Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Markov-Switching Three-Pass Regression Filter

Pierre Guerin, Danilo Leiva-Leon, Massimiliano Marcellino

Working Paper n. 591

This Version: November, 2016

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Markov-Switching Three-Pass Regression Filter*

Pierre Guérin † Danilo Leiva-Leon ‡ Massimiliano Marcellino§

November 30, 2016

Abstract

We introduce a new approach for the estimation of high-dimensional factor models with regime-switching factor loadings by extending the linear three-pass regression filter to settings where parameters can vary according to Markov processes. The new method, denoted as Markov-Switching three-pass regression filter (MS-3PRF), is suitable for datasets with large cross-sectional dimensions since estimation and inference are straightforward, as opposed to existing regime-switching factor models, where computational complexity limits applicability to few variables. In a Monte-Carlo experiment, we study the finite sample properties of the MS-3PRF and find that it performs favorably compared with alternative modelling approaches whenever there is structural instability in factor loadings. As empirical applications, we consider forecasting economic activity and bilateral exchange rates, finding that the MS-3PRF approach is competitive in both cases.

Keywords: Factor model, Markov-switching, Forecasting.

JEL Classification Code: C22, C23, C53.

---

*We would like to thank Dalibor Stevanovic, seminar participants at Queen's University and the 2016 conference of the Canadian Econometric Study Group held at Western University for helpful comments on a previous version of this paper. We also thank Agustín Díaz for excellent research assistance. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada or the Central Bank of Chile.

†Bank of Canada, e-mail: pguerin@bank-banque-canada.ca

‡Central Bank of Chile, e-mail: dleiva@bcentral.cl

§Bocconi University, IGIER and CEPR, e-mail: massimiliano.marcellino@unibocconi.it

1 Introduction

This paper introduces a new approach for the estimation of high-dimensional factor models with regime-switching factor loadings. The literature on factor models has mostly concentrated on situations where the comovements among variables is assumed to be constant over time. However, there is now a large body of literature that has challenged the assumption of constant parameters to model the macroeconomic environment (see, e.g., Sims (1993) or Canova (1993)), as well as the relevance of modelling time variation for macroeconomic forecasting (see, e.g., D'Agostino et al. (2013) and Aastveit et al. (2016)). The importance of incorporating time instabilities in large-scale factor models has gained traction in the literature in recent years (see, e.g., Eickmeier et al. (2015) and Mikkelsen et al. (2015)), but the number of works on this front remains relatively small. Moreover, this literature has so far been restricted to the estimation of models with time-varying factor loadings where time-variation is modelled using random-walk or autoregressive behaviours, which typically restrict the dynamics of time-variation to gradual changes in the factor loadings that may not be appropriate to all situations. The literature has also considered the estimation of factor models with temporal instability (structural breaks) in both factor loadings and the number of factors, see, e.g., Cheng et al. (2016). In contrast, in this paper, we consider factor loadings that vary according to regime-switching processes so as to model recurrent abrupt changes in factor loadings that are potentially highly relevant features in macroeconomic and financial variables.

Our modelling approach builds on Kelly and Pruitt (2015), who develop a new estimator for factor models – the three-pass regression filter (3PRF) – that relies on a series of ordinary least squares regressions. As emphasized in Kelly and Pruitt (2015), the key difference between principal component analysis (PCA) and the 3PRF approach is that PCA summarizes the cross-sectional information based on the covariance within the predictors whereas 3PRF condenses cross-sectional information based on the correlation of the predictors with the target variable of the forecasting exercise, thereby extending partial least squares. In this paper, we extend the 3PRF approach by introducing regime-switching parameters in the linear 3PRF filter. This new framework is denoted as Markov-switching three-pass regression filter (MS-3PRF). A key advantage of this approach is to be well-suited to handle large dimensional factor models, as opposed to the existing regime-switching factor models that can only handle models with limited dimensions due to computational complexity (see, e.g., Camacho et al. (2012)).[1] Our approach is attractive in that our estimation strategy

[1] Groen and Kapetanios (2016) show that partial least squares (and Bayesian methods) perform better than principal components when forecasting based on a large dataset with a weak factor structure. As partial least squares is obtained as a special case of the 3PRF (see Kelly and Pruitt (2015) for details), our method can be also adopted to introduce Markov switching in partial least squares regressions.

only requires to estimate a series of univariate Markov-switching regressions. As such, it is computationally straightforward to implement and offers a great deal of flexibility in modelling time variation in that we do not restrict the regime changes in the cross-sectional dimension to be governed by a single or a limited number of Markov chains.[2]

Empirically, we use the MS-3PRF approach for forecasting selected variables based on a large set of predictors. Since the seminal work of Stock and Watson (2002b), a large literature has developed to improve on the forecasting performance of the principal component approach for macroeconomic forecasting (see, e.g., Forni et al. (2005) and De Mol et al. (2008) among many others). A related work to our paper is Bai and Ng (2008) who find improvements to the principal component approach by using fewer but informative predictors. They also suggest that additional forecasting gains can be obtained when modelling non-linearities. The MS-3PRF approach is related to this strand of the literature given that factors are extracted by modelling the correlation of the predictors with the forecast target so that the estimation of the factors takes into account how informative the predictors are for the target variable. Moreover, the MS-3PRF approach captures non-linearities by modelling parameters that vary according to unobservable Markov chains.

This paper contributes to the literature along two main dimensions. First, theoretically, we provide a new framework for the estimation of high-dimensional factor models with regime switching parameters under classical inference. In a simulation experiment, we study the finite sample accuracy of the MS-3PRF forecasts compared with a number of alternatives. We find that the MS-3PRF performs well when there are instabilities in the DGP modelled via regime-switching parameters. Second, empirically, we provide evidence that the MS-3PRF performs well when forecasting major U.S. macroeconomic variables based on the McCracken and Ng (2015) dataset. Moreover, when forecasting major currencies based on a panel of exchange rates, we also find predictive gains when using the MS-3PRF approach. As such, this provides additional evidence in terms of gains one can draw from the use of factor analysis to forecast exchange rates (see, e.g., Engel et al. (2015)) as well as the importance of modelling non-linearities in this context.

This paper is organized as follows. Section 2 introduces the MS-3PRF approach and discusses its main features. Section 3 presents a Monte-Carlo experiment to study the finite sample accuracy of the MS-3PRF. Section 4 gathers empirical applications devoted to macroeconomic and exchange rate forecasting. Section 5 concludes.

[2]For example, extracting one factor from the MS-3PRF approach using a panel of 130 macroeconomic and financial variables with GDP growth as a target proxy takes about 350 seconds using a laptop with a 2.7 GHz processor and 16 GB RAM.

2 Markov-switching Three-Pass Regression Filter

2.1 The algorithm

There is now a growing literature on dynamic factor models with time-varying parameters. For example, in a Bayesian setting, Del Negro and Otrok (2008) first introduced a dynamic factor model with time-varying factor loadings. In a classical context, see Mikkelsen et al. (2015) and Eickmeier et al. (2015). However, the literature on regime-switching dynamic factor models is limited, and more importantly, restricted to small scale models (see, e.g., Chauvet (1998), Camacho et al. (2012) or Barnett et al. (2016), who use less than 10 variables and focus only on switches in the parameters governing the factor dynamics and not in the factor loadings).[3] The same is true for vector autoregression (VAR) models in that, while there is now a large (both methodological and empirical) literature on time-varying parameter VAR models, the literature using regime-switching VAR models is a lot more narrower, although there are few noticeable exceptions (see, e.g., Sims and Zha (2006) and Hubrich and Tetlow (2015)).

One of the key reasons for the absence of a significant literature on large-scale Markov-switching factor models relates to the computational challenges associated with the estimation of such models. We present here the Markov-switching three-pass regression filter, which circumvents these difficulties and offers a flexible approach in that we impose very few restrictions on the Markov processes driving the changes in the parameters of the model.

The type of setting we have in mind can be informally described as follows. There is a relatively large number N of predictors $\mathbf{x}$ from which we want to extract factors so as to forecast a target variable y. While $\mathbf{x}$ depends on two sets of common factors, say f and g (plus idiosyncratic components), y only depends on f, so that we would like to only extract f from $\mathbf{x}$. In addition, there exist proxi variables, $\mathbf{z}$, whose common components are also only driven by f. This setting is the same as that in Kelly and Pruitt (2015), who introduced the linear 3PRF for estimation of f and forecasting of y, but the key novelty is that we include time variation in the model parameters via Markov processes.

[3] In a Bayesian context, Guérin and Leiva-Leon (2016) develop an algorithm to estimate a high-dimensional factor-augmented VAR model with regime-switching parameters in the factor loadings to study the interactions between monetary policy, the stock market and the connectedness of industry-level stock returns. See also von Ganske (2016), who introduce regime-switching parameters in partial least squares regressions from a Bayesian perspective so as to forecast industry stock returns. Using a Bayesian framework, Hamilton and Owyang (2012) develop a framework for modelling common Markov-switching components in panel datasets with large cross-section and time-series dimensions to estimate turning points in U.S. state-level employment data.

More formally, let us consider the following model:

$$y_{t+1} = \beta_0(S_t) + \boldsymbol{\beta}'(S_t)\mathbf{F}_t + \eta_{t+1}(S_t), \tag{1}$$

$$\mathbf{z}_t = \lambda_0(\mathbf{S}_t) + \boldsymbol{\Lambda}(\mathbf{S_t})\mathbf{F}_t + \omega_t(\mathbf{S}_t), \tag{2}$$

$$\mathbf{x}_t = \phi_0(\mathbf{S}_t) + \boldsymbol{\Phi}(\mathbf{S_t})\mathbf{F}_t + \varepsilon_t(\mathbf{S}_t), \tag{3}$$

where y is the target variable of interest; $\mathbf{F}_t = (\mathbf{f}_t', \mathbf{g}_t')'$ are the $K = K_f + K_g$ common driving forces of all variables, the unobservable factors; S_t denotes a standard Markov-chain driving the parameters of the forecasting equation, while $\mathbf{S}_t = (S_{1,t}, S_{2,t}, ..., S_{N,t})'$ is a vector containing *variable-specific* Markov chains with M regimes driving the parameters of the factor equations, each Markov chain is governed by its own $M \times M$ transition probability matrix,

$$P_i = \begin{pmatrix} p_{i,11} & p_{i,21} & \cdots & p_{i,M1} \\ p_{i,12} & p_{i,22} & \cdots & p_{i,M2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{i,1M} & p_{i,2M} & \cdots & p_{i,MM} \end{pmatrix}, \tag{4}$$

for $i = 1, 2, ..., N$; $\boldsymbol{\beta}(S_t) = (\boldsymbol{\beta}_f'(S_t), \mathbf{0}')'$, so that y only depends on $\mathbf{f}$; $\mathbf{z}$ is a small set of L proxies that are driven by the same underlying forces as y, so that $\boldsymbol{\Lambda}(\mathbf{S_t}) = (\boldsymbol{\Lambda}_f(\mathbf{S}_t), \mathbf{0})$; $\mathbf{x}_t$ is a large set of N variables, driven by both $\mathbf{f}_t$ and $\mathbf{g}_t$; and $t = 1, ..., T$.

To achieve identification, when N and T diverge, the covariance of the loadings is assumed to be the identity matrix in each state, and the factors are orthogonal to one another.[4] For the sake of space, we refer to Kelly and Pruitt (2015) for precise conditions on the factors, allowed temporal and cross-sectional dependence of the residuals, and existence of proper central limit theorems.

Given the model in (1)-(3), our algorithm for the Markov-switching Three-Pass Regression Filter model consists of the following three steps:

- *Step 1:* Time series regressions of each element of $\mathbf{x}$, x_i, on $\mathbf{z}$; that is, run N Markov-switching regressions

$$x_{i,t} = \phi_{0,i}(S_{i,t}) + \mathbf{z}_t'\boldsymbol{\phi}_i(S_{i,t}) + \epsilon_{i,t}(S_{i,t}), \tag{5}$$

[4]More precisely, defining $\mathbf{J}_T = \mathbf{I}_T - \frac{1}{T}\iota_T\iota_T'$ where $\mathbf{I}_T$ is a T-dimensional identity matrix and ι_T a T-vector of ones (and similarly $\mathbf{J}_N$) and assuming that $N^{-1}\boldsymbol{\Phi}'(S_t)\mathbf{J}_N\boldsymbol{\Phi}(S_t) \underset{N\longrightarrow\infty}{\longrightarrow} \mathbf{P}(S_t)$, $N^{-1}\boldsymbol{\Phi}'(S_t)\mathbf{J}_N\phi_0(S_t) \underset{N\longrightarrow\infty}{\longrightarrow} \mathbf{P}_1(S_t)$, $T^{-1}\mathbf{F}'\mathbf{J}_T\mathbf{F} \underset{T\longrightarrow\infty}{\longrightarrow} \Delta_F$, for identification we require, as Kelly and Pruitt (2015), that $\mathbf{P}(\mathbf{S_t}) = \mathbf{I}$, $\mathbf{P}_1(S_t)= \mathbf{0}$ and Δ_F is diagonal, positive definite, and each diagonal element is unique.

where $i = \{1, ..., N\}$, $\epsilon_{i,t}|S_{i,t} \sim NID(0, \sigma^2(S_{i,t}))$ and keep the maximum likelihood estimate of $\boldsymbol{\phi}_i(S_{i,t})$, denoted by $\widehat{\boldsymbol{\phi}}_i(S_{i,t})$. All regime-switching models are estimated via (pseudo) maximum likelihood, and we make a normality assumption about the disturbances to write down the log-likelihood function, which is not required when estimating the linear version of the three-pass regression filter. As mentioned previously, $S_{i,t}$ is a standard Markov chain with M regimes and dynamics driven by constant transition probabilities. It is important to stress that the estimated latent processes $S_{i,t}$ differ across all cross section units i. The pattern of the regime changes in the factor loadings is therefore left unrestricted as opposed to assuming that the changes in the parameters $\phi_{0,i}$ and $\boldsymbol{\phi}_i$ are governed by a single (or a limited number of) Markov chain(s) across all cross section units. Moreover, a different number of regimes could be used across the N first pass regressions. As such, the MS-3PRF approach offers a great deal of flexibility in modelling regime changes.

