 0). \tag{15}$$

We estimate δ by the OLS estimate

$$\hat{\delta} = \left(\sum_{t=1}^{T-h} U_t'U_t\right)^{-1} \left(\sum_{t=1}^{T-h} U_t'Y_{t+h}\right).$$

Therefore, we have $\widehat{vecC} = \hat{\delta}(1 : 2k, 1)$ and $\hat{\omega} = \hat{\delta}(2k+1, 1)$. Hence, we obtain the $\hat{\theta}_t$ and $\hat{\omega}$.

c) Estimation using trigonometric polynomials

Similar to Hermite polynomial estimation, in the trigonometric polynomial estimation method we approximate the θ_t by the orthogonal series

$$\sum_{j=1}^{k} Z_j(.)\gamma_k,$$

where $\{Z_j(.) : 1 \leq j \leq k\}$ is a continuous function, $\gamma = (\gamma_1, \ldots, \gamma_k)'$ is a vector of unknown coefficients, and k is the truncation parameter. Let $v = 2\pi\tau_t$,
$Z_k = (1, cos(v), sin(v); cos(2v), sin(2v); \ldots, cos(kv), sin(kv))'$, and $Z = (Z_1, \ldots, Z_k)'$. Then we may rewrite (6) as,

$$Y_{t+h} = \left(\tilde{F}_t' \otimes Z_k(\tau_t)'\right) vec(\gamma) + \omega W_t + \epsilon_{t+h}, \qquad (t = 1, \ldots, T),$$

and obtain $\hat{\delta} = \begin{pmatrix} \widehat{vec(\gamma)}' & \hat{\omega}' \end{pmatrix}'$ by OLS estimation. Hence, we have $\hat{\theta}_t = \sum_{j=1}^{k} Z_j(2\pi\tau_t)\widehat{vec(\gamma)}$.

After estimating the semi-parametric FAR model using the aforementioned methods, we use the model to obtain forecast for Y_{T+h} as,

$$\hat{Y}_{T+h} = \hat{\theta}_t{}'\tilde{F}_T + \hat{\omega}'W_T.$$

Simulation study

The second part of this chapter contains the results of a simulation study on the finite sample performance of the estimators and forecasts. We compare and contrast the three estimations methods for the nonparametric component. The design of the simulation study and the main results obtained so far are given below.

Design of the simulation study

We consider the following DGP:

$$Y_{t+1} = \alpha_t F_{1t} + \beta_t F_{2t} + \omega Y_t + \epsilon_{t+1}, \quad \epsilon_t \sim N(0,1), \qquad (t = 1, \ldots, T-1), \tag{16}$$

$$F_{1t} = F_{1,t-1} + v_t, \qquad (v_t, F_{2t}) \sim MVN(0, C), \qquad C = (1, \rho | \rho, 1), \tag{17}$$

$$X_{it} = \lambda_i^{(1)} F_{1t} + \lambda_i^{(2)} F_{2t} + e_{it}, \quad \lambda_i \sim N(0,1), \quad e_{it} \sim N(0,1), \quad (t = 1, \ldots, T, i = 1, \ldots, N) \tag{18}$$

The following values are considered for $\{T, N, \rho\}$: $T = 100, 300, 500$; $N = 150$; $\rho = 0.0, 0.5, 0.9$.

Within each of the foregoing scenarios, the following two DGP s are considered for the time varying parameters:

DGP1: $\alpha_t = (1/2)(1 + \tau_t), \beta_t = \tau_t + \tau_t^2$, and $\omega = 0.7$,

DGP2: $\alpha_t = (1/3)sin(2\pi\tau_t), \beta_t = cos(2\pi\tau_t)$, and $\omega = 0.7$, where $\tau_t = t/T$.

We consider the following seven different values for the truncation constant k in the polynomial approximations for the time varying parameters: $k = cT^{1/7}$, $c = [1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5]$; see Dong et al (2015) for a discussion on the rationale for these choices. The bandwidth is chosen to be $h = 1/k$ for the kernel estimation method.

