Valentina Corradi; Norman Swanson

Working Paper

Testing for structural stability of factor augmented forecasting models

Working Paper, No. 2013-14

Provided in Cooperation with:
Department of Economics, Rutgers University

Suggested Citation: Valentina Corradi; Norman Swanson (2013) : Testing for structural stability of factor augmented forecasting models, Working Paper, No. 2013-14, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:
https://hdl.handle.net/10419/94228

Testing for Structural Stability of Factor Augmented Forecasting Models*

Valentina Corradi[1] and Norman R. Swanson[2]

[1]University of Warwick and [2]Rutgers University

April 2013

Keywords: diffusion index, factor loading stability, forecast failure, forecast stability, regression coefficients stability.
JEL classification: C12, C22, C53.

* Valentina Corradi, Department of Economics, University of Warwick, Coventry, CV4 7AL, UK, v.corradi@warwick.ac.uk. Norman Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. For useful discussions and comments on the paper, we would like to thank the editor, Xiaohong Chen, three anonymous referees, Jennifer Castle, Silvia Goncalves, David Hendry, Lajos Horwath, Atsushi Inoue, Brendan McCabe, Costas Milas, Barbara Rossi, Myung Hwan Seo, Moto Shintani, Andy Tremayne. Additionally, we would like to thank seminar participants at the Festschrift conference on Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions in Honor of Halbert White at the University of California, San Diego on May 6-7, 2011, the fifth (2011) CIREQ Conference on Time Series at the University of Montreal, the 2011 Netherlands Econometric Study Group at Erasmus University, the 2012 Econometric Society European Summer Meetings, the 2013 Conference on Change Point Processes at Cass Business School, the Bank of Canada, and the University of Liverpool. Excellent research assistance was provided by Hyun Hak Kim

Abstract

Mild factor loading instability, particularly if sufficiently independent across the different constituent variables, does not affect the estimation of the number of factors, nor subsequent estimation of the factors themselves (see e.g. Stock and Watson (2009)). This result does not hold in the presence of large common breaks in the factor loadings, however. In this case, information criteria overestimate the number of breaks. Additionally, estimated factors are no longer consistent estimators of "true" factors. Hence, various recent research papers in the diffusion index literature focus on testing the constancy of factor loadings. One reason why this is a positive development is that in applied work, factor augmented forecasting models are used widely for prediction, and it is important to understand when such models are stable. Now, forecast failure of factor augmented models can be due to either factor loading instability, regression coefficient instability, or both. To address this issue, we develop a test for the joint hypothesis of structural stability of both factor loadings and factor augmented forecasting model regression coefficients. The proposed statistic is based on the difference between full sample and rolling sample estimators of the sample covariance of the factors and the variable to be forecasted. Failure to reject the null ensures the structural stability of the factor augmented forecasting model. If the null is instead rejected, one can proceed to disentangle the cause of the rejection as being due to either (or both) of the afore mentioned varieties of instability. Standard inference can be carried out, as the suggested statistic has a chi-squared limiting distribution. We also establish the first order validity of (block) bootstrap critical values. Finally, we provide an empirical illustration by testing for the structural stability of factor augmented forecasting models for 11 U.S. macroeconomic indicators.

1 Introduction

The issue of forecast instability arising because of structural instability has received considerable attention in the forecasting literature, in recent years (see e.g. Clements and Hendry (2002), Hendry and Mizon (2005), and Castle, Doornik and Hendry (2010)). Among the main causes of instability, Hendry and Clements (2002) point out the importance of intercept shifts, mainly arising because of shifts in the means of omitted variables. Several ways to cope with forecast failure in regression models have been suggested (see e.g. Clements and Hendry, (2006), and the references cited therein). Moreover, among the different remedies proposed, there is some consensus that forecast pooling is one of the most effective, as discussed in Stock and Watson (2004), where empirical evidence supporting this view is provided. The intuition behind pooling is that, if the intercept shifts are sufficiently uncorrelated across different regressions, then by averaging forecasts we are also averaging out intercept shifts. Following this intuition, Stock and Watson (2009) argue that similar logic should also apply to diffusion index models. If factor loading coefficient instability is sufficiently independent across the different series, then the use of a large numbers of series in factor estimation can average out such instability. In this sense, estimated factors can be quite robust to time varying factor loadings. Indeed, Stock and Watson (2002) formally proved that estimated factors are consistent even in the presence of mild time variation in factor loading coefficients.

However, in the presence of substantial factor loading instability three problems arise. First, information criteria based procedures for detecting the number of breaks (see e.g. Bai and Ng (2002)) overestimate the number of breaks, if implemented using the full sample. This is because a factor model with breaks in the loadings is observationally equivalent to a model with a larger number of factors and stable loadings. Second, the variance of the estimated factors can be degenerate over subsamples. Third, estimated factors are no longer consistent for the "true" ones, in general. Hence, it is not surprising that testing for factor loading stability has received a great deal of attention in recent years. For example, Breitung and Eickmeier (2011) propose tests for the null hypothesis of a structural break in factor loading coefficients. Additionally, direct tests for con-

stancy of all factor loadings, which allow for some spatial correlation, have recently been suggested by Chen, Dolado and Gonzalo (2011) and by Han and Inoue (2012).

In this paper, we go one step further, and we jointly test for the structural stability of factor loadings and of regression coefficients estimated when constructing a factor augmented forecasting model. Heuristically speaking, stability of both factor loadings and regression coefficients implies structural stability of factor augmented forecasting models.

The forecast performance of diffusion index models in the presence of possible factor loading instability has been analyzed by Banerjee, Marcellino and Marsten (2009) via an extensive Monte Carlo. Stock and Watson (2009) disentangle instability into three different components, factor loading instability, factor dynamics instability, and factor model idiosyncratic component induced instability. They suggest using the full sample for factor estimation and using susbamples, or time-varying parameter techniques, for estimating regression coefficients in subsequent (forecasting) regression models. Castle, Clements and Hendry (2012) use *Autometrics* to select the most accurate factor (augmented) models. The use of recursive and rolling techniques for both factor estimation and factor augmented forecasting model estimation is analyzed in a series of prediction experiments by Kim and Swanson (2012).

To the best of our knowledge, there is no consistent test for the null hypothesis of factor augmented forecasting model structural stability. The aim of this paper is thus to fill this gap in the literature. We use the full sample for estimating the number of factors and for factor estimation. We then construct a statistic based on the difference between (i) the sample covariance of the target variable to be forecasted and the estimated factors using a full sample estimation scheme and (ii) the sample covariance of the target variable to be forecasted and the estimated factors using a rolling window estimation scheme. Under the null hypothesis of structural stability, the statistic has a normal limiting distribution. Under the alternative, the statistic diverges. In addition to discussing standard inference using a normal limit distribution result, we also discuss how to construct first order valid bootstrap critical values. In particular, we establish that under the null hypothesis, the actual statistic and a proposed bootstrap statistic

have the same limiting distribution, while under the alternative, the actual statistic diverges and the bootstrap statistic has a normal limiting distribution. It is worth noting that all of our asymptotic results assume only that $\sqrt{T}/N \to 0$, where T is the number of time series observations, and N is the number of variables used to construct factors. This is important because, while in financial applications N is generally larger than T, in macroeconomic applications we typically have $N < T$ (see e.g. Stock and Watson (2002a,b)). If we fail to reject the null hypothesis, then the factor augmented forecasting model is stable and can be reliably used, without worry about possible forecast failure issues (see e.g. Giacomini and Rossi (2009)). On the other hand, rejection of the null implies either instability of the factor loadings and/or instability of the forecasting model regression coefficients. It thus follows that if we subsequently run a test for factor loading stability and fail to reject the null, we then know that the rejection is due to regression coefficient instability. On the other hand, if we subsequently reject the null of factor loading stability, we can proceed to identify and estimate the break points, estimate the factors using the various subsamples, and implement a further test for the stability of the forecasting model coefficients.