- *Step 2:* Cross section regressions of $x_{i,t}$ on $\hat{\boldsymbol{\phi}}_{i,t}$; that is, run T linear regressions

$$x_{i,t} = \alpha_{0,t} + \hat{\boldsymbol{\phi}}'_{i,t}\mathbf{F}_t + \epsilon_{i,t}, \tag{6}$$

where $t = \{1, ..., T\}$ and keep (for each t) the OLS estimates $\mathbf{F}_t$. In this step, the time-varying factor loadings $\hat{\boldsymbol{\phi}}_{i,t}$ can be obtained from the first step of the algorithm by following two alternatives. First, as a weighted average of the regime-specific factor loadings:

$$\hat{\boldsymbol{\phi}}_{i,t} = \sum_{j=1}^{M} \hat{\boldsymbol{\phi}}_i(S_{i,t} = j)P(S_{i,t} = j|\Omega_T), \tag{7}$$

where $P(S_{i,t} = j|\Omega_T)$ is the smoothed probability of being in regime j given the full sample information Ω_T. Second, as a selected loading:

$$\hat{\boldsymbol{\phi}}_{i,t} = \sum_{j=1}^{M} \hat{\boldsymbol{\phi}}_i(S_{i,t} = j)I(P(S_{i,t} = j|\Omega_T)), \tag{8}$$

where $I(\cdot)$ is an indicator function that selects the regime with the highest smoothed probability, $P(S_{i,t} = j|\Omega_T)$, at time t.

- *Step 3:* Time series regressions of y_t on $\hat{\mathbf{F}}_{t-h}$; that is, run one Markov-switching regression for each forecast horizon of interest, h:

$$y_t = \beta_0(S_t) + \hat{\mathbf{F}}'_{t-h}\boldsymbol{\beta}(S_t) + \eta_t(S_t), \tag{9}$$

keep the MLE estimates $\beta_0(S_t)$ and $\boldsymbol{\beta}(S_t)$, and calculate the forecast $\hat{y}_{T+h|T}$ as:

$$\hat{y}_{T+h|T} = \sum_{j=1}^{M} \Big(P(S_{T+h} = j|\Omega_T)\hat{\beta}_0(S_{T+h} = j) + P(S_{T+h} = j|\Omega_T)\hat{\mathbf{F}}'_t\hat{\boldsymbol{\beta}}(S_{T+h} = j) \Big), \tag{10}$$

where $P(S_{T+h} = j|\Omega_T)$ is the predicted probability of being in regime j h-step-ahead given the information available up to time T, Ω_T.

In the third pass of the algorithm, the Markov-chain S_t allows us to model time variation in the intercept of the forecasting regression, which is a common source of forecast failure. Changes in the slope parameters $\boldsymbol{\beta}$ are relevant in that this also allows us to model time variation in the predictive power of the estimated factors $\hat{\mathbf{F}}_t$ for the target variable y_{t+1}. Note that one can estimate a linear model in the third step. We denote this approach as "MS-3PRF (first pass)," while "MS-3PRF (first and third pass)" refers to a situation when considering regime changes in both the first and third pass.

2.2 Assumptions and theoretical properties

The algorithm outlined in the previous subsection rests on a number of assumptions, including the choice of the number of factors and proxi variables to be used in the first step of the algorithm, as well as the number of regimes to consider for the MS-3PRF.

There are several ways to assess the number of regimes in a Markov-switching regression under a classical framework. Just to mention a few, Cho and White (2007) and Carter and Steigewald (2012) suggest the use of a quasi-likelihood ratio test, however, they ignore the Markov property of the variable S_t. Other alternatives consist in calculating goodness of fit measures that trade off the fit of the likelihood against the number of parameters (e.g. Smith et al. (2006)). For ease of illustration of the proposed approach, throughout the simulation exercises and empirical applications, we leave aside this complication and assume that predictor variables experience either $M = 1$ or $M = 2$ regimes. However, the framework can be generalized accordingly.

For the choice of the proxi variables, when there is just one $\mathbf{f}_t$ factor, Kelly and Pruitt (2015) suggest to use directly the target variable y as proxy $\mathbf{z}$. From a predictive point of view, this is a natural choice, since, in this context, one wants to extract a factor that summarizes how related the predictors are to the the predicted variable. They refer to this case as target-proxy 3PRF. In the case of more factors, they propose to use either economic theory to select indicators correlated with the target variable y, or an automated procedure, that can be implemented with the following steps, indicating a proxy by z_j with $j = 1, ..., L$.

- Pass 1: set $z_1 = y$, get the 3PRF forecast $\widehat{y}_t^1$, and the associated residuals $e_t^1 = y_t - \widehat{y}_t^1$.

- Pass 2: set $z_2 = e^1$, get the 3PRF forecast $\widehat{y}_t^2$ using z_1 and z_2 as proxies. Get the associated residuals $e_t^2 = y_t - \widehat{y}_t^2$.

- ...

- Pass L: set $z_L = e^{L-1}$, get the 3PRF forecast $\widehat{y}_t^L$ using $z_1, z_2, ...z_L$ as proxies.

For the choice of the number of factors, Kelly and Pruitt (2015) use appropriate information criteria with asymptotic optimality properties. However, empirically it can be more informative to assess the performance of different number of factors. As in the case of principal component analysis, using more factors than needed reduces forecast efficiency in finite samples but does not introduce a bias, while using fewer factors generates an omitted variable problem and therefore biases both the estimators of the loadings and the forecasts.

Kelly and Pruitt (2015) develop asymptotic theory for the linear 3PRF approach, showing that the 3PRF based forecast converges in probability to the infeasible best forecast as cross section N and sample size T become large. However, we need additional special conditions to be able to claim that their consistency results could be extended to the nonlinear case. Specifically, we need consistency of the parameter estimators for the Markov-switching models in steps 1 and 3. Douc et al. (2004) establish results concerning the consistency and asymptotic normality of the maximum likelihood estimator in Markov-switching models. For the general case of hidden Markov models, Leroux (1992) proved the consistency of the maximum likelihood estimator under mild regularity conditions. For ease of notation, we suppress the i subscript and assume that the same conditions apply to all the predictor variables in $\mathbf{x}$. In particular, for any given variable x_t, let $\{f(\cdot, \theta) : \theta \in \Theta\}$ be a family of densities on a Euclidean space with respect to $\theta_1, \theta_2, ..., \theta_M$ elements of Θ.

The characteristics of the model are parameterized by ϕ which belongs to a parameter space Φ, i.e., we have $p_{jk}(\phi)$, for $j, k = 1, 2, ..., M$, and $\theta_j(\phi) \in \Theta$, for $j = 1, 2, ..., M$. The usual case is $\phi = (p_{11}, p_{12}, ..., p_{MM}, \theta_1, \theta_2, ..., \theta_M)$, and $p_{jk}(\cdot)$ and $\theta_j(\cdot)$ equal to coordinate projections. Letting the true parameter value be denoted by ϕ^*, we assume the following conditions:

- *Assumption 1:* The $M \times M$ transition probability matrix, $[p_{jk}(\phi^*)]$, is irreducible.

- *Assumption 2:* The family of mixtures of at most M elements of $\{f(y, \theta) : \theta \in \Theta\}$ is identifiable. In other words, the finite mixture with M or fewer components determines a unique distribution.

- *Assumption 3:* For each x_t, the density $f(x_t, \cdot)$ is continuous and vanishes at infinity.

- *Assumption 4:* For each j, k, $p_{jk}(\cdot)$ and $\theta_j(\cdot)$ are continuous.
- *Assumption 5:* $E_{\phi^*}[|\log(f(X_1, \theta_j(\phi^*)))|] < \infty$ for $j = 1, 2, .., M$.
- *Assumption 6:* For every $\theta \in \Theta$, $E_{\phi^*}[\sup_{\|\theta-\theta'\|<\delta}(\log(f(X_1, \theta')))^+] < \infty$, for some $\delta > 0$.[5]

Leroux (1992) proved that under assumptions 1-6, the maximum likelihood estimator, $\hat{\phi}$, converges to the true parameter value, ϕ^*, with probability one.

Hence, based on this consistency result of the Markov-switching parameters in the first and third steps, under the above assumptions 1 to 6 and the assumptions in Kelly and Pruitt (2015), the asymptotic results from Kelly and Pruitt (2015) follow through, and thereby the MS-3PRF conserves the theoretical properties from the linear 3PRF. However, as in practical applications both T and N are finite, it is important to also assess the finite sample performance of the MS-3PRF, which is what we do in the next section.

3 Monte Carlo Simulations

In this section, we conduct Monte Carlo simulations to evaluate the finite sample properties of the MS-3PRF, focusing on its predictive performance. We compare the MS-3PRF with competing approaches that have proved to be successful in dealing with large data sets, such as, the linear three-pass regression filter (3PRF) and principal component analysis (PCA). We also use two additional benchmark models – targeted PCA (TPCA) and PC-LARS – from which factors are extracted by the method of principal component from a smaller set of predictors than the N predictors used by PCA. These two additional methods are described in the Appendix; appendix A.1 describes the hard thresholding approach or targeted PCA (TPCA) and appendix A.2 outlines the soft thresholding approach (or PC-LARS). Our simulation exercises compute the full-sample Mean Square Forecasting Errors (MSFE) to predict the target variable, y_t that is generated based on a large set of predictors, $\mathbf{x}_t = (x_{1,t}, x_{2,t}, ..., x_{N,t})'$, driven by a factor structure with regime-switching in the loadings.

[5] Where, $\|\cdot\|$ is the Euclidean distance, and $w^+ = max\{w, 0\}$.

3.1 Design

The data on $\mathbf{x}_t$ and y_t, for $t = \{1, 2, ..., T\}$, are generated following the factor structure proposed in Bates et al. (2013) and Kelly and Pruitt (2015):

$$\mathbf{x}_t = \Phi_t F_t + \varepsilon_t, \tag{11}$$

$$y_{t+1} = \Lambda F_t + \eta_t, \tag{12}$$

where $F_t = (f_t, \mathbf{g}_t')'$, $\Phi_t = (\Phi_{f,t}, \Phi_{g,t})$, $\Lambda = (1, \mathbf{0})$. The relevant and irrelevant factors are generated according to the following dynamics, respectively:

$$f_t = \rho_f f_{t-1} + u_{f,t}, \tag{13}$$

$$\mathbf{g}_t = \rho_g \mathbf{g}_{t-1} + \mathbf{u}_{g,t}, \tag{14}$$

where $u_{f,t} \sim N(0, 1)$, and $\mathbf{u}_{g,t} \sim N(0, \mathbf{\Sigma}_g)$, with $u_{f,t}$ and $\mathbf{u}_{g,t}$ uncorrelated. We consider, $K_g = 4$, irrelevant factors and, $K_f = 1$, relevant factor. The parameters in $\mathbf{\Sigma}_g$ are chosen so that irrelevant factors are dominant; that is, their variances are 1.25, 1.75, 2.25, 2.75 times larger than the relevant factor. The idiosyncratic terms are assumed to follow autoregressive dynamics,

$$\varepsilon_{it} = \alpha \varepsilon_{i,t-1} + v_{i,t}, \tag{15}$$

and to be cross-sectionally correlated; that is, $v_t = (v_{1,t}, v_{1,t}, ..., v_{N,t})'$ and it is $i.i.d$ normally distributed with covariance matrix $\Omega = (\beta^{|i-j|})_{ij}$, as in Amengual and Watson (2007). The starting values for the factors and idiosyncratic terms, f_0, $\mathbf{g}_0$, ε_{i0} are drawn from their respective stationary distributions. The disturbances, η_t, associated to the target variable equation are $i.i.d.$ normally distributed with a variance, σ_η^2, which is adjusted to ensure that the infeasible best forecast has a R^2 of 50%. The free parameters of our Monte Carlo simulations are ρ_f, ρ_g, α, β, N, and T. In line with Stock and Watson (2002a), Bates et al. (2013) and Kelly and Pruitt (2015), we consider $\rho_f = \{0.3, 0.9\}$, $\rho_g = \{0.3, 0.9\}$, $\alpha = \{0.3, 0.9\}$, $\beta = \{0, 0.5\}$, $N = \{100, 200\}$, $T = \{100, 200\}$.

The factor loadings, collected in Φ_t, experience changes between two regimes over time,

$$\Phi_t = \Phi_1 \mathbf{S}_t + \Phi_2 (1 - \mathbf{S}_t), \tag{16}$$

where $\mathbf{S}_t = (S_{1,t}, S_{2,t}, ..., S_{N,t})$ contains N dichotomous state variables, each following distinct dynamics according to a first-order Markov chain. Since the data in $\mathbf{x}_t$ are generated from a factor structure, it is mechanically subject to a certain degree of comovement. Therefore, the nonlinear relationship between the data, $\mathbf{x}_t$, and the factors, F_t, measured by the factor loadings, Φ_t, may also experience a certain degree of comovement.