The main finding of the simulation results indicate that, in terms of the overall forecasting performance (both in-sample and out-of-sample), the semi-parametric FAR model introduced in this chapter is better than its parametric counterpart. Furthermore, we observed that the model estimated using trigonometric polynomials have better in-sample performance, and the kernel estimation method has better out-of-sample performance. We used the expanding window in the aforementioned study. We will perform the simulations with the rolling window as well, and compare with those for the expanding window.

In the third section of this chapter, we consider an empirical application in which we use the same FRED-QD data set. Similar to the chapter 3, we extract four $I(1)$ and four $I(0)$ factors from the data set using PC, and use time lags of dependent variable as the set of observable regressors in the predicting model. We apply the proposed semi-parametric mixture-FAR model for forecasting both stationary and nonstationary variables, namely log(GDP), inflation, and GDP growth rate. For each of these choices, we consider kernel, Hermite, and trigonometric polynomial methods for the nonparametric component.

The main observations of the empirical application are the following for log(GDP), inflation, and GDP growth rate:

(a) Among the three nonparametric methods for modelling the time varying parameters, the trigonometric estimation method performed better in in-sample forecasting, and the kernel estimation method performed better in out-of-sample forecasting.

(b) The semi-parametric mixture-FAR performed better than the basic AR models and the parametric mixture-FAR models. Here we used nonparametric kernel estimation method as it performed better than the other two estimation methods in out-of-sample forecasting.

Chapter 5: Forecasting time-series using semi-parametric mixture-FAR model with multi-level factor structure

This chapter contains an extension of the semi-parametric FAR model discussed in chapter 4. By allowing the factors to have a multi-level structure, we consider both pervasive (global) and non-pervasive (regional) factors. The main structure of the two-level factor model is

$$x_{s,it} = \gamma'_{s,i} H_t + \lambda'_{s,i} R_{s,t} + e_{s,it} \quad (i = 1, \ldots, N_s;\ s = 1, \ldots, S;\ t = 1, \ldots, T) \tag{19}$$

where H_t is an $r_G \times 1$ vector of global factors, $R_{s,t}$ is an $r_s \times 1$ vector of regional/sector factors, $\gamma_{s,i}$ and $\lambda_{s,i}$ are factor loadings, $e_{s,it}$ is the set of idiosyncratic errors, S and N_s are respectively the number of categories and number of variables in category s such that $N = N_1 + \ldots + N_S$. Then, we have,

$$X_t = \Lambda^m F_t^m + e_t = \Lambda^{(m1)} E_t^m + \Lambda^{(m2)} G_t^m + e_t, \qquad (t = 1, \ldots, T) \tag{20}$$

where $F_t^m = (H'_t, R'_{1,t}, \ldots, R'_{S,t})'$, $\Lambda^m = (\Gamma', \Lambda)$, $\Lambda = diag(\Lambda_1, \ldots, \Lambda_S)$, $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_S)$, and E_t^m and G_t^m are nonstationary and stationary factors. We estimate the factors by sequential least squares. Let $\tilde{F}_t^m (t = 1, \ldots, T)$ denote the estimated multi-level factor at time t. Then we use $\tilde{F}_t^m$ as the predictors in the regression model for forecasting Y_{T+h}, using the model

$$Y_{t+h} = \theta_t \tilde{F}_t^m + \omega W_t + \epsilon_{t+h}, \qquad (t = 1, \ldots, T; h > 0). \tag{21}$$

We refer to the above model as *two-level semi-parametric mixture-FAR* model - 'two-level' because there are global and regional factors. The novelty of this chapter is that it introduces the multi-level factor structure to the semi-parametric FAR that we introduced in the previous chapter. The coefficients in model (21) are estimated using the non-parametric Kernel smoothing method. This chapter does not include any mathematical proofs for any theorems.