In an empirical illustration we test for the structural stability of factor augmented forecasting models of 11 U.S. macroeconomic variables, including: the unemployment rate, personal income less transfer payments, the 10 year Treasury-bond yield, the consumer price index, the producer price index, non-farm payroll employment, housing starts, industrial production, M2, the S&P 500 index, and gross domestic product, using an extended version of the Stock and Watson macroeconomic dataset first examined in Kim and Swanson (2012). In this illustrative example, we implement a simple version of our test wherein critical values are constructed using the block bootstrap, and we find that structural instability is prevalent for approximately one half of the variables for which we construct factor augmented forecasting models. However, the dynamic structure of the forecasting models that we construct is quite simple, with explanatory variables consisting solely of diffusion indices; and hence our findings are meant only as a starting point to a thorough empirical investigation.

The rest of this paper is organized as follows. In Section 2, we define the set-up

and introduce the test for joint structural stability. Section 3 establishes the asymptotic properties of the suggested statistic. In Section 4, we establish the asymptotic first order validity of block bootstrap critical values. Finally, Section 5 reports the findings of our empirical illustration based on the use of a largescale macroeconomic dataset, and Section 6 concludes. All proofs are gathered in an Appendix.

2 Set-Up

We begin by outlining the factor model used in the sequel. Let

$$X_t = \mu_0 + \Lambda_{0,t} F_{0,t} + u_t, \tag{1}$$

where X_t is a $N \times 1$ vector, $\Lambda_{0,t}$ is a $N \times r$ factor loading matrix, μ_0 is a $N \times 1$ intercept vector, $F_{0,t}$ is the unobserved $r \times 1$ factor vector, and u_t is an error term.

Our objective is to predict a scalar target variable, y_{t+h}, where h denotes the forecast horizon. For the sake of simplicity, we develop our methodology in the context of predictive models based only on diffusion indices. Namely, consider the following forecasting model:

$$\begin{aligned} y_{t+h} &= \alpha_0 + \beta_{0,1,t} F_{0,1,t} + ... + \beta_{0,r,t} F_{0,r,t} + \epsilon_{t+h} \\ &= \alpha_0 + F_{0,t}' \beta_{0,t} + \epsilon_{t+h}. \end{aligned} \tag{2}$$

Needless to say, we can augment the model in (2) with both additional regressors and lagged factors. As such generalizations do not change any of our results, we focus our discussion on this simpler model. For a complete discussion of the usefulness of factor augmented models for forecasting, see Banerjee, Marcellino and Marsten (2010), Dufour and Stevanovic (2011), and the references cited therein.

We allow for two sources of potential structural instability. The first potential source stems from the structural relation between the covariates X_t and the factors $F_{0,t}$, and is captured by the loading factor matrix, $\Lambda_{0,t}$. The second source stems from the structural relation between the factors and the variable to be predicted, and it is captured by $\beta_{0,t}$. In principle, there may be two additional source of instability. One concerns potential

shifts in the intercept term in either (1) or (2). Another arises because one can also explicitly model factor dynamics (see e.g. Stock and Watson (2009)). For the sake of simplicity, in the sequel we just focus on loading and regression coefficient stability. In particular, we are concerned with testing the following hypotheses:

$$H_0 : \Lambda_{0,t} = \Lambda_0 \text{ and } \beta_{0,t} = \beta_0 \text{ for all } t,$$

versus

$$H_A : \Lambda_{0,t} = \begin{cases} \Lambda_{0,1} \text{ for } t/T \leq \tau_\lambda^{(1)} \\ \Lambda_{0,2} \text{ for } \tau_\lambda^{(1)} + 1 \leq t/T \leq \tau_\lambda^{(2)} \\ \vdots \\ \Lambda_{0,q_\lambda+1} \text{ for } \tau_\lambda^{(q_\lambda)} + 1 \leq t/T \leq 1 \end{cases}$$

and/or

$$\beta_{0,t} = \begin{cases} \beta_{0,1} \text{ for } t/T \leq \tau_\beta^{(1)} \\ \beta_{0,2} \text{ for } \tau_\beta^{(1)} + 1 \leq t/T \leq \tau_\beta^{(2)} \\ \vdots \\ \beta_{0,q_\beta} \text{ for } \tau_\beta^{(q_\beta)} + 1 \leq t/T \leq 1 \end{cases}$$

Note that both the case in which only a fraction αN, $0 < \alpha < 1$, of loadings have breaks, and the case in which the first r_1 factors have constant loadings, say, while the remaining $r - r_1$ factors have loadings which break, is covered by the alternative hypothesis above.

We proceed in three steps. First, we use the information criterion of Bai and Ng (2002), say the BN statistic, to determine the number of breaks, using the full sample. Second, we estimate the factors, again using the full sample. Third, we construct an estimator of the sample covariance between y_{t+h} and the estimated factors, using both the full sample and rolling windows of observations. The difference between the full sample and the rolling estimator of the covariance between y_{t+h} and the estimated factors is the key ingredient of our statistic.

If the "true" model has r factors, and if factor loadings are constant, then the information criterion used will detect r factors, with probability one. However, as already pointed out by Breitung and Eickmeier (2011), Chen, Dolado and Gonzalo (2011), and Han and Inoue (2012), in the presence of factor loading instability, the BN statistic will

overestimate the number of factors. Suppose that, in the definition of H_A, $q_\lambda = 1$, then the model in (1) can be written as:

$$X_t = \mu_0 + \Lambda_{0,1} F_t 1\{1 \leq t \leq \tau_\lambda^{(1)} T\} + \Lambda_{0,2} F_t 1\{\tau_\lambda^{(1)} T < t \leq T\} + u_t,$$

and the BN statistic applied over the full sample will detect $2r$ factors. We then estimate the $2r$ factors as:

$$\begin{aligned} &\left(\widehat{G}_{t,T,N}, \widehat{\Lambda}_{T,N}\right) \\ &= \arg\min_{\Lambda,G} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\left(X_{i,t} - \frac{1}{T} \sum_{t=1}^{T} X_t \right) - \lambda_i' G_t \right)^2. \end{aligned}$$

More generally, $\widehat{G}_{N,T} = \widehat{F}_{N,T}$ is a $T \times r$ matrix if $\Lambda_{0,t} = \Lambda_0$ for all t, otherwise $\widehat{G}_{N,T} = \begin{pmatrix} \widehat{F}_{1,N,T} & \widehat{F}_{2,N,T} & \cdots & \widehat{F}_{q_\lambda+1,N,T} \end{pmatrix}$, which is a $T \times (q_\lambda + 1) \times r$ matrix. Now, under factor loading stability:

$$\widehat{G}_{N,T} - F_{0,T} H_F' = o_p(1),$$

while if there are q_λ breaks in the factor loadings, then

$$\begin{aligned} &\widehat{G}_{N,T} - G_T H_G' \\ = \;& o_p(1), \end{aligned}$$

where H_F and H_G are defined in the Appendix, and

$$G_T = \begin{pmatrix} F_{0,T_1}^{(1)} & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & F_{0,T_{q_\lambda}}^{(q_\lambda+1)} \end{pmatrix}, \tag{3}$$

with $F_{0,T_1}^{(1)} = \left(F_{0,1}, ..., F_{0,T\tau_\lambda^{(1)}} \right)'$ a $T\tau_\lambda^{(1)} \times r$ matrix and $F_{0,T_{q_\lambda}}^{(q_\lambda+1)} = \left(F_{0,T\tau_\lambda^{(q_\lambda)}+1}, ..., F_{0,T} \right)'$ a $(T - T\tau_\lambda^{(q_\lambda)}) \times r$ matrix.

We now construct a full sample estimator of the covariance between the factors and

the variable to be predicted. First, using the full sample, construct

$$\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t.$$

Then, additionally construct a rolling estimator based on windows of R observations, as follows:

$$\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j,\ t\geq R.$$

The statistic is:

$$Z_{P,R}=\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t-\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right)\right), \tag{4}$$

where $P+R=T$, and $P/R\to\pi$, with $0<\pi<\infty$. Note that for simplicity, in the above expression (and in the sequel) we assume that $h=1$.

Note that in the above statistic, it is not our objective to examine the difference between a full sample and a rolling estimator of the β_s. The reason is that we wish to avoid degeneracy associated with breaks that occur after the first R observations. For example, suppose that $q_\lambda=1$ and $T\tau_\lambda^{(1)}=\alpha R$, with $\alpha>1$. In this case, for all $R\leq t\leq\alpha R$,

$$\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_j\widehat{G}_j'\xrightarrow{p}H_G\begin{pmatrix}\Sigma_{01}&0\\0&0\end{pmatrix}H_G',$$

where Σ_{01} is $r\times r$, and 0 denotes a $r\times r$ block of zeros. Then, $\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}\widehat{G}_{j-1}'$ is singular and cannot be inverted. For the same reason, we consider only a full sample estimator of the factors. The reason is that in the above case, we would not be able to identify two factors using rolling windows, $[t-R+1,t]$ for $R\leq t\leq\alpha R$.