To provide a more realistic data generating process that is relevant for economic data where data are generally weakly dependent (as opposed to *i.i.d.*), we model comovement in the factor loadings, which is translated in modeling comovement in the Markovian state variables contained in $\mathbf{S}_t$. In doing so, let $\tilde{S}_{i,t}$ be the state vector of the i-th sequence at time t. If the i-th sequence is in state 1 at time t then we write $\tilde{S}_{i,t} = (1,0)'$, and if it is in state 2 at time t, then we write $\tilde{S}_{i,t} = (0,1)'$. First, we generate a "seed" sequence variable $\tilde{S}_{0,t}$. For time t, compute $(q, 1-q)' = P_{00}\tilde{S}_{0,t}$, where

$$P_{00} = \begin{pmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{pmatrix}, \tag{17}$$

is the transition probability matrix, and the realization of the sequence at time $t+1$ is defined as

$$\tilde{S}_{0,t+1} = \begin{cases} (1,0)' & \text{If } \; q \geq \theta \\ (0,1)' & \text{otherwise} \end{cases} \tag{18}$$

where θ is drawn from a $U[0,1]$. Next, we generate a Markov chain $\tilde{S}_{i,t}$ conditional on the dynamics of $\tilde{S}_{0,t}$, using the following system

$$\begin{bmatrix} (q_0, 1-q_0)' \\ (q_i, 1-q_i)' \end{bmatrix} = \begin{bmatrix} \lambda_{00}P_{00} & \lambda_{0i}P_{0i} \\ \lambda_{i0}P_{i0} & \lambda_{ii}P_{ii} \end{bmatrix} \begin{bmatrix} \tilde{S}_{0,t} \\ \tilde{S}_{i,t} \end{bmatrix}, \tag{19}$$

where the coefficients λ measure the comovement between both Markov chains, with $\lambda_{jk} \geq 0$, and $\sum_{k=1}^{2} \lambda_{jk} = 1$. The matrix P_{jk} collects the transition probabilities from the states in the k-th sequence to the states in the j-th sequence.[6] Accordingly, q_k represents the state probability distribution of the k-th sequence at time $t+1$, from which the realization $\tilde{S}_{i,t+1}$ can be generated as follows,

$$\tilde{S}_{i,t+1} = \begin{cases} (1,0)' & \text{If } \; q_i \geq \theta \\ (0,1)' & \text{Otherwise} \end{cases} . \tag{20}$$

For simplicity we assume that $P_{0i} = P_{i0} = P_{ii} = P_{00}$, and that $p_{11} = p_{22}$, with $p_{11} = 0.9$. Also, we set $\lambda_{00} = \lambda_{ii} = 0.2$, to induce a relatively large degree of comovement between the state variables. Given $\tilde{S}_{0,t}$, we repeat the same procedure for $i = \{1, 2, ..., N\}$, to get all the elements in $\mathbf{S}_t$. Finally, the elements in Φ_κ, for $\kappa = \{1, 2\}$, are generated from a $N(\phi_\kappa, \sigma_\phi)$, with $\phi_1 = 0.5$, $\phi_2 = 1.5$, and $\sigma_\phi = 0.1$.

We also study the case when $\mathbf{x}_t$ and y_t are generated by following the same processes described above but with the factor loadings being driven by Markov chains that are totally independent from each other and do not experience comovement. Finally, we assess the performance of the proposed nonlinear approaches in estimating the factors under different features of regime-switching factor loadings dynamics.

[6] Ching et al. (2002) proposed a multivariate Markov chain approach for modeling multiple categorical data sequences.

3.2 Models and evaluation criteria

We perform $L = 500$ Monte Carlo replications for each configuration of parameters ρ_f, ρ_g, α, and β, and sample sizes, T and N. Once $\mathbf{x}_t$ and y_t are generated, we apply the MS-3PRF to extract the factor and predict the target variable. In particular, first, we estimate a time series (Markov-switching) regression, $x_{i,t} = \mathbf{z}_t'\boldsymbol{\phi}_i(S_{i,t}) + \epsilon_{i,t}$, for $i = \{1, 2, ..., N\}$. For simplicity, we take the proxy variable as the target variable, $\mathbf{z}_t' = y_t$. Second, we run a cross section (OLS) regression, $x_{i,t} = \hat{\boldsymbol{\phi}}_{i,t}'\mathbf{F}_t + \epsilon_{i,t}$, for $t = 1, 2, ..., T$, using the weighted average of the regime-switching factor loadings obtained in the previous step. Third, we run a time series (OLS) regression, $y_t = \hat{\mathbf{F}}_{t-1}'\boldsymbol{\beta} + \eta_t$, and produce the forecast, $\hat{y}_{t+1} = \hat{\mathbf{F}}_t'\hat{\boldsymbol{\beta}}$, obtained with the MS-3PRF approach introduced in this paper. Also, we produce forecasts with the version of the MS-3PRF when the loadings are selected instead of being averaged (MSS-3PRF); that is, the time-varying loadings are set to the regime-specific loadings of the most likely regime. To ease the computational burden, in our Monte Carlo simulations, we do not model regime switches in the third pass of the algorithm. This is not detrimental for our simulation exercise, since we are only interested in studying situations characterized by instabilities in the factor loadings (and not time instability in the relation between the predicted variable and the predictors). However, in the empirical applications, we consider the case of regime switches in the third pass.

We compare the predictive performance of the two variants of our proposed method with several benchmark methodologies. First, we compute the forecast obtained with the linear version of the 3PRF, proposed in Kelly and Pruitt (2015). Second, Bates et al. (2013) show that principal components estimation (PCA) methods can be applied to consistently estimate dynamic factor models under certain instabilities in the loadings. Therefore, we compute the forecast obtained with the method of principal components. Third, Bai and Ng (2008) argue that the principal components methodology, as it stands, does not take into account the predictive ability of $\mathbf{x}_t$ for y_{t+h} when the factors are estimated. Therefore, Bai and Ng (2008) propose to use only predictors that are informative for y_t in estimating the factors in order to take explicitly into account that the object of interest is ultimately the forecast of y_t. Accordingly, we also compute the forecast obtained with targeted PCA (TPCA). Fourth, we consider the Elastic Net soft-thresholding rules, which are special cases of the 'Least Angle Regression' (LARS) algorithm developed in Efron et al. (2004), and compute the forecast using the PC-LARS method. Ultimately, our focus is on comparing the median in-sample MSFE over the L replications associated to each of the six methods (two MS-3PRF approaches and four competitors) to evaluate their relative predictive performance. Finally, in our simulation experiments, we assume that the number of relevant factors and the number of regimes are known (i.e., across all procedures, we extract one factor and the non-linear models consider the case of two regimes).

3.3 Results

Table 1 reports the simulation results associated to the different configurations of parameters, methodologies and degrees of instability (regime-switching) in the data. In particular, the upper part of Table 1 present the MSFE for the cases when all factor loadings exhibit regime-switching dynamics. First, the MSS-3PRF exhibits the lowest MFSE for most of the cases, indicating that it performs best in terns of predictive performance. In particular, the MSS-3PRF outperforms the MS-3PRF, linear 3PRF, PCA, TPCA and PC-LARS. Notice also that, in general, the MS-3PRF exhibits the second best forecasting performance, suggesting that the nonlinear frameworks, MS-3PRF and especially MSS-3PRF, are able to capture in a better way instabilities in the relationship between the set of predictors and its common factor.

Second, TPCA tends to provide the best performance when the irrelevant factors, and the idiosyncratic terms, have a low autocorrelation, but the relevant factor is highly autocorrelated. If the irrelevant factors behave close to white noise processes, while the relevant factors experience a high autocorrelation, then TPCA would be able to easily identify the variables highly associated with the relevant factors so as to obtain a reliable estimation of the relevant factors, as long as the target variable to be predicted, y_t, does not approximately behave as a white noise. Also notice that the linear 3PRF outperforms in most cases PCA and PC-LARS. This implies that in the presence of instabilities in the loadings, the linear 3PRF takes better advantage of both the time series and cross sectional dimensions to provide a more accurate estimation of the underlying factor than PCA and PC-LARS.

Third, the scenarios associated with low autocorrelation in the irrelevant factors, ρ_g, yield the highest MSFE. The fact that irrelevant factors (i) behave close to a white noise, (ii) are linked to $\mathbf{x}_t$ trough regime-switching loadings, and (iii) are dominant, make them able to introduce a relatively large amount of noise into the set of predictors, creating more difficulty for all the methods to provide more accurate estimates of the underlying relevant factor and consequently better forecasts for y_t. In particular, when not only the irrelevant factors, but also the idiosyncratic terms are closer to behave as a white noise (that is, $\rho_g = 0.3$ and $\alpha = 0.3$) the forecasting performance of all methods deteriorates due to the reason just described. These results indicate that in the presence of instabilities in the loadings, a lower autocorrelation in any of the components driving the observed data (predictors and target variables) is detrimental for the predictive performance of all factor extraction methods studied in this paper.

Fourth, the performance of the methods varies with the size of both cross sectional and time series dimensions. In particular, while a larger N helps to improve the predictive ability of the methods, a larger T appears to be detrimental for the predictive ability.

Notice that when holding constant the cross sectional dimension N and letting grow the time series dimension T; that is, when comparing the case of $T = 100, N = 100$ versus the case of $T = 200, N = 100$, there is an overall increase in the corresponding MSFE. However, when keeping constant T and letting grow N, that is, when comparing the case of $T = 100, N = 100$ versus the case of $T = 100, N = 200$, there is an overall reduction in the MSFE, pointing to an increase in the predictive ability across all the methods. As such, this suggests that higher ratio of N to T is better. This is because a larger cross sectional dimension is helpful for more precise estimation of the underlying factor, and thereby leads to better predictive performance.

When dealing with large dimensional data sets, the substantial heterogeneity in the data may be accompanied by different degrees of instabilities contained in the predictors $\mathbf{x}_t$. Therefore, we repeat these simulation exercises along the lines of Bates et al. (2013) and let only a subset J of variables, randomly selected from a uniform distribution, exhibit regime-switching factor loadings. The lower part of Table 1 reports the case when the share of variables experiencing instabilities in their loadings is 0.25. The results are relatively similar to the ones obtained with all the variables experiencing instabilities in the loadings in that the MSS-3PRF obtains the best forecasting performance. However, the method showing the second best forecasting performance is TPCA. Another important difference is that the increase in the predictive ability of the methods when N increase and T remains constant is not systematically observed through all the scenarios under consideration, as it was the case when all the variables where experiencing instabilities in the loadings. This is simply because in this case there is less cross sectional information that can help to capture the nonlinearities in the time dimension of the data. We performed the same simulation exercise when the share of variables experiencing instabilities in the loadings is 0.50 and 0.75. The results remained qualitatively unchanged. However, to conserve space, we do not report the results for these DGPs, but these results are available upon request.

The results reported in Table 1 were based on a data generating process where the factor loadings associated to each predictor variable were driven by their own Markov chain, and these Markov chains were assumed to experience a high degree of interdependence between them in order to mimic the behavior that macroeconomic and financial data usually tend to exhibit. However, we are interested in assessing the performance of the methods when the assumption of interdependent Markov chains is no longer valid; that is, when the loadings depend on Markov chains that are independent from each other. The upper part of Table 2 reports the MSFE for the case when factors loading are driven by independent Markov chains. The results provide the same consistent message obtained from the previous exercises, the MSS-3PRF shows the most accurate predictive performance among the methods under consideration, followed by the MS-3PRF. These results imply that regardless of the relationship between the Markov chains driving the factor loadings, the nonlinear methods proposed in this paper consistently outperform the linear competing methods.

Finally, we compare the forecasting performance of the methods under consideration when the dynamics of the factor is also subject to regime changes. In particular, we repeat the Monte Carlo exercises assuming that the DGP of the factor is given by

$$f_t = \rho_{f,t} f_{t-1} + u_{f,t}, \tag{21}$$

where the dynamics of the autoregressive coefficient depends on the "seed" Markov chain,

$$\rho_{f,t} = \rho_{1f} S_{0,t} + \rho_{0f}(1 - S_{0,t}). \tag{22}$$

The lower part of Table 2 shows the corresponding MSFE for the different scenarios, showing that MSS-3PRF and the MS-3PRF perform better than the linear benchmarks for most of the cases. Notice that, in this case, the MS-3PRF model is not correctly specified given that there is no regime switching in the factor loadings.

Overall, conditional on our DGPs, we can conclude that, on average, the MSS-3PRF is the framework best able to capture instabilities in the relationship between the set of predictors and its common factor, followed by the MS-3PRF. Both nonlinear frameworks outperform linear approaches. Regarding the linear frameworks, in general, TPCA outperforms the 3PRF and the PC-LARS, and PCA obtains the weakest forecasting performance.

4 Empirical Applications

The first application is related to exchange rate forecasting. This is highly relevant given that it has long been recognized that non-linearities play an important role in the dynamics of exchange rates (see, e.g., the early contribution in Chinn (1991) and more recently Rossi (2013) and Abbate and Marcellino (2016)). However, it is only recently that the literature on exchange rate forecasting has concentrated on the role and importance of factors for predicting exchange rates (see, e.g., Engel et al. (2015) in a linear context). Putting the MS-3PRF approach at work in the context of exchange rate forecasts is highly relevant in that this allows us to combine the non-linear dynamics observed in exchange rate movements with the factor structure driving systematic variations in exchange rates, which has recently gained traction in the exchange rate forecasting literature. Our second empirical application is a standard macroeconomic forecasting application in that we use the McCracken and Ng (2015) dataset so as to forecast economic activity in the United States.

4.1 Forecasting exchange rates

In this first forecasting exercise, we construct factors from a cross-section of nominal bilateral U.S. dollar (USD) exchange rates against a panel of twenty-six currencies. We extract factors from the MS-3PRF, the MSS-3PRF, the linear 3PRF, PCA, TPCA and PC-LARS. We then use the resulting factors to forecast selected bilateral exchange rates. (All currency pairs use the USD as numéraire.) The choice of the dataset draws from the exercise in Greenaway-McGrevy et al. (2016). The dataset is monthly and the full sample size extends from January 1995 to December 2015. The data are obtained from the International Financial Statistics of the International Monetary Fund, and the monthly data are taken as the monthly average of daily data. The dataset consists of the currencies of Australia (AUS), Brazil (BRA), Canada (CAN), Chile (CHI), Columbia (COL), the Czech Republic (CZE), the euro (EUR), Hungary (HUN), Iceland (ICE), India (IND), Israel (ISR), Japan (JPN), Korea (KOR), Mexico (MEX), Norway (NOR), New Zealand (NZE), the Philippines (PHI), Poland (POL), Romania (ROM), Singapore (SIN), South Africa (RSA), Sweden (SWE), Switzerland (SUI), Taiwan (TAI), Turkey (TUR) and the United Kingdom (GBR).[7]

We consider forecast horizons, h, ranging from 1 month to 12 months and report predictive results for selected major currencies: the euro, the British pound, the Japanese yen and the Canadian dollar. The first estimation sample runs from February 1995 to March 2007, and it is recursively expanded until we reach the end of the estimation sample. Hence, the forecast evaluation period extends from April 2007 to December 2015. As in the Monte-Carlo experiment, we compare the forecasts obtained from the MS-3PRF with forecasts derived from principal component analysis, the linear 3PRF, targeted PCA (TPCA) and PC-LARS. Moreover, we use two different versions of the MS-3PRF, one with regime-switching only in the first step (i.e., in the factor loadings), and one with regime-switching parameters in the first and third step (i.e., in the factor loadings and the parameters of the forecasting equation). In the first step of the algorithm for the MS-3PRF approach, we model regime changes in all parameters of the model (i.e., intercept, slope and innovation variance), since we obtained stronger fit – as measured by the SIC – with such a specification. Note also that we consider a model with two factors across all methods. The choice of the number of factors follows the modelling choices in Engel et al. (2015), Greenaway-McGrevy et al. (2016) and Verdelhan (2015), but, qualitatively, our results are robust to the use of one or three factors in the predictive equation. We also include the MSS-3PRF approaches in the set of models we consider (both versions; that is, with regime switches in the first pass only and regime switches in the first and third pass).