In the second part of this chapter, we consider the semi-parametric mixture-FAR with two-level factor structure for forecasting a few key macroeconomic variables. We use the FREQ-QD data set for 1959:Q1 - 2018:Q4. The data set is categorized into 13 groups, such as output and income, price, interest rate, and exchange rate. We considered one-step ahead forecasting of log(GDP), inflation, and GDP growth rate. Forecasting performance of the new semi-parametric FAR model is evaluated for the following models:

(i) AR(1) and AR(4);

(ii) The basic mean model, $Y_{t+1} = \mu + \epsilon_{t+1}$ (only for forecasting $I(0)$ variables).

(iii) Parametric FAR with two-level factors.

(iv) Both parametric and semi-parametric FAR models with one-level factor structure - these are the models discussed in Chapters 3 and 4.

In total, we considered 21 models. Forecast performance is evaluated in terms of out-of-sample R-square, denoted R^2_{os}. The main observations from the empirical application are stated below:

(a) For forecasting log(GDP), overall the one-level semi-parametric mixture-FAR models performed better than each of the other models considered.

(b) For forecasting inflation, two-level semi-parametric FAR models performed better compared to the other models.

(c) For forecasting GDP growth rate, either a two-level or a one-level semi-parametric FAR model performed the best, but none of these two performed the best always.

Overall, multi-level semi-parametric FAR models performed better than the competing ones for forecasting any of the three macroeconomic variables. Thus, incorporation of multi-level factors together with time-varying parameters, resulted in improved forecast performance. Therefore, the new ideas introduced in this chapter make an improvement over the previous chapters, at least in the empirical example studied.

Chapter 6: Conclusion

The final chapter of my thesis will contain all the concluding remarks and possible future extensions.

3 Statement of progress and timetable

3.1 Statement of progress

Chapter 1 will be written after completing the rest of the thesis. Most of material for Chapter 2 is already in the copies of the two papers attached. Chapter 3 is practically complete; we need to re-run numerical studies with a minor change to the code; this could be completed in two weeks. Most of the numerical studies for Chapter 4 have been completed. We are exploring the extent of mathematical content to be included in Chapter 4; this work remains to be completed. Chapter 5 is in progress. This chapter is an empirical study, and we do not plan to include any mathematical proofs. Most of the empirical work for Chapter 5 has been completed. The concluding chapter will be written at the end.

3.2 Timetable

March 2021 - May 2021

- Completing the paper (chapter 5) on multi-level FAR model.
- Deriving the theoretical results with semi-parametric FAR model for chapter 4.

May 2021

- Completing the paper on semi-parametric FAR model with theoretical contributions (chapter 4)

June 2021 - August 2021

- Completing the introduction, literature review, and conclusion chapters and submit the thesis.

References

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

Bai, J. and Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, 74(4):1133–1150.

Beck, G. W., Hubrich, K., and Marcellino, M. (2009). Regional inflation dynamics within and across euro area countries and a comparison with the united states. *Economic Policy*, 24(57):142–184.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics*, 120(1):387–422.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194.

Chen, J., Gao, J., and Li, D. (2012). Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics*, 171(1):71–85.

Choi, I. (2017). Efficient estimation of nonstationary factor models. *Journal of Statistical Planning and Inference*, 183:18–43.

Eickmeier, S. (2005). Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model. Discussion paper 2, Studies of the Economic Research Centre.

Gao, J. and Hawthorne, K. (2006). Semiparametric estimation and testing of the trend of temperature series. *The Econometrics Journal*, 9(2):332–355.

Li, D., Chen, J., and Gao, J. (2011). Non-parametric time-varying coefficient panel data models with fixed effects. *The Econometrics Journal*, 14(3):387–408.

Moon, H. R. and Perron, B. (2007). An empirical analysis of nonstationarity in a panel of interest rates with factors. *Journal of Applied Econometrics*, 22(2):383–400.

Stock, J. H. and Watson, M. W. (1998a). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. Technical report, National Bureau of Economic Research.

Stock, J. H. and Watson, M. W. (1998b). Diffusion indexes. Technical report, National Bureau of Economic Research.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.