In summary, our suggested statistic is more in the spirit of a Hausman type test, than in the spirit of tests for structural stability. Indeed, we are comparing two estimators which converge to the same probability limit under the null, but to different probability limits under the alternative. Furthermore, as shown in the Appendix, under the null, the full sample estimator is the more efficient estimator.

Han and Inoue (2012) test the null of constancy of factor loadings using a sup-type

statistic of the following sort:

$$\sup_{\tau_\lambda \in (0,1)} \sqrt{T} \left(\frac{1}{T\tau_\lambda} \sum_{t=1}^{T\tau_\lambda} \widehat{G}_{t-1} \widehat{G}'_{t-1} - \frac{1}{T(1-\tau_\lambda)} \sum_{t=T\tau_\lambda+1}^{T(1-\tau_\lambda)} \widehat{G}_{t-1} \widehat{G}'_{t-1} \right).$$

Their statistic has optimality properties versus the alternative of one common break. However, in our case, we cannot separately identify the contribution of breaks in factor loadings from breaks in the regression coefficients.[1]. Finally, it is worth noting that Chen, Dolado and Gonzalo (2011) propose a Lagrange Multiplier test for constancy of loadings, in which they regress the first factors over the remaining factors. The logic underlying their test is that under the null of structural stability, the estimated factors are consistent for the "true" ones, and so are orthogonal to each other.

3 Asymptotics

Below we state the assumptions that are needed in order to establish the asymptotic properties $Z_{P,R}$. As the number of estimated factors differs under H_0 and under H_A, for notational simplicity, we separately state the assumptions needed in order to establish the limiting distribution under the null, and the assumptions needed under the alternative. Hereafter, for a matrix B, $\|B\| = (tr(B'B))^{1/2}$, and C denotes a generic constant.

Under the null hypothesis, we require the following assumptions.

Assumption AN:

AN1: **(i)** For $i = 1, ..., N$, $(F_{0,t}, u_{it})$ is $\alpha-$mixing with size $-4(4+\psi)/\psi$, $\psi > 0$. **(ii)** For $i = 1, ..., N$ and $j = 1, ..., r$, $\sup_t \mathrm{E}\left(|F_{0,jt}|^{2k}\right) \leq C$, $\sup_t \mathrm{E}\left(|u_{it}|^{2k}\right) \leq C$, with $k > 2(2+\psi)$, $\inf_t \det\left(\mathrm{E}\left(F_{0,t}F'_{0,t}\right)\right) > 0$. **(iii)** For $i = 1, ..., N$, $j = 1, ..., r$, $\sup_{i,j} |\lambda_{0,ij}| \leq C$. **(iv)** $\mathrm{E}(u_{it}) = 0$, $\mathrm{E}(F_{0,t}u_{it}) = 0$.

AN2: Let $\sigma_{ij,ts} = \mathrm{E}(u_{it}u_{js})$, $\sup_{t,s} |\sigma_{ij,ts}| = \delta_{ij}$, $\sup_{i,j} |\sigma_{ij,ts}| = \delta_{ts}$, and $\gamma_{st} = \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N} u_{is}u_{it}\right)$.

[1] In principle, we could construct a statistic based on:

$$\sup_{\tau \in (0,1)} \sqrt{T} \left(\frac{1}{T\tau} \sum_{t=1}^{T\tau} \widehat{G}_t y_t - \frac{1}{T(1-\tau)} \sum_{t=T\tau+1}^{T(1-\tau)} \widehat{G}_t y_t \right).$$

However, in general this statistic does not have optimality properties versus either breaks in the loadings or breaks in the coefficients, or both.

(i) $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\delta_{ij}\leq C$, $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\delta_{ts}\leq C$. **(ii)** $\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}|\gamma_{st}|\leq C$. **(iii)** $\sup_{t,s}\mathrm{E}\left(N^{-1/2}\sum_{i=1}^{N}|u_{it}u_{is}-\mathrm{E}(u_{it}u_{is})|^{4}\right)\leq C$. **(iv)** For all t, $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{0,i}u_{it}$ satisfy a central limit theorem. **(v)** $\Lambda_0'\Lambda_0/N-\Sigma_{\Lambda}^{\dagger}=o(1)$, with $\Sigma_{\Lambda}^{\dagger}$ of rank r.

AN2': (i) and **(iii)-(v)** as in AN2, **(ii)** $\frac{1}{R^2}\sum_{j=t-R+1}^{t}\sum_{k=1}^{T}|\gamma_{kj}|=O\left(\frac{1}{R}\right)$ uniformly in t.

AN3: (i) For each t, $\mathrm{E}\left\|\frac{1}{\sqrt{NT}}\sum_{s=1}^{T}\sum_{i=1}^{N}F_{0,s}\left(u_{it}u_{is}-\mathrm{E}\left(u_{it}u_{is}\right)\right)\right\|^{2}\leq C$.

(ii) $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}F_{0,t}\Lambda_0 u_t\right\|^{2}\leq C$, where $u_t=(u_{1t},...,u_{Nt})$.

(iii) $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_{0,i}u_{it}\right\|\right)\leq C$.

AN4: (i) y_t is $\alpha-$mixing, with size $-4\left(4+\psi\right)/\psi$, $\psi>0$. **(ii)** $\sup_t\mathrm{E}\left(|y_t|^{2k}\right)\leq C$, with $k>2\left(2+\psi\right)$ and $\mathrm{E}(F_{0,t}\epsilon_t)=0$. **(iii)** $\mathrm{E}\left(\epsilon_t^{2k}\right)\leq C$ and for $i=1,...,N$, $\mathrm{E}(\epsilon_t u_{it})=0$.

AN5: (i) For all t, $\mathrm{E}\left(\left|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\epsilon_{s+1}\left(u_{it}u_{is}-\mathrm{E}\left(u_{it}u_{is}\right)\right)\right|^{2}\right)\leq C$.

(ii) $\mathrm{E}\left(\left|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\lambda_{0,i}\epsilon_{s+1}u_{it}\right|^{2}\right)\leq C$.

AN5': (i) For all t, $\mathrm{E}\left(\sup_{s>R}\left|\frac{\sqrt{P}}{R\sqrt{N}}\sum_{j=s-R}^{s}\sum_{i=1}^{N}\epsilon_{j+1}\left(u_{it}u_{ij}-\mathrm{E}\left(u_{it}u_{ij}\right)\right)\right|^{2}\right)\leq C$.

(ii) For all t, $\mathrm{E}\left(\sup_{s>R}\left\|\frac{\sqrt{P}}{R\sqrt{N}}\sum_{j=s-R}^{s}\sum_{i=1}^{N}\lambda_{0,i}\epsilon_{j+1}u_{it}\right\|^{2}\right)\leq C$.

Assumptions AN1-AN5 are standard assumption on loadings and factors (AN1), spatial and time dependence of u_{it} (AN2), moment and weak dependence among factors, loading, and idiosyncratic errors (AN3), moment and dependence of the variable to be forecasted and the regression error (AN4), and weak dependence between idiosyncratic and regression errors (AN5). AN1-AN5 concern full sample estimation of factors and regression coefficients, and largely coincide with Assumptions A1-A5 in Goncalves and Perron (2013). AN2' and AN5' state additional conditions on spatial and time dependence of idiosyncratic and regression errors, and on weak dependence among idiosyncratic and regression errors for the case of rolling estimation.