[7] Data for the euro before January 1999 and Taiwan were obtained from the U.S. Federal Reserve G.5 table (monthly average of daily data).

All exchange rate series are taken as the first difference of their logarithm before performing factor analysis. In the case of the 3PRF approaches, we use the automatic proxy-selection procedure from Kelly and Pruitt (2015); that is, we use the exchange rate we are interested in forecasting as a target proxy when extracting the first factor and then proceed sequentially as outlined in Table 2 from Kelly and Pruitt (2015). For PCA, TPCA and PC-LARS, we standardize the data recursively in the estimation before estimating the factors. In contrast, the 3PRF approaches do not require to standardize the data before estimation.

For the prediction step, in the linear cases; that is, for PCA, TPCA, PC-LARS, 3PRF, MS-3PRF (first pass) and MSS-3PRF (first pass), the h-period ahead forecasts for a specific currency $R^j_{t+h|t}$ are constructed in level based on the following equation

$$R^j_{t+h|t} = R^j_t(1 + \hat{\alpha} + \boldsymbol{F'_t}\hat{\boldsymbol{\beta}}), \tag{23}$$

where $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ are obtained from the following regression (for simplicity of the notation, we omit h subscripts from the coefficients α and $\boldsymbol{\beta}$)

$$\Delta r^j_{t,h} = \alpha + \boldsymbol{F}'_{t-h}\boldsymbol{\beta} + \epsilon_t, \tag{24}$$

where $\Delta r^j_{t,h}$ indicates the h-period change in the logarithm of the exchange rate R^j_t (i.e., $\Delta r^j_{t,h} = \ln(R^j_t) - \ln(R^j_{t-h})$), $\boldsymbol{F}_t$ indicates the factors extracted from either PCA, TPCA, PC-LARS, 3PRF or MS-3PRF approaches. In the case of the MS-3PRF (first and third pass) and MSS-3PRF (first and third pass), equation (24) is modified as follows

$$\Delta r^j_{t,h} = \alpha(S^j_t) + \boldsymbol{F}'_{t-h}\boldsymbol{\beta}(S^j_t) + \epsilon_t(S^j_t), \tag{25}$$

where S^j_t is a two-regime Markov-chain – distinct across all predicted currencies j – with constant transition probabilities. In the case of the MS-3PRF (first and third pass), as commonly done in forecasting exercises with Markov-switching models, the forecasts are calculated as a weighted average of forecasts conditional on the parameters being in a given regime. The predicted probabilities of being in a given regime k, h periods ahead are obtained recursively as

$$P(S^j_{t+h|t} = k) = \sum_{i=1}^{2} p^j_{ik} P(S^j_{t+h-1|t} = k), \tag{26}$$

where p^j_{ik} indicates the constant transition probabilities and 2 is the total number of regimes.

In the case of the MSS-3PRF (first and third pass) approach, the forecasts are calculated conditional on being in a given regime, i.e., the predicted probabilities are obtained as

$$P(S^j_{t+h|t} = k) = I(P(S^j_{t|t} = k)), \tag{27}$$

where $I(\cdot)$ is an indicator variable that indicates the regimes with the highest smoothed probability at the origin of the forecast horizon. As such, this corresponds to the approach often used to plot (regime-specific) impulse responses in MS-VAR models (see, e.g., Hubrich and Tetlow (2015)).

As an illustration of the MS-3PRF approach to extract factors from a panel of exchange rates, Figure 1 reports the Markov-switching factor loadings based on the MS-3PRF approach using the Canadian dollar as a target proxy; that is, the factor loadings associated with the first factor. From this figure, one can see that there is substantial time variation in factor loadings for a number of currencies (e.g., the Australian dollar or the New Zealand dollar) whereas for other currencies, there is little time variation in the factor loadings (e.g., the euro and Swiss franc).

A number of additional comments are in order. First, we do not use observable factors to model the dynamics of exchange rates. Verdelhan (2015) finds that exchange rate variations are driven by a two-factor structure: a U.S. dollar factor that serves as a proxy for global macroeconomic risk and a carry factor which is interpreted as capturing uncertainty risk. Our analysis differs from this study in that our focus is on out-of-sample predictive ability. Hence, we do not aim at providing a structural interpretation to the factors we extract.

Table 3 reports point forecasting results for specific currencies: the Canadian dollar (CAD), the euro (EUR), the Japanese yen (JPY) and the British pound (GBP), all relative to the USD. These are G7 currencies, and among the most traded currency pairs according to the BIS Triennal Central Bank Survey.[8] The point forecast results are presented as the MSFE of a specific approach relative to the MSFE obtained from the no-change forecast. The no-change forecast is the standard benchmark in the exchange rate forecasting literature (see, e.g., Rossi (2013)). This table also reports the results of the Diebold and Mariano (1995) test of equal out-of-sample predictive accuracy using the no-change forecast as a benchmark.[9] First, the models' forecasting performance relative to the no-change forecast is typically the strongest for forecast horizon $h = 1$ (except for the JPY-USD). The improvement in forecast accuracy relative to the random walk is also statistically significant according to the Diebold and Mariano test of equal MSFE when forecasting the Canadian dollar at forecast horizon $h = 1$ across most approaches (this is also true to a lesser extent for the British pound). Second, the PC-LARS approach performs best for forecast horizon $h = 1$ when forecasting the euro and the British pound, albeit it is closely followed by PCA and TPCA in those cases. Third, for the Canadian dollar and the Japanese yen, modelling

[8]See `http://www.bis.org/publ/rpfx16fx.pdf`.

[9]The Diebold and Mariano (1995) test of equal out-of-sample predictive accuracy is reported to give a sense of statistical significance of the point forecasting results. However, this test is based on the population MSPE (not the actual MSPE) so that this test tends to reject the null of equal MSPEs too often.

time variation in the forecasting equation is relevant in that this leads to substantial forecasting improvement over the no-change forecast at distant forecast horizons h={9,12} for the Japanese yen and h={2,3,6,9,12} for the Canadian dollar, using the MS-3PRF (first and third pass) approach. [10]

Next, Table 4 shows the directional accuracy forecasting results, which are broadly in line with the point forecast results. Under the null hypothesis of no directional accuracy, one would expect a success ratio of 0.5. In this table, we also report the results of the Pesaran and Timmermann (2009) test to evaluate the statistical significance of the directional accuracy results. Across all forecasting approaches, the success ratios tend to be stronger for forecast horizon $h = 1$, except for the JPY-USD. It is also interesting to note that the success ratios are especially strong at distant forecast horizons for selected currencies, as high as 67.3 per cent for the CAD-USD and 79.7 per cent for the JPY-USD in the case of the MS-3PRF and MSS-3PRF with regime changes in the first and third pass.

4.2 Forecasting economic activity

In this application, we use the McCracken and Ng (2015) dataset to forecast eight major quarterly U.S. variables: GDP, Consumption, Investment, Exports, Imports, Total Hours, GDP inflation and PCE inflation.[11] We implemented the following outlier corrections to the predictors: observations of the transformed series with absolute median deviations larger than 6 times the inter quartile range were replaced with the median value of the preceding five observations. The full sample extends from 1960Q3 to 2015Q3. In the forecasting exercise, the first estimation sample extends from 1960Q3 to 1984Q4, and it is recursively expanded until we reach the end of the sample. We consider forecast horizons, h, ranging from 1 quarter to 8 quarters. We use eight competing approaches: principal component analysis (PCA) from which we extract 5 factors from the underlying dataset, but we only use the first one in the forecasting equation; PCA where hard thresholding has been performed before extracting the first principal component to forecast (TPCA); PCA where soft thresholding has been performed before extracting the first principal component to forecast (PC-LARS); linear 3PRF; MS-3PRF and MSS-3PRF with regime-switching parameters in the first pass only; and MS-3PRF and MSS-3PRF with regime-switching

[10] Admittedly, in the case of the euro, the forecasting performance of the MS-3PRF (first and third pass) and the MSS-3PRF (first and third pass) approaches deteriorate as the forecast horizon lengthens, suggesting that it is not always relevant to model regime shifts in the forecasting equation.

[11] Data descriptions and details on data transformation are available online at `https://research.stlouisfed.org/econ/mccracken/fred-databases/Appendix_Tables_Update.pdf`, the slight modifications we made to the original dataset are reported in the appendix.

parameters in the first and third passes. For the 3PRF approaches, we use one factor and use the predicted variable as a target proxy in the first step of the 3PRF approach (target-proxy 3PRF).[12]

We first report results from an in-sample exercise. Figure 2 shows the estimated factors across all six methods (PCA, TPCA, PC-LARS, 3PRF, MS-3PRF and MSS-3PRF; the latter three methods use GDP growth as a target variable). This shows that the factor estimates are relatively similar across approaches, and that they closely follow the U.S. business cycle. As in McCracken and Ng (2015), we calculate diffusion indices ($\hat{F}_t$) based on the partial sums of the factor estimates $\hat{f}_t$; that is, $\hat{F}_t = \sum_{j=1}^{t} \hat{f}_j$. (The reason for doing so is that diffusion indices summarize information contained in the trend as opposed to the "raw" factors that are estimated on stationary data so that the resulting factors are too volatile for turning point analysis.) The factors f_t are extracted with the six aforementioned methods. We then implemented the Bry and Boschan (1971) algorithm to estimate expansions and recessions from these diffusion indices.[13] The resulting classification of U.S. business cycles obtained from the MS-3PRF diffusion index has the strongest correlation with the NBER dummy variable of expansions and recessions (0.563) followed by the 3PRF diffusion index (0.543), MSS-3PRF diffusion index (0.506), PC-LARS diffusion index (0.497), PCA diffusion index (0.486) and TPCA diffusion index (0.467). Moreover, only the MS-3PRF and MSS-3PRF approaches obtain a perfect classification of recessions, while PCA, TPCA and 3PRF diffusion indices have a near perfect classification of recessions (these three methods identify the 1973-1974 recession with a one-quarter lag). As such, this suggests that the MS-3PRF approach has important information related to the state of the business cycle that is not necessarily reflected in competing approaches.

An attractive feature of Markov-switching models is their ability to endogenously estimate regimes. Figure 3 shows a heatmap of the smoothed probability of being in the first regime (associated with adverse business cycle conditions) for the in-sample factor loadings obtained from the first step of the MS-3PRF approach. First, it is interesting to note that, across all series, there is substantial time variation in the smoothed probability, suggesting that there is evidence in favor of regime shifts in the factor loadings. Second, the timing of the shifts in the factor loadings coincides with the changes in business cycle phases

[12] When estimating the number of factors using information criteria, it is common to find a large number of factors summarizing the co-movements of U.S. macroeconomic variables (e.g., McCracken and Ng (2015) estimate eight factors in the FRED-MD monthly macroeconomic database). However, in the forecasting exercise, in line with the literature, we use the first factor in the predictive equation. This corresponds to a real economic activity factor that closely follows the U.S. business cycle dynamics (see figure 2). Using the first two factors in the predictive equation led to little changes in the forecasting performance.

[13] The Bry and Boschan (1971) is a non-parametric method to estimate cycles in time series. We implemented the quarterly version of the Bry and Boschan (1971) algorithm from Harding and Pagan (2002), using the GAUSS code available at `http://www.ncer.edu.au/resources/data-and-code.php`.

for a large number of series (e.g., output and income, as well as labor market variables). Additional evidence on regime shifts in the factor loadings is provided in Figure 4. This figure shows that there is substantial variation in the factor loadings for the unemployment rate and industrial production related to the state of the business cycle. Selected financial and credit variables (S&P500 returns and consumer loans) also exhibit substantial time variation, suggesting that the assumption of constant factor loadings often employed with this type of dataset is likely to be too restrictive.

A few additional comments related to the out-of-sample forecasting exercise are required. First, note that macroeconomic variables are typically subject to substantial revisions and different publication lags. In this empirical exercise, we abstract from this issue, and consider revised data. While this is not a fully realistic approach from a practitioners' perspective, there is no reason to think that one specific approach would benefit more from this simplification. Hence, this remains a useful forecasting exercise to compare the relative merits of each forecasting approach. Second, across all approaches, quarterly factors are extracted from the monthly dataset of McCracken and Ng (2015), where quarterly data are taken as quarterly averages of monthly data before performing factor analysis. Obviously, alternative temporal aggregation schemes could be adopted, but we found that the in-sample correlation of the real activity factor was very strong compared with a situation where one would use the last monthly observation of the quarter as a quarterly observation before performing factor analysis (about 0.95 between these two aggregation schemes across the different factor approaches). [14] Our temporal aggregation scheme is standard in the literature (see, e.g., section 6.1 in Stock and Watson (2016)) and we leave the issue of a mixed-frequency setting to future research.[15] Third, the forecasts are constructed as follows

$$y_{t+h|t} = \hat{\alpha} + \hat{\beta}(L)\hat{f}_t + \hat{\gamma}(L)y_t, \tag{28}$$

where $\beta(L)$ and $\gamma(L)$ are finite order lag polynomials, whose lag lengths are obtained with the SIC at the beginning of the forecasting exercise, using a maximum lag length of 6 for $\gamma(L)$ and 3 for $\beta(L)$. All predicted variables y_t are taken as the first difference of their logarithm. For the MS-3PRF and MSS-3PRF with switches in the first and third passes, we consider regime-switching parameters in all parameters of equation (28) and in the variance of the error term.

[14]This result still holds when doing PCA at a monthly frequency and then aggregating the factor at a quarterly frequency.

[15]As a side note, the first pass of the 3PRF filter could possibly accommodate mixed-frequency data using the techniques outlined in Foroni et al. (2015); whereas, in the third pass of the filter, unrestricted MIDAS polynomials could be used as in Hepenstrick and Marcellino (2016), and regime switching parameters in the mixed-frequency predictive equation could be modelled as in Guérin and Marcellino (2013).

Tables 5 and 6 show the out-of-sample forecasting results. All results are reported relative to the forecasts obtained from PCA. Hence, a number below 1 indicates that a given approach outperforms PCA. We also report the results of the Diebold and Mariano (1995) test of equal out-of-sample predictive accuracy using PCA as a benchmark. Overall, across all forecast horizons and predicted variables (64 cases), the MS-3PRF and MSS-3PRF obtain the best forecasting results in 35 cases, the linear 3PRF in 16 cases, PC-LARS in 10 cases and TPCA in 1 case. In the remainder of the cases, PCA performs best. It is interesting to note that the MSS-3PRF (first and third pass) approach performs best for forecasting inflation (both PCE inflation and GDP inflation) and it does so significantly according to the Diebold and Mariano (1995) test at long forecast horizons (i.e., for $h > 4$ for GDP inflation and $h > 3$ for PCE inflation). The rationale for the strong forecasting performance of the MSS-3PRF (first and third pass) when forecasting inflation at distant forecast horizons is that over the full sample we look at, there is a decline in the rate of inflation. Moreover, as the forecast horizon increases, forecasts tend to converge toward the full sample mean of inflation. However, forecasts from the MSS-3PRF (first and third pass) approach give zero weight to the high inflation observed in the early part of the sample as opposed to the forecasts from the competing approaches, which necessarily include information from high inflation episodes at distant forecast horizons. When forecasting aggregate economic activity (GDP), the MSS-3PRF approach performs best at short forecasting horizon and the linear 3PRF tend to perform best for distant forecast horizons, and the improvements in forecast accuracy relative to PCA are typically statistically significant according to the Diebold and Mariano (1995) test. The MS-3PRF forecasting approaches perform particularly well relative to PCA when predicting export, import and hours worked, whereas PC-LARS performs well for forecasting investment and hours worked at relatively short forecast horizons (i.e., $h < 4$).

5 Conclusion

In this paper, we extended the linear three-pass regression filter to settings where parameters can vary according to Markov processes, introducing the Markov-Switching Three-Pass Regression Filter. A key advantage of our framework is to circumvent the computational difficulties associated with the estimation of a large scale dynamic factor model with regime-switching parameters without foregoing flexibility in modelling choices.

In both simulation and empirical examples, our method compares favorably with existing alternatives in terms of forecasting performance. The MS-3PRF approach is also attractive beyond forecasting applications. For example, the MS-3PRF approach would easily allow one to model regime-switching correlations often observed in finance in a high-dimensional setting. This could be relevant in the context of the growing literature aiming at measuring network connectedness among financial firms or asset classes (see, e.g., Billio et al. (2012) or Diebold and Yilmaz (2014)). Likewise, the MS-3PRF framework could be used in the context of structural factor-augmented VAR models that are commonly used in macroeconomics. Overall, thanks to its generality and ease of implementation, the MS-3PRF approach offers a promising framework to model regime changes in high-dimensional settings for a large class of applications in macroeconomics and finance.

References

Aastveit, K., Carreiro, A., Clark, T., and Marcellino, M. (2016). Have standard VARs remained stable since the crisis? *Journal of Applied Econometrics*, forthcoming.

Abbate, A. and Marcellino, M. (2016). Modelling and forecasting exchange rates with time-varying parameter models. *Journal of the Royal Statistical Society, Series A*, forthcoming.

Amengual, D. and Watson, M. W. (2007). Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel. *Journal of Business & Economic Statistics*, 25:91–96.

Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317.

Barnett, W. A., Chauvet, M., and Leiva-Leon, D. (2016). Real-time nowcasting of nominal GDP with structural breaks. *Journal of Econometrics*, 191(2):312–324.

Bates, B. J., Plagborg-Moller, M., Stock, J. H., and Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177(2):289–304.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559.

Bry, G. and Boschan, C. (1971). Cyclical Analysis of Time Series: Procedures and Computer Programs. *National Bureau of Economic Research*.

Camacho, M., Pérez-Quirós, G., and Poncela, P. (2012). Markov-switching dynamic factor models in real time. CEPR Discussion Papers 8866, C.E.P.R. Discussion Papers.

Canova, F. (1993). Modelling and Forecasting Exchange Rates with a Bayesian Time-Varying Coefficient Model. *Journal of Economic Dynamics and Control*, 17:233–261.

Carter, A. and Steigewald, D. (2012). Testing for Regime Switching: A Comment. *Econometrica*, 80:1809–1812.

Chauvet, M. (1998). An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching. *International Economic Review*, 39(4):969–96.

Cheng, X., Liao, Z., and Schorfheide, F. (2016). Shrinkage Estimation of High-Dimensional Factor Models with Structural Instabilities. *Review of Economic Studies*, forthcoming.

Ching, W.-K., Fung, E., and Ng, M. (2002). A multivariate Markov chain model for categorical data sequences and its applications in demand predictions. *IMA Journal of Management Mathematics*, 13:187–199.

Chinn, M. D. (1991). Some linear and nonlinear thoughts on exchange rates. *Journal of International Money and Finance*, 10(2):214–230.

Cho, J. S. and White, H. (2007). Testing for Regime Switching. *Econometrica*, 75:1671–1720.

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.

De Mol, C., Giannone, D., and Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2):318–328.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

Diebold, F. X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Douc, R., Moulines, E., and Rydén, T. (2004). Asymptotic properties of the maximum likelihood estimor in autoregressive models with Markov regime. *Annals of Statistics*, 32(5):2254–2304.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least Angle Regression. *Annals of Statistics*, 32(2):407–499.

Eickmeier, S., Lemke, W., and Marcellino, M. (2015). Classical time varying factor-augmented vector auto-regressive modelsestimation, forecasting and structural analysis. *Journal of the Royal Statistical Society Series A*, 178(3):493–533.

Engel, C., Mark, N. C., and West, K. D. (2015). Factor Model Forecasts of Exchange Rates. *Econometric Reviews*, 34(1-2):32–55.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, 100:830–840.

Foroni, C., Guérin, P., and Marcellino, M. (2015). Using low frequency information for predicting high frequency variables. *Norges Bank Working Paper*, 2015/13.

Greenaway-McGrevy, R., Mark, N., Sul, D., and Wu, J.-L. (2016). Identifying Exchange Rate Common Factors. *Mimeo Notre Dame*.

Groen, J. and Kapetanios, G. (2016). Revisiting useful approaches to data-rich macroeconomic forecasting. *Computational Statistics & Data Analysis*, 100:221–239.

Guérin, P. and Leiva-Leon, D. (2016). Monetary Policy, Stock Market and Sectoral Comovement. *Mimeo*.

Guérin, P. and Marcellino, M. (2013). Markov-Switching MIDAS Models. *Journal of Business & Economic Statistics*, 31(1):45–56.

Hamilton, J. D. and Owyang, M. T. (2012). The Propagation of Regional Recessions. *The Review of Economics and Statistics*, 94(4):935–947.

Harding, D. and Pagan, A. (2002). Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics*, 49(2):365–381.

Hepenstrick, C. and Marcellino, M. (2016). Forecasting with Large Unbalanced Datasets: The Mixed-Frequency Three-Pass Regression Filter. *SNB Working Papers*, 2016-04.

Hubrich, K. and Tetlow, R. J. (2015). Financial stress and economic dynamics: The transmission of crises. *Journal of Monetary Economics*, 70(C):100–115.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

Leroux, B. (1992). Maximum-likelihood estimation for hidden Markov models. *Stochastic Processes and their Applications*, 40:127–143.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. Technical report.

von Ganske, J. (2016). A Regime Switching Partial Least Squares Approach to Forecasting Industry Stock Returns. *Mimeo Edhec-Risk Institute*.

McCracken, M. W. and Ng, S. (2015). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, forthcoming.

Mikkelsen, J. G., Hillebrand, E., and Urga, G. (2015). Maximum Likelihood Estimation of Time-Varying Loadings in High-Dimensional Factor Models. CREATES Research Papers 2015-61, School of Economics and Management, University of Aarhus.

Pesaran, M. H. and Timmermann, A. (2009). Testing Dependence Among Serially Correlated Multicategory Variables. *Journal of the American Statistical Association*, 104(485):325–337.

Rossi, B. (2013). Exchange Rate Predictability. *Journal of Economic Literature*, 51(4):1063–1119.

Sims, C. (1993). A 9 Variable Probabilistic Macroeconomic Forecasting Model. *Business Cycles, Indicators and Forecasting, NBER studies in business cycles*, 28:179–214.

Sims, C. A. and Zha, T. (2006). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review*, 96(1):54–81.

Smith, A., Prasad, N., and Tsai, C.-L. (2006). Markov-switching Model Selection using Kullback-Leibler Divergence. *Journal of Econometrics*, 134:553–577.

Stock, J. and Watson, M. (2002a). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, 97:1167–1179.

Stock, J. H. and Watson, M. W. (2002b). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business & Economic Statistics*, 20(2):147–62.

Stock, J. H. and Watson, M. W. (2016). Factor Models and Structural Vector Autoregressions in Macroeconomics. *Mimeo*.

Verdelhan, A. (2015). The Share of Systematic Variation in Bilateral Exchange Rates. *Journal of Finance*, forthcoming.

Appendix

A.1 Description of the hard thresholding forecasting approach

The hard thresholding algorithm consists of the following steps (This description partly stems from Bai and Ng (2008).)

1. For each variable $x_{i,t}$, perform a time series regression of the variable to forecast y_t on $x_{i,t}$ and a constant. Let t_i denote the t statistic associated with $x_{i,t}$
2. Let k^*_α be the number of series whose $|t_i|$ exceeds a threshold significance level, α. In our application, we use a threshold of 1.65, which corresponds to a one-sided 5 per cent significance level for the t test.
3. Let $\chi_t(\alpha) = (x_{[1t]}, ..., x_{[k^*_\alpha]})$ be the corresponding set of predictors. Estimate f_t from $\chi_t(\alpha)$ by the method of principal component.
4. Estimate equation (28) to calculate the h period ahead forecast y_{t+h}.

This approach is denoted as *TPCA*.

A.2 Description of the soft thresholding forecasting approach

The soft thresholding approach we adopt follows from the least angle regressions (LARS) method described in Bai and Ng (2008). In detail, we select the set of the first K predictors $x_{i,t}$ selected by forward stagewise selection regressions to extract principal component(s). In the macroeconomic forecasting application, we use $K = 30$ predictors, since it is the number of predictors retained by Bai and Ng (2008) and Kelly and Pruitt (2015) when forecasting macroeconomic variables with a similar dataset than ours. For the exchange rate forecasting application, we retain the first $K = 10$ predictors ordered by LARS to extract principal components, which corresponds to slightly more than a third of the total number of predictors (26). Finally, in the Monte Carlo experiments, we set $K = 30$ across all DGPs.

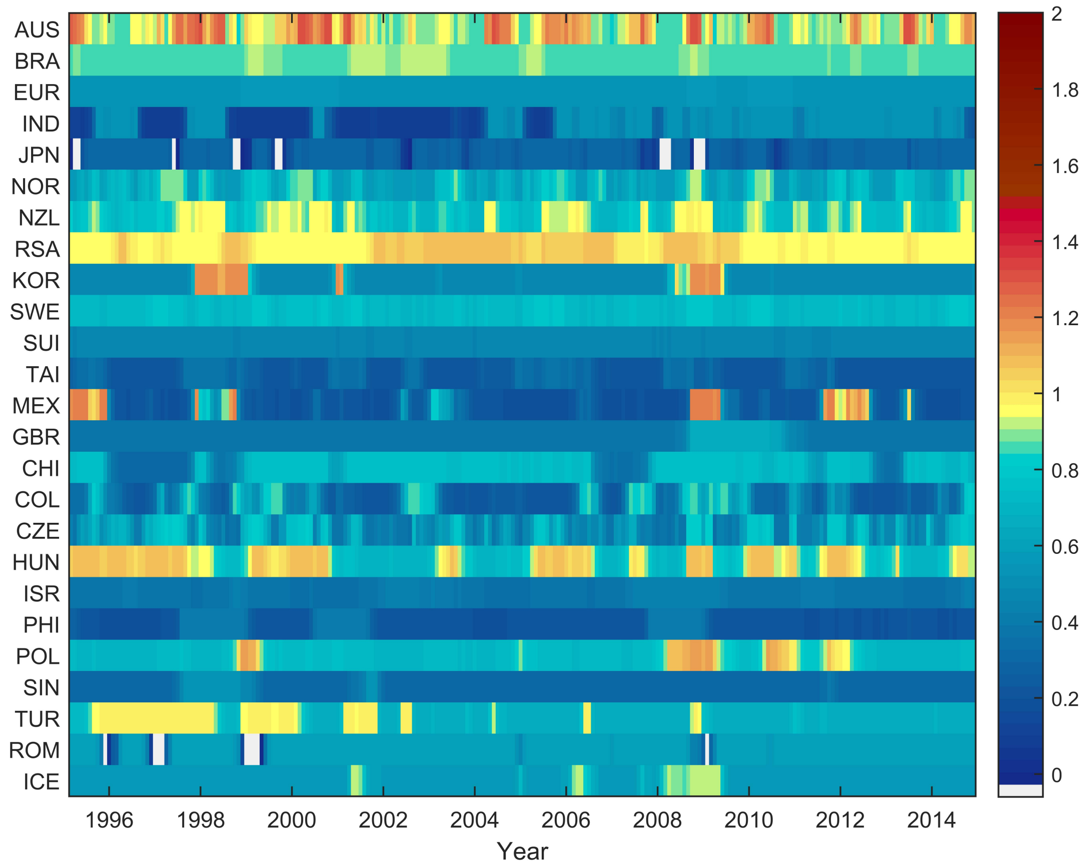
This approach is denoted as *PC-LARS*.