Turning now to the data generating process under the alternative, for notational simplicity, suppose that we have q_λ common breaks in the factor loadings, and q_β common

breaks in the regression coefficients, where $\max\{q_\lambda, q_\beta\} \geq 1$. Hereafter, let:

$$G_t = \begin{cases} \left(\underbrace{F_t}_{r\times 1}, \underbrace{0,...,0}_{q_\lambda r\times 1}\right) 1\left\{1 < t \leq \tau_\lambda^{(1)} T\right\} \\ \vdots \\ \left(\underbrace{0,...,0}_{q_\lambda r\times 1}, \underbrace{F_t}_{r\times 1}\right) 1\left\{\left(\tau_\lambda^{(q_\lambda)} T + 1\right) \leq t \leq T\right\} \end{cases},$$

so that G_t is an $(q_\lambda + 1)r \times 1$ matrix. If $q_\lambda = \min\{q_\lambda, q_\beta\} = 0$, then $\tau_\lambda^{(1)} = 1$ and $G_t = F_{0,t}$, which is an $r \times 1$ vector. We now state the assumptions required under the alternative hypothesis.

Assumption AA:

AA1: **(i)** For $i = 1, ..., N$, (G_t, u_{it}) is $\alpha-$mixing with size $-4(4+\psi)/\psi$, $\psi > 0$. **(ii)** For $i = 1, ..., N$ and $j = 1, ..., (q_\lambda + 1)r$, $\sup_t \mathrm{E}\left(|G_t|^{2k}\right) \leq C$, $\sup_t \mathrm{E}\left(|u_{it}|^{2k}\right) \leq C$, with $k > 2(2+\psi)$, and following the notation in (3), $\frac{F_{0,T_1}^{(1)} F_{0,T_1}^{(1)\prime}}{T} \xrightarrow{p} \tau_\lambda^{(1)} \Sigma_{01}$, $\frac{F_{0,T_{q_\lambda}}^{(q_\lambda+1)} F_{0,T_{q_\lambda}}^{(q_\lambda+1)\prime}}{T} \xrightarrow{p} (1 - \tau_\lambda^{(q_\lambda)})\Sigma_{0 q_\lambda+1}$. **(iii)** For $i = 1, ..., N$, $j = 1, ..., (q_\lambda + 1)r$, $\sup_{i,j} |\lambda_{0,ij}| \leq C$. **(iv)** $\mathrm{E}(u_{it}) = 0$, $\mathrm{E}(G_t u_{it}) = 0$.

AA2: Let $\sigma_{ij,ts} = \mathrm{E}(u_{it} u_{js})$, $\sup_{t,s} |\sigma_{ij,ts}| = \delta_{ij}$, $\sup_{i,j} |\sigma_{ij,ts}| = \delta_{ts}$, and $\gamma_{st} = \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^N u_{is} u_{it}\right)$.
(i) $\frac{1}{N}\sum_{i=1}^N \sum_{j=1}^N \delta_{ij} \leq C$ and $\frac{1}{T}\sum_{t=1}^T \sum_{s=1}^T \delta_{ts} \leq C$. **(ii)** $\frac{1}{T}\sum_{s=1}^T \sum_{t=1}^T |\gamma_{st}| \leq C$.
(iii) $\sup_{t,s} \mathrm{E}\left(N^{-1/2}\sum_{i=1}^N |u_{it}u_{is} - \mathrm{E}(u_{it}u_{is})|^4\right) \leq C$. **(iv)** For all[2] $\tau_\lambda^{(h)} + 1 \leq t/T \leq \tau_\lambda^{(h+1)}$, $0 \leq h \leq q_\lambda$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_{0,h,i} u_{it}$ satisfy a central limit theorem. **(v)** $\Lambda_{h,0}'\Lambda_{h,0}/N - \Sigma_{\Lambda_h}^\dagger = o(1)$, with $\Sigma_{\Lambda_h}^\dagger$ of rank r.

AA2': **(i)** and **(iii)-(v)** as in AA2. **(ii)** $\frac{1}{R^2}\sum_{j=t-R+1}^t \sum_{k=1}^T |\gamma_{kj}| = O\left(\frac{1}{R}\right)$ uniformly in t.

AA3: **(i)** For each t, $\mathrm{E}\left\|\frac{1}{\sqrt{NT}}\sum_{s=1}^T \sum_{i=1}^N G_s(u_{it}u_{is} - \mathrm{E}(u_{it}u_{is}))\right\|^2 \leq C$.
(ii) For $0 \leq h \leq q_\lambda$, $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^T G_t \Lambda_{0,h} u_t 1\left\{\tau_\lambda^{(h)} + 1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\}\right\|^2 \leq C$, where $u_t = (u_{1t}, ..., u_{Nt})$.
(iii) For $0 \leq h \leq q_\lambda$, $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_{0,i} u_{it} 1\left\{\tau_\lambda^{(h)} + 1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\}\right\|\right) \leq C$.

[2] With an abuse of notation, we set $\tau_\lambda^{(q_\lambda+1)} = 1$.

AA4: **(i)** y_t is $\alpha-$mixing, with size $-4(4+\psi)/\psi$, $\psi > 0$. **(ii)** For[3] $0 \leq h \leq q_\lambda$, $0 \leq h' \leq q_\beta$, $\sup_t \mathrm{E}\left(\left|y_t 1\left\{\tau_\beta^{(h')}+1 \leq t/T \leq \tau_\beta^{(h'+1)}\right\}\right|^{2k}\right) \leq C$, with $k > 2(2+\psi)$ and $\mathrm{E}(G_t\epsilon_t) = 0$. **(iii)** $\mathrm{E}\left(\left(\epsilon_t 1\left\{\tau_\beta^{(h')}+1 \leq t/T \leq \tau_\beta^{(h'+1)}\right\}\right)^{2k}\right) \leq C$, and, for $i = 1, ..., N$, $\mathrm{E}\Big(\epsilon_t u_{it} 1\left\{\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\} 1\left\{\tau_\beta^{(h')}+1 \leq t/T \leq \tau_\beta^{(h'+1)}\right\}\Big) = 0$.

AA5: **(i)** For all t, $\mathrm{E}\Big(\Big|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\epsilon_{s+1}\left(u_{it}u_{is}-\mathrm{E}\left(u_{it}u_{is}\right)\right)$
$1\left\{\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\} 1\left\{\tau_\beta^{(h')}+1 \leq s/T \leq \tau_\beta^{(h'+1)}\right\}\Big|^2\Big) \leq C$.
(ii) For all t,
$\mathrm{E}\left(\left|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\lambda_{0,h,i}\epsilon_{s+1}u_{it}\right|^2 1\left\{\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\} 1\left\{\tau_\beta^{(h')}+1 \leq s/T \leq \tau_\beta^{(h'+1)}\right\}\right) \leq C$.

AA5': **(i)** For all t,
$\mathrm{E}\Big(\sup_{s>R}\Big|\frac{\sqrt{P}}{R\sqrt{N}}\sum_{j=s-R}^{s}\sum_{i=1}^{N}\epsilon_{j+1}\left(u_{it}u_{ij}-\mathrm{E}\left(u_{it}u_{ij}\right)\right)$
$1\left\{\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\} 1\left\{\tau_\beta^{(h')}+1 \leq s/T \leq \tau_\beta^{(h'+1)}\right\}\Big|^2\Big) \leq C$.
(ii) For all t, $\mathrm{E}\Big(\sup_{s>R}\Big\|\frac{\sqrt{P}}{R\sqrt{N}}\sum_{j=s-R}^{s}\sum_{i=1}^{N}\lambda_{0,h,i}\epsilon_{j+1}u_{it}$
$1\left\{\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}\right\} 1\left\{\tau_\beta^{(h')}+1 \leq s/T \leq \tau_\beta^{(h'+1)}\right\}\Big\|^2\Big) \leq C$.

Assumption AA1 imposes restrictions on the pseudo true factors, G_t, and on the loadings Λ_{0h}, $0 \leq h \leq q_\lambda$. Assumption AA2 perfectly mirrors AN2, while AA3 controls moment and weak dependence among factors, loadings, and idiosyncratic errors, over each subsample, $\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}$. Assumption AA4 imposes conditions on the moments of the variable to be forecasted and on the regression error over each subsample, $\tau_\beta^{(h')}+1 \leq t/T \leq \tau_\beta^{(h'+1)}$, $0 \leq h' \leq q_\beta$. AA5 controls the dependence between idiosyncratic and regression errors, over each subsample, $\tau_\lambda^{(h)}+1 \leq t/T \leq \tau_\lambda^{(h+1)}$, $0 \leq h \leq q_\lambda$, and $\tau_\beta^{(h')}+1 \leq s/T \leq \tau_\beta^{(h'+1)}$, $0 \leq h' \leq q_\beta$. Finally, AA2' and AA5' are related assumptions for the case of rolling estimation.