A.3 Additional Details on the macroeconomic forecasting exercise

In the macroeconomic forecasting empirical application, we use the May 2016 vintage of the McCracken and Ng (2015) dataset as available online at `https://research.stlouisfed.org/econ/mccracken/fred-databases/`. We use the exact same transformation as suggested by McCracken and Ng (2015). However, due to missing observations, we omit the following five series in our analysis (fred mnemonics are in parenthesis): "New Orders for Consumer Goods" (ACOGNO), "New Orders for Nondefense Capital Goods" (ANDENOx), "Trade-weighted U.S. Dollar Index: Major Currencies" (TWEXMMTH), "Consumer Sentiment Index" (UMCSENTx) and the "VXO" (VXOCLSx).

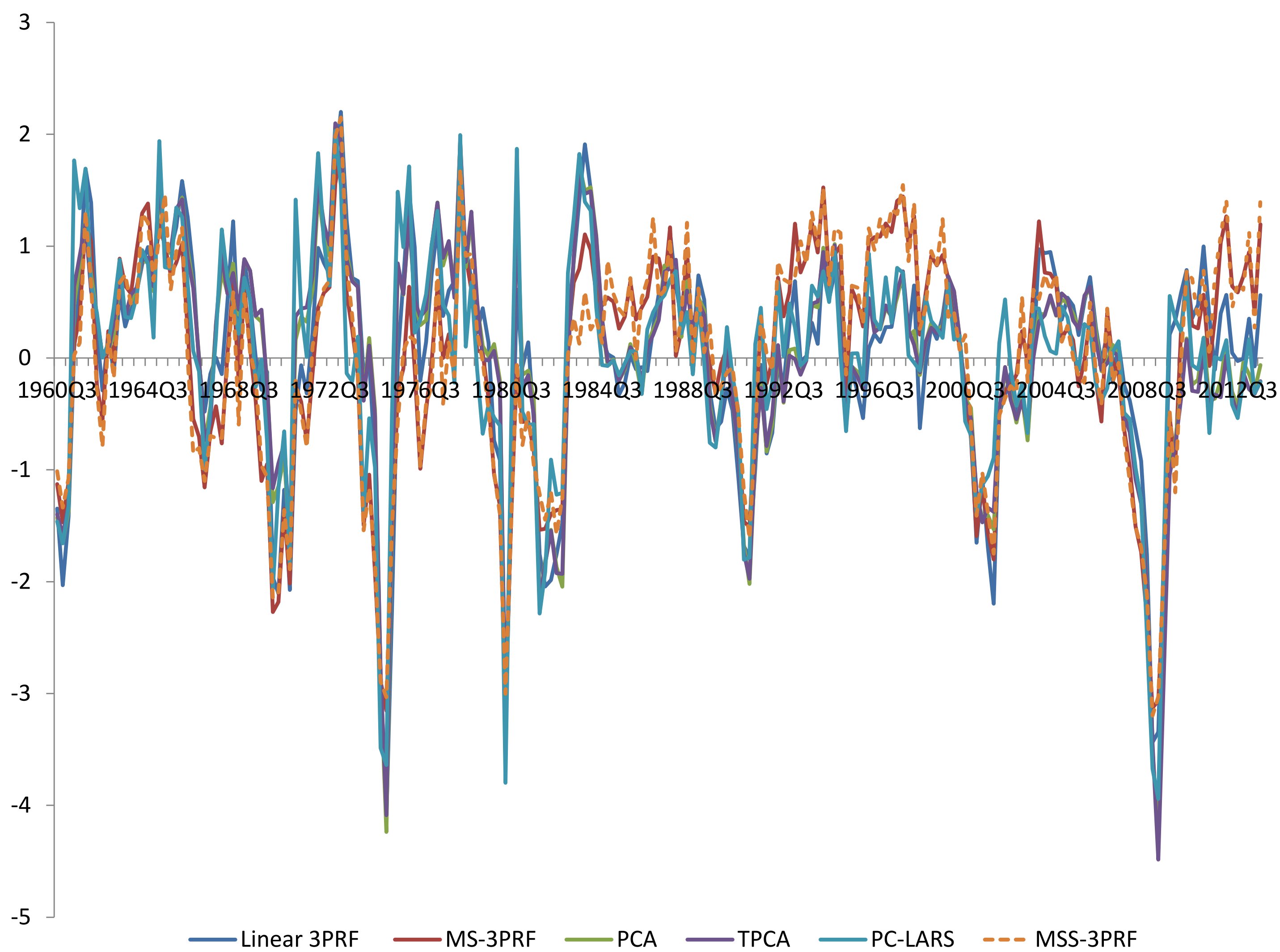
The eight variables we forecast in the macroeconomic forecasting application are : Gross Domestic Product (GDPQ@USNA), Personal Consumption Expenditures (CQ@USNA), Gross Private Domestic Investment (IQ@USNA), Exports of Goods & Services (XQ@USNA), Imports of Goods & Services (MQ@USNA), Business Sector: Hours of All Persons (LXBH@USECON), Gross Domestic Product: Chain Price Index (JGDP@USNA) and Personal Consumption Expenditures: Chain Price Index (JC@USNA). (Haver/Analytics mnemonics are in parenthesis.)

Figure 1: Markov-switching factor loadings – Canadian dollar as a target proxy



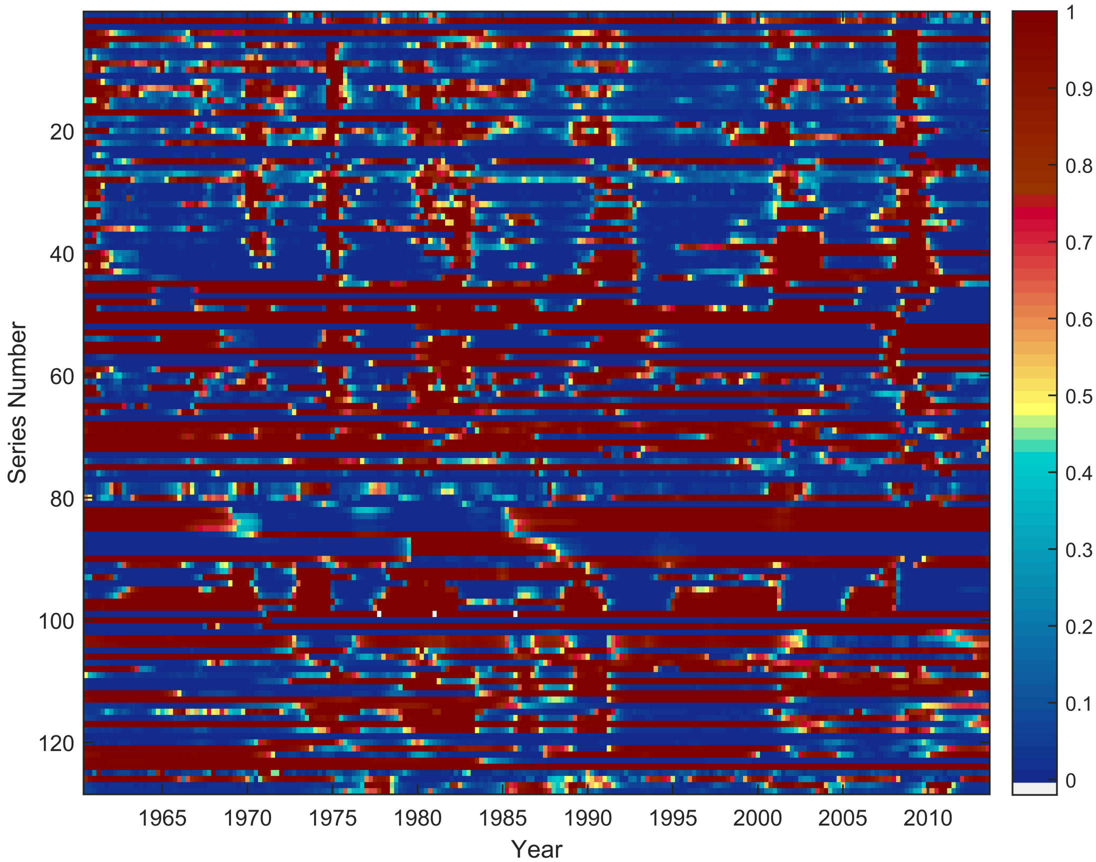
Note: Dark red indicates higher values for the factor loadings obtained with the MS-3PRF approach with the Canadian dollar as a target proxy. Three-letter country codes follow the convention from the International Olympic Committee except for Taiwan labelled as TAI.

Figure 2: Factor estimates across different approaches



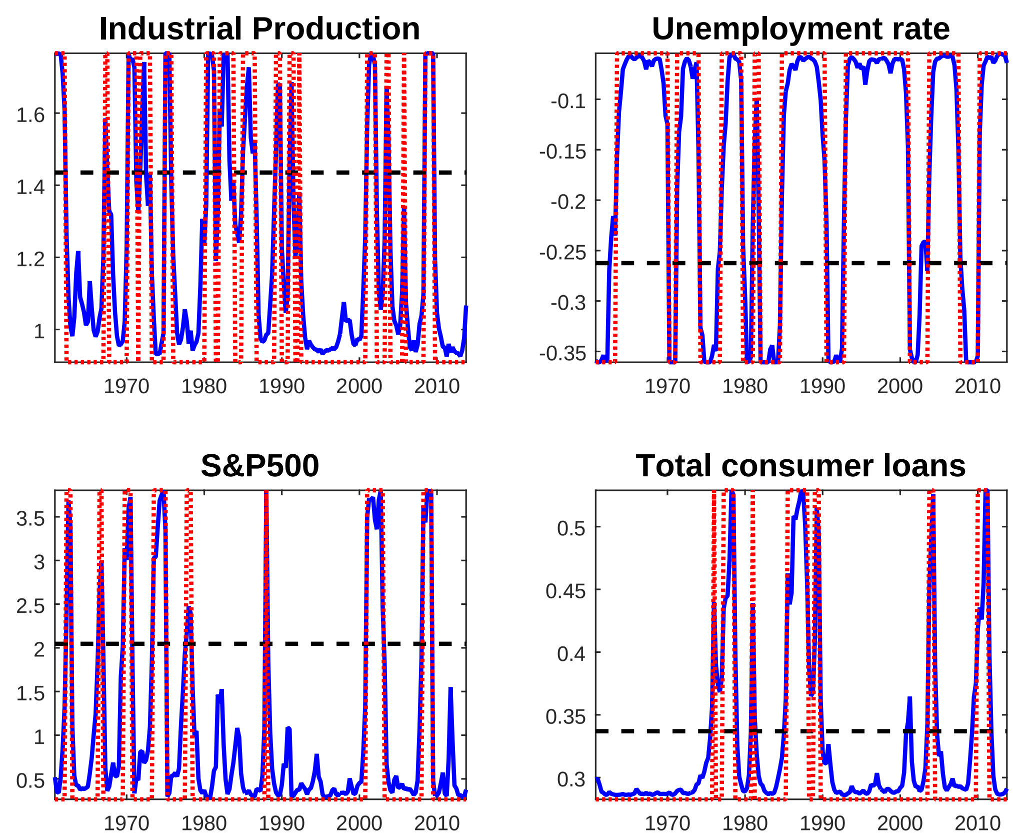
Note: Factor estimates across different approaches: linear 3PRF, Markov-switching 3PRF (both MS-3PRF and MSS-3PRF), PCA, TPCA and PC-LARS. GDP growth is used as a target proxy for the 3PRF approaches.

Figure 3: Probability of being in the first regime for the factor loadings



Note: Dark red indicates higher value for the probability of being in the first regime, which is normalized to correspond to the lowest intercept of the two regimes. Factor loadings are obtained from the MS-3PRF approach with GDP growth as a target proxy.

Figure 4: Factor loadings for selected variables



Note: Factor loadings for selected variables obtained from the linear 3PRF (black dashed line), MS-3PRF (blue solid line) and MSS-3PRF (red dotted line). GDP growth is used as a target proxy for the 3PRF approaches.

Table 1: Simulation results with different degrees of instabilities in the loadings

ρ_f	ρ_g	α	β	MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS	MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS
Instability in 100 per cent of the loadings															
				T=100 , N=100						T=100 , N=200					
0.3	0.9	0.3	0	2.00	**1.97**	2.00	2.05	2.00	2.05	1.99	**1.98**	2.01	2.07	2.00	2.07
0.3	0.9	0.3	0.5	2.00	**1.97**	2.00	2.05	2.00	2.05	1.98	**1.96**	1.97	2.03	1.96	2.02
0.3	0.9	0.9	0	1.91	**1.86**	1.94	2.08	2.00	2.07	1.91	**1.85**	1.97	2.05	1.98	2.06
0.3	0.9	0.9	0.5	1.93	**1.90**	1.97	2.08	1.99	2.08	1.91	**1.82**	1.95	2.07	1.98	2.06
0.9	0.3	0.3	0	8.57	8.51	8.55	8.60	**8.26**	8.54	9.12	9.05	9.13	9.11	**8.78**	9.04
0.9	0.3	0.3	0.5	8.67	8.50	8.72	8.77	**8.50**	8.72	8.23	8.21	8.35	8.29	**8.02**	8.25
0.9	0.3	0.9	0	8.02	**7.62**	8.49	8.69	8.20	8.75	8.01	**7.60**	8.77	8.89	8.56	8.84
0.9	0.3	0.9	0.5	8.44	**8.04**	8.86	8.99	8.68	8.99	8.01	**7.51**	8.64	8.80	8.35	8.71
				T=200 , N=100						T=200 , N=200					
0.3	0.9	0.3	0	2.03	2.02	2.03	2.06	**2.02**	2.07	2.04	**2.04**	2.06	2.10	2.04	2.10
0.3	0.9	0.3	0.5	2.05	**2.03**	2.05	2.09	2.04	2.09	2.05	**2.02**	2.05	2.09	2.03	2.09
0.3	0.9	0.9	0	1.96	**1.91**	2.00	2.08	2.01	2.08	1.94	**1.90**	1.99	2.07	1.99	2.06
0.3	0.9	0.9	0.5	1.96	**1.92**	2.01	2.08	2.02	2.08	1.95	**1.90**	2.00	2.08	1.99	2.08
0.9	0.3	0.3	0	9.23	9.14	9.24	9.26	**9.08**	9.26	9.46	9.33	9.42	9.36	**9.27**	9.34
0.9	0.3	0.3	0.5	9.27	9.25	9.26	9.25	**9.22**	9.25	9.31	**9.17**	9.30	9.36	9.26	9.31
0.9	0.3	0.9	0	8.64	**8.40**	9.10	9.17	8.98	9.14	8.94	**8.69**	9.42	9.55	9.43	9.53
0.9	0.3	0.9	0.5	8.71	**8.39**	9.13	9.19	9.04	9.19	9.12	**8.95**	9.58	9.62	9.39	9.60
Instability in 25 per cent of the loadings															
				T=100 , N=100						T=100 , N=200					
0.3	0.9	0.3	0	1.93	**1.89**	1.93	2.04	1.92	2.04	1.91	**1.86**	1.93	2.04	1.90	2.04
0.3	0.9	0.3	0.5	1.96	**1.92**	1.98	2.08	1.94	2.08	1.94	**1.89**	1.96	2.04	1.90	2.05
0.3	0.9	0.9	0	1.88	**1.79**	1.88	2.12	2.02	2.13	1.89	**1.80**	1.90	2.15	2.03	2.15
0.3	0.9	0.9	0.5	1.89	**1.83**	1.91	2.12	2.01	2.12	1.86	**1.79**	1.88	2.12	1.99	2.11
0.9	0.3	0.3	0	8.67	**8.34**	9.02	9.32	8.53	9.14	8.10	**7.81**	8.39	8.64	7.96	8.31
0.9	0.3	0.3	0.5	8.37	8.05	8.53	8.76	**7.93**	8.56	8.48	**8.15**	8.71	9.03	8.16	8.71
0.9	0.3	0.9	0	6.99	**6.41**	7.59	8.48	7.56	8.36	7.20	**6.70**	8.14	8.56	7.72	8.19
0.9	0.3	0.9	0.5	7.19	**6.70**	8.18	8.88	8.05	8.73	7.41	**6.91**	8.30	8.74	7.95	8.41
				T=200 , N=100						T=200 , N=200					
0.3	0.9	0.3	0	1.93	1.89	1.95	2.05	**1.89**	2.03	1.94	**1.92**	1.98	2.07	1.93	2.06
0.3	0.9	0.3	0.5	1.94	1.92	1.95	2.05	**1.92**	2.06	1.96	1.92	1.98	2.08	**1.92**	2.07
0.3	0.9	0.9	0	1.88	**1.81**	1.89	2.09	1.96	2.09	1.88	**1.79**	1.88	2.10	1.96	2.11
0.3	0.9	0.9	0.5	1.86	**1.81**	1.89	2.07	1.95	2.07	1.92	**1.82**	1.91	2.12	1.99	2.12
0.9	0.3	0.3	0	8.76	**8.46**	9.01	9.11	8.74	8.94	8.91	**8.56**	9.09	9.24	8.90	8.89
0.9	0.3	0.3	0.5	8.62	**8.38**	8.91	9.10	8.76	8.89	8.64	**8.35**	8.91	9.05	8.64	8.86
0.9	0.3	0.9	0	7.94	**7.38**	8.68	9.09	8.63	8.86	7.85	**7.31**	8.89	9.33	8.79	8.98
0.9	0.3	0.9	0.5	7.85	**7.30**	8.74	9.13	8.69	9.00	7.87	**7.27**	8.83	9.16	8.75	8.82

Note: The table reports the median MSFE based on 500 replications. Serial correlation in the factors is governed by ρ_f and ρ_g, while α and β govern serial and cross sectional correlation in the predictors' residuals, respectively. Entries in bold represent the lowest median MSFE for each specification. See text for additional details.

Table 2: Simulation results with different relationships of loadings instabilities

Regime changes in the factor loadings are governed by independent Markov chains															
ρ_f	ρ_g	α	β	T=100 , N=100						T=100 , N=200					
				MS-3PRF	MSS-3PRF	3PRF	PCA	TPC	LARS	MS-3PRF	MSS-3PRF	3PRF	PCA	TPC	LARS
0.3	0.9	0.3	0	1.97	**1.94**	1.98	2.06	2.00	2.06	1.97	**1.97**	1.96	2.05	1.97	2.04
0.3	0.9	0.3	0.5	1.97	**1.94**	1.98	2.04	1.97	2.04	1.91	**1.86**	1.93	2.08	2.00	2.09
0.3	0.9	0.9	0	1.91	**1.86**	1.93	2.08	2.00	2.09	1.91	**1.84**	1.94	2.10	1.99	2.10
0.3	0.9	0.9	0.5	1.88	**1.84**	1.93	2.08	1.96	2.07	1.89	**1.83**	1.94	2.08	1.99	2.08
0.9	0.3	0.3	0	8.69	**8.44**	8.69	8.78	8.49	8.72	8.65	**8.51**	8.57	8.59	8.32	8.57
0.9	0.3	0.3	0.5	8.50	**8.33**	8.55	8.68	8.34	8.60	8.60	**8.36**	8.55	8.68	8.45	8.57
0.9	0.3	0.9	0	8.05	**7.79**	8.57	8.71	8.15	8.61	7.97	**7.74**	8.54	8.79	8.38	8.65
0.9	0.3	0.9	0.5	8.08	**7.66**	8.72	8.78	8.54	8.84	7.76	**7.39**	8.55	8.70	8.26	8.61
ρ_f	ρ_g	α	β	T=200 , N=100						T=200 , N=200					
				MS-3PRF	MSS-3PRF	3PRF	PCA	TPC	LARS	MS-3PRF	MSS-3PRF	3PRF	PCA	TPC	LARS
0.3	0.9	0.3	0	2.00	**1.98**	2.01	2.06	2.00	2.06	2.02	2.00	2.00	2.06	**1.99**	2.06
0.3	0.9	0.3	0.5	2.02	2.00	2.00	2.06	**2.00**	2.05	2.01	**1.98**	2.00	2.07	2.00	2.06
0.3	0.9	0.9	0	1.94	**1.88**	1.96	2.07	1.98	2.06	1.92	**1.86**	1.96	2.06	1.97	2.06
0.3	0.9	0.9	0.5	1.93	**1.89**	1.97	2.06	1.98	2.06	1.94	**1.89**	1.99	2.08	1.98	2.07
0.9	0.3	0.3	0	8.99	**8.88**	9.10	9.17	9.03	9.13	9.22	**9.06**	9.29	9.30	9.12	9.28
0.9	0.3	0.3	0.5	9.22	9.16	9.27	9.24	**9.13**	9.24	9.01	8.86	8.90	8.92	**8.80**	8.83
0.9	0.3	0.9	0	8.90	**8.60**	9.32	9.48	9.24	9.45	8.77	**8.34**	9.33	9.43	9.15	9.38
0.9	0.3	0.9	0.5	8.83	**8.39**	9.24	9.34	9.21	9.33	8.72	**8.41**	9.19	9.39	9.17	9.30

Markov-switching autoregressive dynamics for the relevant factor f																
ρ_{f0}	ρ_{f1}	ρ_g	α	β	T=100 , N=100						T=100 , N=200					
					MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS	MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS
0.1	0.9	0.9	0.3	0	3.84	**3.71**	3.84	4.00	3.92	3.98	3.72	**3.63**	3.71	3.85	3.72	3.82
0.1	0.9	0.9	0.3	0.5	3.65	**3.55**	3.67	3.76	3.72	3.75	3.60	**3.47**	3.66	3.82	3.71	3.80
0.1	0.9	0.9	0.9	0	3.44	**3.23**	3.59	3.86	3.78	3.86	3.33	**3.11**	3.50	3.70	3.60	3.69
0.1	0.9	0.9	0.9	0.5	3.66	**3.47**	3.82	4.05	3.96	4.07	3.53	**3.31**	3.68	3.90	3.78	3.89
0.3	0.7	0.3	0.3	0	2.64	2.63	2.63	2.65	**2.59**	2.64	2.57	2.55	2.56	2.57	**2.51**	2.57
0.3	0.7	0.3	0.3	0.5	2.61	2.59	2.58	2.61	**2.51**	2.60	2.58	2.56	2.56	2.59	**2.51**	2.58
0.3	0.7	0.3	0.9	0	2.40	**2.29**	2.49	2.60	2.49	2.59	2.39	**2.30**	2.47	2.60	2.48	2.60
0.3	0.7	0.3	0.9	0.5	2.41	**2.34**	2.48	2.62	2.49	2.61	2.38	**2.29**	2.46	2.58	2.46	2.57
ρ_{f0}	ρ_{f1}	ρ_g	α	β	T=200 , N=100						T=200 , N=200					
					MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS	MS-3PRF	MSS-3PRF	3PRF	PCA	TPCA	PC-LARS
0.1	0.9	0.9	0.3	0	4.08	**4.01**	4.09	4.21	4.16	4.19	4.02	**3.90**	3.98	4.03	3.99	4.03
0.1	0.9	0.9	0.3	0.5	4.09	**4.02**	4.12	4.24	4.19	4.22	4.03	**3.99**	4.02	4.13	4.04	4.11
0.1	0.9	0.9	0.9	0	3.77	**3.50**	3.96	4.11	4.02	4.11	3.89	**3.62**	4.20	4.37	4.28	4.35
0.1	0.9	0.9	0.9	0.5	3.89	**3.72**	4.19	4.37	4.29	4.35	3.92	**3.68**	4.12	4.24	4.19	4.20
0.3	0.7	0.3	0.3	0	2.75	2.76	2.72	2.74	**2.69**	2.73	2.78	2.76	2.74	2.78	**2.71**	2.77
0.3	0.7	0.3	0.3	0.5	2.73	2.72	2.69	2.72	**2.67**	2.71	2.74	2.74	2.69	2.72	**2.65**	2.71
0.3	0.7	0.3	0.9	0	2.57	**2.50**	2.66	2.75	2.65	2.74	2.51	**2.42**	2.60	2.69	2.60	2.68
0.3	0.7	0.3	0.9	0.5	2.54	**2.49**	2.66	2.72	2.66	2.73	2.54	**2.49**	2.63	2.71	2.61	2.70

Note: The table reports the median MSFE based on 500 replications. Serial correlation in the relevant factor during each of the two regimes is governed by ρ_{f0} and ρ_{f1}, serial correlation in the irrelevant factors is governed by ρ_g, while α and β govern serial and cross sectional correlations in the predictors' residuals. Entries in bold represent the lowest median MSFE for each specification. See text for additional details.

Table 3: Out-of-Sample Exchange Rate Forecasting: Selected Currencies (RMSFE)

Forecast horizon	1	2	3	6	9	12
CAD–USD						
PCA	**0.871***	0.951	0.980	1.001	1.094	1.175
TPCA	0.872*	0.955	0.986	0.996	1.088	1.182
PC-LARS	0.872	**0.949**	0.986	0.998	1.095	1.157
Linear 3PRF	0.910*	0.959	0.997	1.001	1.043	1.057
MS-3PRF (first pass)	0.915*	0.975	0.994	1.006	1.018	1.045
MS-3PRF (first and third pass)	0.994	1.080	**0.964**	**0.938**	**0.895**	**0.949**
MSS-3PRF (first pass)	0.916*	0.953	0.981	0.987	1.019	1.047
MSS-3PRF (first and third pass)	1.150	1.204	0.995	0.972	1.148	1.072
EUR–USD						
PCA	0.988	1.033	1.000	1.034	1.085	1.138
TPCA	0.995	1.049	1.010	1.061	1.109	1.155
PC-LARS	**0.961**	1.028	1.010	1.072	1.141	1.207
Linear 3PRF	1.014	1.010	1.016	1.049	1.081	1.148
MS-3PRF (first pass)	0.986	1.028	1.011	1.022	1.071	1.127
MS-3PRF (first and third pass)	1.078	1.037	1.117	1.170	1.126	1.270
MSS-3PRF (first pass)	0.996	1.049	1.049	1.088	1.111	1.174
MSS-3PRF (first and third pass)	1.234	1.108	1.306	1.329	1.463	2.043
JPY–USD						
PCA	1.091	1.093	1.108	1.070	1.101	1.145
TPCA	1.081	1.111	1.098	1.066	1.085	1.121
PC-LARS	1.066	1.066	1.078	1.078	1.074	1.117
Linear 3PRF	1.019	1.046	1.055	1.031	1.068	1.105
MS-3PRF (first pass)	1.019	1.035	1.038	1.055	1.075	1.094
MS-3PRF (first and third pass)	1.035	0.984	1.025	1.126	0.879	0.877*
MSS-3PRF (first pass)	1.032	1.023	1.024	1.045	1.082	1.072
MSS-3PRF (first and third pass)	1.119	1.075	1.187	1.236	**0.695****	**0.872**
GBP–USD						
PCA	0.803*	0.977	1.043	1.045	1.064	1.091
TPCA	0.796*	0.965	1.023	1.043	1.064	1.091
PC-LARS	**0.784***	0.953	1.029	1.035	1.073	1.114
Linear 3PRF	0.875*	**0.934**	**0.989**	1.029	1.059	1.094
MS-3PRF (first pass)	0.893	0.980	1.014	1.034	1.084	1.110
MS-3PRF (first and third pass)	1.679	1.201	1.049	1.082	1.131	1.079
MSS-3PRF (first pass)	0.889	0.982	1.013	1.039	1.078	1.113
MSS-3PRF (first and third pass)	1.337	1.382	1.351	1.422	1.408	1.252

Note: This table shows the Relative Mean Square Forecast Error (RMSFE) for selected currency pairs (CAD–USD, EUR–USD, JPY–USD and GBP–USD) using PCA, TPCA, PC-LARS, linear 3PRF, MS-3PRF (first pass), MS-3PRF (first and third pass), MSS-3PRF (first pass) and MSS-3PRF (first and third pass) as forecasting approaches. An entry smaller than 1 indicates that a given approach outperforms the random walk model. Entries in bold indicate the best performing approach for a specific horizon. Statistically significant reductions in the MSFE relative to the random-walk according to the Diebold-Mariano test are indicated by asterisks (* denotes significance at the 10% level and ** denotes significance at the 5% level).