Theorem 1: **(i)** *Let AN1-AN5, AN2', and AN5' hold, and let* $\sqrt{T}/N \to 0$, and $P/R \to \pi > 0$. *Then, under* H_0*:*

$$Z_{P,R} \xrightarrow{d} N\left(0, \Omega_0\right),$$

[3] With an abuse of notation, we set $\tau_\beta^{(q_\beta+1)} = 1$.

where

$$\Omega_0 = 1\{\pi \leq 1\}\left(\pi - \frac{\pi^2}{3}\right)V_{\beta_0} + 1\{\pi > 1\}\left(1 - \frac{1}{3\pi}\right)V_{\beta_0} - \frac{\pi}{1+\pi}V_{\beta_0}$$

and

$$V_{\beta_0} = H_F \sum_{j=-\infty}^{\infty} \mathrm{E}\left(F_{0,1}F'_{0,1+j}\epsilon_1\epsilon_{1+j}\right)H'_F,$$

with

$$H_F = \mathrm{p}\lim_{\mathrm{N,T}\to\infty} V_{N,T}^{-1}\frac{\widehat{F}'F_0}{T}\frac{\Lambda'\Lambda}{N},$$

with $V_{N,T}$ an $r \times r$ diagonal matrix with the largest r eigenvalues of XX'/NT in decreasing order.

(ii) Let AA1-AA5, AA2', and AA5' hold, and let $\sqrt{T}/N \to 0$, and $P/R \to \pi > 0$. *Then, under H_A there exists $\varepsilon > 0$, such that:*

$$\lim_{P\to\infty} \Pr\left(P^{-1/2}\|Z_{P,R}\| \geq \varepsilon\right) = 1.$$

Corollary 1: **(i)** *Let AN1-AN5, AN2', and AN5' hold, and let* $\sqrt{T}/N \to 0$, and $P/R \to \pi > 0$. *Then, under H_0:*

$$Z'_{P,R}\widehat{\Omega}_T^{-1}Z'_{P,R} \xrightarrow{d} \chi^2_r,$$

where

$$\widehat{\Omega}_T = 1\{\pi \leq 1\}\left(\pi - \frac{\pi^2}{3}\right)\widehat{V}_{T,\beta_0} + 1\{\pi > 1\}\left(1 - \frac{1}{3\pi}\right)\widehat{V}_{T,\beta_0} - \frac{\pi}{1+\pi}\widehat{V}_{T,\beta_0}$$

and

$$\widehat{V}_{T,\beta_0} = \frac{1}{T}\sum_{t=l_T}^{T-l_T}\sum_{j=-l_T}^{l_T}\omega_j\widehat{G}_{t-1}\widehat{G}'_{t-1-j}\widehat{\epsilon}_t\widehat{\epsilon}_{t-j}, \tag{5}$$

with $\omega_j = 1 - \frac{j}{l_T - 1}$, $\widehat{\epsilon}_t = y_t - \widehat{G}'_{t-1}\widehat{\beta}_T$ and $\widehat{\beta}_T = \left(\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}\widehat{G}'_{t-1}\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t$.

(ii) Let AA1-AA5, AA2', and AA5' hold, and let $\sqrt{T}/N \to 0$, and $P/R \to \pi > 0$.

Then, under H_A there exists $\varepsilon > 0$, such that:

$$\lim_{P\to\infty} \Pr\left(P^{-1}\left(Z'_{P,R}\widehat{\Omega}_T^{-1}Z'_{P,R}\right) \geq \varepsilon\right) = 1.$$

Note that $\widehat{V}_{T,\beta_0}$ has to be constructed using residuals from a full sample estimator of the regression coefficients. This ensures that under the alternative, $\widehat{V}_{T,\beta_0}$, though no longer consistent for V_{β_0}, is nevertheless bounded in probability.

Note also that the statistic $Z_{P,R}$ is constructed without recentering. Indeed, there is no need to recenter, provided that α_0 in (2) is constant; in fact the contribution of a non-zero intercept cancels out, both in finite and large samples. On the other hand, if α_0 were time varying, then the statistic would diverge even if loadings and β_s were stable. To allow for possible shifts in the intercept in (2) one would have to recenter y_{t+1} around the full sample and rolling sample estimators of the mean.
If instead the intercept in (1) were time varying, then the information criterion would still detect the correct number of factors. However, the estimated factors would not be consistent for the "true" ones, and in general inference based on $Z_{P,R}$ would lead to rejection of the null.

Suppose that instead of forecasting y_{t+1} using only the estimated factors, we also have additional variables, say W_t, as well as lagged factors. In this case, we can still use the same statistic as in (4), simply replacing $\widehat{G}_t$ with $\widehat{Z}_t = \left(\widehat{G}_t, \widehat{G}_{t-1}, ..., \widehat{G}_{t-p}, W_t\right)$. The null is now that of joint stability of factors loadings and of all regression coefficients - including coefficients on the factors, on their lags, and on additional variables.

In finite samples, the statistic $Z_{P,R}$ depends on the choice of how to split T into R and P (i.e. on the choice of π). Heuristically, the smaller is π, the better is the finite sample size and the worse is the finite sample power, while larger π values lead to the reverse trade-off. Rossi and Inoue (2012) suggest tests for out of sample predictive accuracy that are robust to the choice of π. In particular, they suggest viewing π as a nuisance parameter, which can be integrated out by constructing appropriate summary statistics over a range of π values. Along the same lines, Hansen and Timmermann (2012) suggest minimizing related statistics over, say, $\underline{\pi} \leq \pi \leq \overline{\pi}$. The same approach can be followed

in our set-up. Nevertheless, as shown in the proof of part (ii) of Theorem 1, what really matters is the relative location of breaks in loadings with respect to breaks in regression coefficients, and whether these break points are close or far from π.

As discussed in the introduction, the limiting distribution in Theorem 1 is obtained under the Assumption that $\sqrt{T}/N \to 0$. In fact, as shown in Bai and Ng (2006), if $\sqrt{T}/N \to 0$, one can ignore factor estimation error. The case of $\sqrt{T}/N \to c > 0$ is analyzed in Goncalves and Perron (2013) who have analytically derived the closed form of the -bias term due to factor estimation error.

Finally, it should be noted that if the model in (1) and (2) is misspecified, then provided loadings and regression coefficients are stable, we will not reject H_0. This is because, while factor and regression coefficient estimators are no longer consistent for the "true" ones, the full sample and rolling sample estimators of the covariance between estimated factors and target variable still converge to the same probability limit. The only exception would be the case in which misspecification is due to the omission of a variable with a shift in the mean, say, as in this case the full sample and rolling estimators no longer converge to the same object.

4 Bootstrap Critical Values

Inference based on asymptotic chi-squared critical values should be expected to be relatively unreliable in medium and small samples, particularly in cases where the estimator $\widehat{V}_{T,\beta_0}$ is not well behaved. Thus, it may be preferable to rely on simple to construct bootstrap critical values.

Let $W_t = \left(y_t,\ \widehat{G}_{t-1}\right)$, $t = 2, ..., T$, and resample b blocks of length l, $bl = T$, with $l = o(\sqrt{T})$. This yields a bootstrap sample, say, $W_t^* = \left(y_t^*, \widehat{G}_{t-1}^*\right)$, with $W_i^*, ..., W_T^* = W_{I_{1+1}}, ..., W_{I_1+l}, W_{I_2+1}, ..., W_{I_b+1}, ..., W_{I_b+l}$, where for $i = 0, ..., T-l$, I_i is a discrete iid uniform random variable on the interval $[0, T-l-1]$. Note that resampling is done only once, from the full sample. If we were to instead resample over rolling windows, we would be imposing artificial independence in the sequence of bootstrap rolling estimators. However, note that in the rolling estimation case, observations in the middle of the

sample are used more frequently than observations at either the beginning or the end of the sample. This introduces a location bias in any bootstrap statistic, as any block from the original sample has the same probability of being selected. As explained in detail in Corradi and Swanson (2006), the location bias problem is solved by simply recentering the bootstrap statistic around the full sample. In the present context, this results in simply recentering both $\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}^{*}y_j^{*}$ and $\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}^{*}y_t^{*}$ around $\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t$, and these two recentering terms clearly cancel out. Hence, the bootstrap statistic is simply:

$$\left\|Z_{P,R}^{*}\right\| = \left\|\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}^{*}y_j^{*}\right)-\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}^{*}y_t^{*}\right)\right\|.$$

Note that we have directly resampled the factors estimated using the full sample. This is because, for $\sqrt{T}/N\to 0$, factor estimation error is asymptotically negligible, and thus we do not need to bootstrap estimated factors. If instead, $\sqrt{T}/N\to c>0$, then factor estimation error matters, and results in asymptotic bias. Goncalves and Perron (2013) establish sufficient conditions for the validity of a wild bootstrap procedure in the case where $c>0$.