Table 4: Out-of-Sample Exchange Rate Forecasting: Selected Currencies (Success Ratios)

Forecast horizon	1	2	3	6	9	12
CAD–USD						
PCA	0.637**	0.602**	0.575*	0.522	0.504	0.407
TPCA	0.637**	0.620**	0.593**	0.522	0.487	0.407
PC-LARS	0.611**	0.602**	0.513	0.487	0.469	0.407
Linear 3PRF	0.611**	**0.620****	0.558**	0.513	0.522	0.478
MS-3PRF (first pass)	0.620**	0.575**	0.540	0.487	0.522	0.460
MS-3PRF (first and third pass)	0.558	0.566**	**0.690****	**0.655****	0.646**	0.673**
MSS-3PRF (first pass)	**0.646****	0.611**	0.558	0.504	0.522	0.460
MSS-3PRF (first and third pass)	0.504	0.522	0.620**	0.575	**0.673****	**0.681****
EUR–USD						
PCA	0.549	0.487	0.540	0.504	0.416	0.372
TPCA	0.558	0.487	0.566	0.487	0.443	0.354
PC-LARS	**0.566***	0.531	0.549	0.540	0.460	0.389
Linear 3PRF	0.531	0.496	0.540	0.496	0.434	0.398
MS-3PRF (first pass)	0.549	0.513	**0.584**	**0.540**	0.487	0.460
MS-3PRF (first and third pass)	0.540	0.496	0.522	0.451	0.425	0.398
MSS-3PRF (first pass)	0.549	0.496	0.540	0.443	0.460	0.416
MSS-3PRF (first and third pass)	0.496	**0.575**	0.540	0.487	0.425	0.345
JPY–USD						
PCA	0.460	0.496	0.469	0.496	0.504	0.425
TPCA	0.460	0.434	0.451	0.504	0.478	0.425
PC-LARS	0.531	0.504	0.496	0.487	0.469	0.425
Linear 3PRF	0.575**	0.549*	**0.531**	0.522	0.460	0.504
MS-3PRF (first pass)	0.496	0.487	0.443	0.443	0.487	0.460
MS-3PRF (first and third pass)	**0.593****	**0.575***	0.487	0.522	0.726**	0.584
MSS-3PRF (first pass)	0.513	0.513	0.531	0.531	0.407	0.469
MSS-3PRF (first and third pass)	0.504	0.566**	0.496	**0.566**	**0.797****	**0.611***
GBP–USD						
PCA	0.593**	0.540	0.513	0.434	0.522	0.575
TPCA	0.593**	0.540	0.504	0.469	**0.558****	**0.584****
PC-LARS	**0.611****	0.593**	0.531	**0.531**	0.425	0.496
Linear 3PRF	0.584**	**0.593****	**0.602****	0.513	0.504	0.478
MS-3PRF (first pass)	0.549	0.566*	0.549	0.504	0.504	0.451
MS-3PRF (first and third pass)	0.460	0.540	0.593**	0.504	0.425	0.522
MSS-3PRF (first pass)	0.522	0.522	0.558	0.504	0.531	0.558
MSS-3PRF (first and third pass)	0.460	0.522	0.531	0.496	0.416	0.504

Note: This table shows the Success Ratios for selected currency pairs (CAD–USD, EUR–USD, JPY–USD and GBP–USD) using PCA, TPCA, PC-LARS, linear 3PRF, MS-3PRF (first pass), MS-3PRF (first and third pass), MSS-3PRF (first pass) and MSS-3PRF (first and third pass) as forecasting approaches. Under the null hypothesis of no directional accuracy, one would expect a success ratio of 0.5. Higher ratios indicate an improvement over the no-change forecast. Entries in bold indicate the best performing approach for a specific horizon. Statistically significant improvements in directional accuracy according to the Pesaran-Timmermann test are indicated by asterisks (* denotes significance at the 10% level and ** denotes significance at the 5% level).

Table 5: Out-of-Sample Macroeconomic Forecasting

Forecast horizon	1	2	3	4	5	6	7	8
				GDP				
MS-3PRF (first pass)	1.151	0.967*	0.975*	0.973	**0.936****	0.976	1.013	0.979
MS-3PRF (first and third pass)	1.095	1.029	1.306	1.196	1.026	1.147	1.169	1.104
MSS-3PRF (first pass)	1.150	**0.961****	0.975*	0.943**	0.943**	0.973	1.020	0.959
MSS-3PRF (first and third pass)	**0.962**	1.114	1.172	1.246	1.226	1.172	1.118	1.109
Linear 3PRF	1.077	0.971**	1.009	0.983	0.943**	**0.945****	**0.978****	**0.916****
TPCA	0.966**	0.986**	0.990**	0.996	0.997	1.006	1.003	1.004
PC-LARS	1.024	0.963*	**0.950****	**0.935****	0.944**	0.962	1.003	0.950*
				Consumption				
MS-3PRF (first pass)	1.054	0.963	1.001	0.933	**0.950**	0.956	0.955	0.956
MS-3PRF (first and third pass)	1.073	1.029	1.368	1.099	1.164	1.058	1.109	1.168
MSS-3PRF (first pass)	1.035	**0.958**	0.987	**0.920**	0.956	0.944*	0.952	0.953
MSS-3PRF (first and third pass)	1.055	1.149	1.256	1.421	1.309	1.467	1.300	1.311
Linear 3PRF	0.970	0.959**	**0.969****	0.949**	0.954**	**0.926****	**0.939****	**0.945****
TPCA	**0.939****	0.961**	0.981**	0.986*	0.997	1.008	1.020	1.009
PC-LARS	1.028	1.011	0.972**	1.016	0.982	0.990	0.961	0.985
				Investment				
MS-3PRF (first pass)	0.933	0.996	1.039	1.039	1.000	0.995	1.013	1.010
MS-3PRF (first and third pass)	1.181	1.106	1.186	1.397	1.207	1.140	1.104	1.065
MSS-3PRF (first pass)	0.963	0.998	1.047	1.056	1.001	1.006	1.015	1.015
MSS-3PRF (first and third pass)	**0.912**	1.062	1.141	1.139	1.081	1.043	1.029	1.043
Linear 3PRF	0.997	0.990	1.034	1.023	0.996	**0.989**	**0.992**	**0.988**
TPCA	1.010	0.998	0.999	1.001	1.005	1.009	1.003	1.005
PC-LARS	0.913	**0.977**	**0.980**	**0.991**	**0.995**	1.003	1.021	1.007
				Export				
MS-3PRF (first pass)	1.175	**0.984**	1.009	**0.963**	1.020	1.004	1.007	0.982
MS-3PRF (first and third pass)	1.003	1.031	1.108	1.082	1.019	1.084	1.003	**0.954****
MSS-3PRF (first pass)	1.192	0.985	0.998	0.965	1.015	1.008	1.003	0.989
MSS-3PRF (first and third pass)	**0.927****	1.072	1.094	1.037	1.036	1.057	1.064	1.045
Linear 3PRF	1.003	0.996	**0.991**	0.983	1.003	**0.983**	**0.996**	0.999
TPCA	1.061	1.029	1.021	1.006	1.008	1.002	0.999	0.998
PC-LARS	1.117	1.027	1.018	1.000	1.010	1.003	1.006	1.003

Note: This table reports the mean squared prediction error of a given approach relative to the MSPE of PCA for forecast horizons ranging from 1 quarter to 8 quarters ahead. Linear 3PRF uses a single target proxy, MS-3PRF (first pass) and MSS-3PRF (first pass) are regime-switching 3PRFs based on a single target proxy and regime-switching parameters in the first pass only; MS-3PRF (first and third pass) and MSS-3PRF (first and third pass) are regime-switching 3PRFs based on a single target proxy and regime-switching parameters in the first and third pass. For these approaches, the target proxy is the variable to forecast. TPCA is PCA where hard thresholding was performed before extracting the first principal component to forecast. PC-LARS is PCA where soft thresholding was performed before extracting the first principal component to forecast. Boldface indicates the best performing procedure for a specific horizon and variable. The first estimation sample extends from 1960Q3 to 1984Q4, and it is recursively expanded as we progress in the forecasting exercise. The full evaluation sample runs from 1985Q1 to 2015Q3. Statistical reductions in MSPE relative to PCA according to the Diebold and Mariano (1995) test are indicated by asterisks (* denotes significance at the 10% level and ** denotes significance at the 5% level).

Table 6: Out-of-Sample Macroeconomic Forecasting

Forecast horizon	1	2	3	4	5	6	7	8
				Import				
MS-3PRF (first pass)	1.129	0.981	0.938	0.994	0.921*	1.011	0.993	0.948
MS-3PRF (first and third pass)	**0.992**	**0.947**	**0.925**	0.987	0.928	0.967	1.009	0.981
MSS-3PRF (first pass)	1.187	0.983	0.935	0.997	**0.916***	1.007	1.026	0.953
MSS-3PRF (first and third pass)	1.035	0.961	0.963	1.051	0.973	1.026	1.062	1.027
Linear 3PRF	1.028	0.994	0.976	**0.978***	0.943*	**0.936****	0.962*	**0.940***
TPCA	1.012	0.995**	0.997	1.004	1.012	1.011	1.010	1.018
PC-LARS	1.058	0.991	1.004	1.000	0.977	0.977	**0.962****	0.972
				Hours				
MS-3PRF (first pass)	1.057	1.037	1.013	**0.995**	**0.992**	1.006	**0.972**	**0.982**
MS-3PRF (first and third pass)	1.037	1.253	1.565	1.500	1.417	1.159	1.327	1.217
MSS-3PRF (first pass)	1.088	1.032	1.017	0.995	0.997	1.011	0.984	0.990
MSS-3PRF (first and third pass)	1.069	1.001	1.017	1.109	1.319	1.401	1.515	1.491
Linear 3PRF	1.023	1.026	1.046	1.023	1.009	1.009	1.003	0.992
TPCA	1.010	1.001	1.003	1.005	1.009	1.003	1.005	1.008
PC-LARS	**0.864****	**0.980***	**0.993**	1.006	1.010	1.004	1.017	1.030
				GDP inflation				
MS-3PRF (first pass)	1.018	1.094	1.215	1.242	1.289	1.296	1.312	1.320
MS-3PRF (first and third pass)	1.003	1.099	1.725	1.611	1.528	1.393	1.407	1.459
MSS-3PRF (first pass)	1.018	1.096	1.181	1.228	1.263	1.276	1.288	1.315
MSS-3PRF (first and third pass)	**0.965**	**0.787**	**0.798**	**0.784**	**0.710***	**0.616****	**0.588****	**0.570****
Linear 3PRF	1.109	1.263	1.320	1.093	1.068	1.038	1.038	1.048
TPCA	1.010	1.026	1.037	1.056	1.050	1.059	1.031	1.044
PC-LARS	0.996	1.031	1.140	1.294	1.250	1.221	1.132	1.085
				PCE inflation				
MS-3PRF (first pass)	1.033	1.029	1.069	1.059	1.051	1.102	1.124	1.104
MS-3PRF (first and third pass)	1.004	1.065	1.547	1.867	2.076	2.353	2.469	2.346
MSS-3PRF (first pass)	1.020	1.020	1.056	1.050	1.028	1.085	1.093	1.050
MSS-3PRF (first and third pass)	**0.957**	**0.863**	**0.938**	**0.834***	**0.801****	**0.835***	**0.778****	**0.707****
Linear 3PRF	1.038	1.045	1.074	1.049	1.076	1.100	1.082	1.114
TPCA	1.031	1.023	1.025	1.023	1.029	1.028	1.049	1.043
PC-LARS	1.069	1.130	1.169	1.168	1.132	1.100	1.077	1.046

Note: This table reports the mean squared prediction error of a given approach relative to the MSPE of PCA for forecast horizons ranging from 1 quarter to 8 quarters ahead. Linear 3PRF uses a single target proxy, MS-3PRF (first pass) and MSS-3PRF (first pass) are regime-switching 3PRFs based on a single target proxy and regime-switching parameters in the first pass only; MS-3PRF (first and third pass) and MSS-3PRF (first and third pass) are regime-switching 3PRFs based on a single target proxy and regime-switching parameters in the first and third pass. For these approaches, the target proxy is the variable to forecast. TPCA is PCA where hard thresholding was performed before extracting the first principal component to forecast. PC-LARS is PCA where soft thresholding was performed before extracting the first principal component to forecast. Boldface indicates the best performing procedure for a specific horizon and variable. The first estimation sample extends from 1960Q3 to 1984Q4, and it is recursively expanded as we progress in the forecasting exercise. The full evaluation sample runs from 1985Q1 to 2015Q3. Statistical reductions in MSPE relative to PCA according to the Diebold and Mariano (1995) test are indicated by asterisks (* denotes significance at the 10% level and ** denotes significance at the 5% level).