Theorem 2: **(i)** *Let AN1-AN5, AN2', and AN5' hold, and as* $T,N,P,R\to\infty$, *let* $\sqrt{T}/N\to 0$, $P/R\to\pi>0$, $l,b\to\infty$ *and* $l/\sqrt{T}\to 0$. *Then, under* H_0*:*

$$P\left(\omega:\sup_{x\in\mathbb{R}^{+}}\left|P_{P,R}^{*}\left(\left\|Z_{P,R}^{*}\right\|\leq x\right)-P\left(\|Z_{P,R}\|\leq x\right)\right|>\varepsilon\right)\to 0.$$

(ii) *Let AA1-AA5, AA2', and AA5' hold, and as* $T,P,R\to\infty$, *let* $P/R\to\pi>0$, $l,b\to\infty$ *and* $l/\sqrt{T}\to 0$. *Then, under* H_A*:*

$$P\left(\omega:\left(\|Z_{P,R}\|-\left\|Z_{P,R}^{*}\right\|\right)>\varepsilon\right)\to 1.$$

The above result suggests proceeding in the following manner. Carry out B bootstrap replications (B large), and compute the percentiles of the empirical distribution of the B bootstrap statistics. Then, (do not) reject at level $\alpha\%$, if $\|Z_{P,R}\|$ is (smaller than or equal

to) larger than the $(1-\alpha)th-$percentile of the empirical bootstrap statistic distribution. As alluded to above, the main advantage of comparing $\|Z_{P,R}\|$ with $\left\|Z^*_{P,R}\right\|$ is that we do not have to estimate the covariance matrix $\widehat{V}_{T,\beta_0}$ in (5). However, it should be stressed that our test is non-pivotal and hence bootstrap critical values, although first order valid, cannot provide refinements. An alternative, proceeding as in Goncalves and White (2004), is to compute the bootstrap counterpart of $Z'_{P,R}\widehat{\Omega}_T^{-1}Z'_{P,R}$, namely $Z^{*\prime}_{P,R}\widehat{\Omega}_T^{*-1}Z^{*\prime}_{P,R}$, where

$$\widehat{\Omega}^*_T = 1\{\pi \leq 1\}\left(\pi - \frac{\pi^2}{3}\right)\widehat{V}^*_{T,\beta_0} + 1\{\pi > 1\}\left(1 - \frac{1}{3\pi}\right)\widehat{V}^*_{T,\beta_0} - \frac{\pi}{1+\pi}\widehat{V}^*_{T,\beta_0}$$

$$\widehat{V}^*_{T,\beta_0} = \frac{1}{b}\sum_{i=1}^{b}\left(\left(\frac{1}{\sqrt{l}}\sum_{j=1}^{l}\left(\widehat{G}_{I_i+j}\left(y_{I_i+j} - \widehat{G}_{I_i+j}\widehat{\beta}^*_T\right)\right)\right)\left(\frac{1}{\sqrt{l}}\sum_{j=1}^{l}\left(\widehat{G}_{I_i+j}\left(y_{I_i+j} - \widehat{G}_{I_i+j}\widehat{\beta}^*_T\right)\right)\right)'\right),$$

and $\widehat{\epsilon}^*_j = y^*_j - \widehat{G}^{*\prime}_{j-1}\widehat{\beta}^*_T$, $\widehat{\beta}^*_T = \left(\frac{1}{T}\sum_{t=1}^{T}\widehat{G}^*_{t-1}\widehat{G}^{*\prime}_{t-1}\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\widehat{G}^*_{t-1}y^*_t$. Needless to say, because of $\widehat{V}^*_{T,\beta_0}$, the computation of the empirical distribution of $Z^{*\prime}_{P,R}\widehat{\Omega}_T^{*-1}Z^{*\prime}_{P,R}$ is more involved than that of $\left\|Z^*_{P,R}\right\|$. Also, because of factor estimation error, it is not easy to assess whether and under which additional conditions inference based on $Z^{*\prime}_{P,R}\widehat{\Omega}_T^{*-1}Z^{*\prime}_{P,R}$ may provide refinements, when compared with the use of asymptotic critical values.

5 Empirical Illustration

We illustrate the implementation of the proposed test statistic, $\|Z_{P,R}\|$ by carrying out experiments where the objective is to construct predictions of the same 11 macroeconomic variables examined in Armah and Swanson (2010), as summarized in Table 1. Prediction models are assumed to be constructed according to the generic specification given in equations (1) and (2). As discussed above, the number of factors is estimated based on the full sample of data, using the approach of Bai and Ng (2002). Factors are constructed using the macroeconomic dataset first introduced of Stock and Watson (2002a,b), and extended by and Kim and Swanson (2012,2013). In total, we have 155 monthly variables for the period 1962:10 - 2009:5, so that $N = 155$ and $T = 560$. We construct bootstrap critical values. In implementation of the bootstrap, block lengths

of $l = \{2, 5, 10\}$ were tried, and $B = 300$. Additionally, a variety of different values of R (and hence P, given that $T = R + P$) were tried, including $R = \{200, 240, 280, 320, 360, 400\}$, corresponding to values of P/R including $\{9/5, 4/3, 1, 3/4, 5/9, 2/5\}$.

A selected subset of our empirical findings are collected in Tables 2-4, based on predictions constructed for h =1 (Table 2), h =3 (Table 3), and h =12 (Table 4). Related results for other parameter permutations were qualitatively similar, and are available upon request from the authors. In the tables, entries are given for: (i) the test statistic; (ii) the 95th, 90th, and 50th percentiles of the empirical bootstrap distribution; and (iii) the p-value based on the empirical bootstrap distribution.

A number of conclusions emerge from examination of the results in the tables. First, results are largely the same, regardless of the value of l (with Panel A corresponding to $l = 2$, Panel B corresponding to $l = 5$, and Panel C corresponding to $l = 10$), and regardless of the forecast horizon (compare results in Tables 2,3, and 4). Across all of these permutations, and for all values of R, the null of stability is rejected for personal income less transfer payments, the consumer price index, non-farm payroll employment, industrial production, M2, and gross domestic product.[4] Failures to reject the null of stability occur in a similarly consistent manner across all experimental permutations, for the unemployment rate, the 10 year Treasury-bond yield, the producer price index, and housing starts. The only variable that consistently fails to reject the null for small values of R and rejects for large values of R is the S&P 500 index. In conclusion, the test appears to be performing as desired, in the sense that it is robust to values of R, h, and l.

6 Concluding Remarks

We have developed a simple to implement test for the structural stability of factor augmented forecasting models. Our null hypothesis involves jointly testing stability of factor loading and forecast model coefficients. We estimate the number of factors and the factors themselves using the full sample. Then, we construct a statistic based on

[4] More precisely, it should be noted that there are two or three cases (across all 9 block length - forecast horizon permutations) where NPE and M2 fail to reject the null for the smallest values of R.

the difference between a full sample and a rolling estimator of the sample covariance of estimated factors and variable to be predicted. The statistic has a chi-squared limiting distribution. Also, asymptotically valid critical values can be very simply constructed using the *block* bootstrap. In an empirical illustration, we show that the test is robust across various parameters of interest, such as block lengths, ex-ante prediction periods, and forecast horizons.

7 Appendix

The proof of Theorem 1 relies on Lemmas A1-A4.

Lemma A1: Given AN1-AN5, and recalling that, under H_0, $\widehat{G}_t = \widehat{F}_t$,then:

(i)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left\|\widehat{F}_t - H_F F_{0,t}\right\|^2 = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right),$$

(ii)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1} H_F F_{0,t}\left(\widehat{F}_t - H_F F_{0,t}\right)' = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right),$$

and

(iii)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left(\widehat{F}_t - H_F F_{0,t}\right)'\epsilon_{t+1} = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right).$$

with

$$H_F = V^{-1}\frac{\widehat{F}' F_0}{T}\frac{\Lambda'\Lambda}{N},$$

where V is an $r \times r$ diagonal matrix with the largest r eigenvalues of XX'/NT, in decreasing order.

Lemma A2: Given AN1, AN2', AN3, AN4 and AN5', and recalling that, under H_0, $\widehat{G}_t = \widehat{F}_t$, then:

(i)

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\left\|\widehat{F}_j - H_F F_{0,j}\right\|^2 = O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{T}\right\}\right),$$

(ii)

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t} H_F F_{0,j}\left(\widehat{F}_j - H_F' F_{0,j}\right)' = O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{T}\right\}\right),$$

and

(iii)

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\left(\widehat{F}_j - H_F F_{0,j}\right)'\epsilon_{j+1} = O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{T}\right\}\right).$$

Lemma A3: Given AA1-AA5:

(i)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left\|\widehat{G}_t - H_G G_t\right\|^2 = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right),$$

(ii)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1} H_G G_t\left(\widehat{G}_t - H_G G_t\right)' = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right),$$

and

(iii)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left(\widehat{G}_t - H_G G_t\right)'\epsilon_{t+1} = O_p\left(\max\left\{\frac{\sqrt{T}}{N}, \frac{1}{\sqrt{T}}\right\}\right),$$

with

$$H_G = U^{-1}\frac{\widehat{G}'G}{T}\frac{\Lambda_G'\Lambda_G}{N},$$

where U is an $(q_\lambda+1)r \times (q_\lambda+1)r$ diagonal matrix with the largest $(q_\lambda+1)r$ eigenvalues of XX'/NT, in decreasing order, and where Λ_G is an $N \times (q_\lambda+1)r$ matrix, defined as $\Lambda_G = \left(\begin{array}{cccc} \Lambda_{0,1} & ... & \Lambda_{0,q_{\lambda+1}} \end{array}\right)$.

Lemma A4: Given AA1, AA2', AA3, AA4 and AA5':

(i)

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\left\|\widehat{G}_t - H_G G_t\right\|^2 = O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{T}\right\}\right)$$

and

(ii)

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t} H_G G_t\left(\widehat{G}_t - H_G G_t\right)' = O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{T}\right\}\right).$$

Proof of Lemma A1: Assumption AN(i)-(ii) ensures that Assumption A in Bai and Ng (2006: BN06) is satisfied. In particular, given AN1(i)-(ii), $\frac{1}{T}\sum_{t=1}^{T} F_{0,t}F_{0,t}' - \Sigma_F = o_p(1)$, with Σ_F of rank r. Assumptions AN1(ii)-(iv) and AN2(v) ensure that Assumption B in BN06 is satisfied, AN2(i)-(iv) ensures their Assumption C holds, and AN3 ensures that Assumption E in BN06 is satisfied. Finally, AN3 and AN5 correspond to Assumptions 4 and 5 in Goncalves and Perron (2013: GP13). Indeed Assumption 4 in GP13 is a slightly weakened version of Assumption D in BN06, who instead require loading

factors, idiosyncratic errors, and regressor errors to be mutually independent. All other assumptions in GP13 are the same as in BN06. The statements in (i) and (ii) thus follow from Lemma A1 in BN06, and the statement in (iii) follows from Lemma A1 in GP13.

Proof of Lemma A2: **(i)** From the proof of Theorem 1 in Bai and Ng (2002),

$$\widehat{F}_{j,t-R} - H_F F_{0,j} = \widehat{V}^{-1}\left(\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\gamma_{kj} + \frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\zeta_{kj} + \frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\eta_{kj} + \frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\xi_{kj}\right), \tag{6}$$

where $\widehat{V}$ is an $r \times r$ diagonal matrix containing the largest r eigenvalues of XX'/NR, $\gamma_{kj} = \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N} u_{ik}u_{ij}\right)$, $\zeta_{kj} = \frac{1}{N}\sum_{i=1}^{N}\left(u_{ij}u_{ik} - \gamma_{kj}\right)$, $\eta_{kj} = \frac{1}{N}\sum_{i=1}^{N}\lambda'_{0,i,t}F_{0,k}u_{ij}$ and $\xi_{kj} = \frac{1}{N}\sum_{i=1}^{N}F'_{0,j}\lambda_{0,i,t}u_{ik}$. Furthermore, from Bai and Ng (2002, p.213),

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\widehat{F}_j - H_F F_{0,j}\right\|^2 \\
&\leq \frac{4}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left(\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\gamma_{kj}\right\|^2 + \left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\zeta_{kj}\right\|^2\right. \\
&\left. + \left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\eta_{kj}\right\|^2 + \left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\xi_{kj}\right\|^2\right).
\end{aligned}$$

Now,

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\gamma_{kj}\right\|^2 \\
&\leq \frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\left(\frac{1}{T}\sum_{k=1}^{t}\left\|\widehat{F}_k\right\|^2\right)\left(\frac{1}{T}\sum_{j=t-R+1}^{t}\sum_{k=1}^{T}\gamma_{kj}\right) \\
&= O_p\left(\frac{\sqrt{P}}{T}\right),
\end{aligned}$$

given that, as follows from AN2'(ii), $\frac{1}{T^2}\sum_{k=1}^{T}\sum_{j=t-R+1}^{t}\gamma_{k,j}^2 = O_p\left(\frac{1}{T}\right)$ uniformly in t, by Lemma 1(i) in Bai and Ng (2002), and given that $\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_k\right\|^2 = O_p(1)$, as follows

from AN1(i)-(iv) and AN2'(ii)-(iii). Additionally,

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\zeta_{kj}\right\|^2 \\
&=\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\sum_{k=1}^{T}\sum_{k'=1}^{T}\frac{1}{T^2}\widehat{F}_k'\widehat{F}_{k'}\zeta_{kj}\zeta_{k'j} \\
&\leq\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\left[\left(\frac{1}{T^2}\sum_{k=1}^{T}\sum_{k'=1}^{T}\left(\widehat{F}_k'\widehat{F}_{k'}\right)^2\right)^{1/2}\right. \\
&\left.\left(\frac{1}{T^2}\sum_{k=1}^{T}\sum_{k'=1}^{T}\left(\sum_{j=t+1-R}^{t-1}\zeta_{kj}\zeta_{k'j}\right)^2\right)^{1/2}\right] \\
&\leq\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\left(\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_k\right\|^2\right)\left(\frac{1}{T^2}\sum_{k=1}^{T}\sum_{k'=1}^{T}\left(\sum_{j=t+1-R}^{t-1}\zeta_{kj}\zeta_{k'j}\right)^2\right)^{1/2} \\
&=O_p\left(\frac{\sqrt{P}}{N}\right),
\end{aligned}$$

since,

$$\begin{aligned}
\mathrm{E}\left(\left(\sum_{j=t+1-R}^{t-1}\zeta_{kj}\zeta_{k'j}\right)^2\right)&=\sum_{j=t+1-R}^{t-1}\sum_{j'=t+1-R}^{t-1}\mathrm{E}\left(\zeta_{kj}\zeta_{k'j}\zeta_{kj'}\zeta_{k'j'}\right) \\
&\leq R^2\max_{j,k}\mathrm{E}\left(\zeta_{kj}^4\right)\leq R^2N^{-2}C,
\end{aligned}$$

which follows from AN2'(iii). Letting $u_j=(u_{1,j},...,u_{N,j})$,

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\eta_{kj}\right\|^2 \\
&=\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_kF_{0,k}'\Lambda_0'u_j/N\right\|^2 \\
&\leq\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\frac{1}{N^2}\left\|u_j'\Lambda_0'\right\|^2\left(\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_{k,t-R}\right\|^2\right)\left(\frac{1}{T}\sum_{k=1}^{T}\left\|F_{0,t}'\right\|^2\right) \\
&=O_p\left(1\right)O_p\left(\frac{\sqrt{P}}{N}\right),
\end{aligned}$$

given that $\frac{1}{T}\sum_{k=1}^{t}\widehat{F}_k\widehat{F}_k' = O_p(1)$, as follows from AN1, AN2'(ii) and AN2'(v); and given that $\frac{1}{R}\sum_{j=t-R+1}^{t-1}\left\|\frac{\Lambda_0}{N}u_j\right\|^2 = O_p\left(\frac{1}{N}\right)$ uniformly in t, as follows from AN2(ii) and ARN(iv). Finally, $\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left\|\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\xi_{kj}\right\|^2$ can be treated analogously. As $\widehat{V}^{-1} = O_p(1)$ the statement in (i) follows immediately.

(ii) The proof follows, using a similar argument as that used to prove (iii) below.

(iii) We modify the proof of Lemma A1 in GP13. It is immediate to see that:

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\left(\widehat{F}_j - H_F F_{0,j}\right)\epsilon_{j+1} \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\widehat{V}^{-1}\frac{1}{R}\sum_{j=t+1-R}^{t}\left(\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\gamma_{kj}\epsilon_{j+1} + \frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\zeta_{kj}\epsilon_{j+1}\right. \\
&\left.+\frac{1}{T}\sum_{k=1}^{T}\widehat{F}_k\eta_{kj}\epsilon_{j+h} + \frac{1}{T}\sum_{k=1}^{T}\widehat{F}_{k,t-R}\xi_{kj}\epsilon_{j+h}\right) \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-}\widehat{V}^{-1}\left(I_t + II_t + III_t + IV_t\right).
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}I_t &= \frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\frac{1}{T}\sum_{k=1}^{T}H_F F_{0,k}\gamma_{kj}\epsilon_{j+1} \\
&\quad+\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\frac{1}{T}\sum_{k=1}^{T}\left(\widehat{F}_k - H_F F_{0.k}\right)\gamma_{kj}\epsilon_{j+1} \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}I_{1t} + \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}I_{2t}
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}I_{1t}\right) \\
&\leq \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\frac{1}{R^2}\sum_{j=t-R+1}^{t}\sum_{k=1}^{T}\left|\gamma_{kj}\right|\left(\mathrm{E}\left(\left\|H_F F_{0,k}\right\|^2\right)\right)^{1/2}\left(\mathrm{E}\left(\epsilon_{j+1}\right)\right)^{1/2} \\
&= O\left(\frac{\sqrt{P}}{R}\right)O(1),
\end{aligned}$$

as $\frac{1}{R^2}\sum_{j=t-R+1}^{t}\sum_{k=1}^{T}\left|\gamma_{kj}\right| = O\left(\frac{1}{R}\right)$, which follows by AN2'(ii). Hence, given AN1(ii) and AN4(iii), $\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h} I_{1t} = O_p\left(\frac{\sqrt{P}}{R}\right)$. Now,

$$
\begin{aligned}
&\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h}\left\|I_{2t}\right\| \\
&\leq \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\left\|\frac{1}{T}\sum_{k=1}^{T}\left(\widehat{F}_k - H_F F_{0,k}\right)\frac{1}{R}\sum_{j=t-R+1}^{t}\gamma_{kj}\epsilon_{j+1}\right\| \\
&\leq \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\left(\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_k - H_F F_{0,k}\right\|^2\right)^{1/2}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\sum_{k=1}^{T}\left|\gamma_{kj}\right|^2\frac{1}{R}\sum_{j=t-R+1}^{t}\epsilon_{j+h}^2\right)^{1/2} \\
&= O_p\left(\max\left\{\frac{\sqrt{P}}{R},\frac{\sqrt{P}}{N}\right\}\right)O_p(1),
\end{aligned}
$$

which follows from part (i), and since the term in the second bracket is $O_p(1)$, uniformly in t, given AN2'(ii), AN1(ii), and AN4(ii). Now,

$$
\begin{aligned}
\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} II_t &= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\frac{1}{TR}\sum_{k=1}^{T} H_F F_{0,k}\zeta_{kj}\epsilon_{j+1} \\
&\quad + \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\frac{1}{T}\sum_{k=1}^{T}\left(\widehat{F}_k - H_F F_{0,k}\right)\frac{1}{R}\sum_{j=t-R+1}^{t}\zeta_{kj}\epsilon_{j+1} \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h} II_{1t} + \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h} II_{2t}.
\end{aligned}
$$

Recalling the definition of ζ_{kj},

$$
\begin{aligned}
&\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} II_{1t} \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\frac{1}{NTR}\sum_{k=1}^{T}\sum_{j=t-R+1}^{t}\sum_{i=1}^{N} H_F F_{0,k}\left(u_{ij}u_{ik} - \mathrm{E}\left(u_{ij}u_{ik}\right)\right)\epsilon_{j+h} \\
&= \frac{1}{\sqrt{P}}\frac{1}{\sqrt{NT}}\sum_{t=R+1}^{T-1}\frac{1}{R}\sum_{j=t-R+1}^{t} m_{j,t}\epsilon_{j+1},
\end{aligned}
$$

where $m_{j,t} = \frac{1}{\sqrt{NT}}\sum_{k=1}^{T}\sum_{i=1}^{N} H_F F_{0,k}\left(u_{ij}u_{ik} - \mathrm{E}\left(u_{ij}u_{ik}\right)\right)$. We need to show that $\frac{1}{R}\sum_{j=t-R+1}^{t} m_{j,t}\epsilon_{j+h} =$

$O_p(1)$, uniformly in t, thus implying that $\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h} II_{1t} = O_p\left(\frac{\sqrt{P}}{\sqrt{NT}}\right)$. Now,

$$\left\|\frac{1}{R}\sum_{j=t-R+1}^{t} m_{j,t}\epsilon_{j+h}\right\| \leq \left(\frac{1}{R}\sum_{j=t-R+1}^{t}\|m_{j,t}\|^2\right)^{1/2}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\epsilon_{j+h}^2\right)^{1/2}$$
$$= \left(\frac{1}{R}\sum_{j=t-R+1}^{t-h}\|m_{j,t}\|^2\right)^{1/2} O_p(1),$$

because $\frac{1}{R}\sum_{j=t-R+1}^{t}\epsilon_{j+h}^2 = O_p(1)$ uniformly in t, as it satisfies a uniform law of large numbers, given AN1(i)-(iii) and AN4(ii). Also,

$$\frac{1}{R}\sum_{j=t-R+1}^{t}\mathrm{E}\|m_{j,t}\|^2 \leq \mathrm{E}\left\|\frac{1}{\sqrt{NT}}\sum_{k=1}^{T}\sum_{i=1}^{N} H_F F_0\left(u_{ij}u_{ik} - \mathrm{E}\left(u_{ij}u_{ik}\right)\right)\right\|^2 = O(1),$$

given AN3(i). Recalling the statement in part (i),

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\|II_{2t}\|$$
$$\leq \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\left(\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_k - H_F F_{0,k}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{k=1}^{T}\left|\frac{1}{R}\sum_{j=t-R+1}^{t}\zeta_{kj}\epsilon_{j+h}\right|^2\right)^{1/2}$$
$$= O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{R}\right\}\right) o_p(1),$$

as

$$\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}\left|\sup_{t\geq R}\frac{1}{R}\sum_{j=t-R+1}^{t}\zeta_{kj}\epsilon_{j+h}\right|^2$$
$$= \frac{1}{NR}\frac{1}{T}\sum_{k=1}^{T}\mathrm{E}\left|\sup_{t\geq R}\sqrt{\frac{1}{RN}}\sum_{j=t-R+1}^{t}\sum_{i=1}^{N}\left(u_{ij}u_{ik} - \mathrm{E}\left(u_{ij}u_{ik}\right)\right)\epsilon_{j+h}\right|^2$$
$$= O\left(\frac{P}{NR}\right) = o(1) \text{ uniformly in } t,$$

given AN5'(i). Now,

$$\begin{aligned}
&\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} III_t \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t}\frac{1}{RT}\sum_{k=1}^{T} H_F F_{0,k}\eta_{kj}\epsilon_{j+h} \\
&+ \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\frac{1}{RT}\sum_{j=t+1-R}^{t}\sum_{k=1}^{T}\left(\widehat{F}_k - H_F F_{0,k}\right)\eta_{kj}\epsilon_{j+1} \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} III_{1t} + \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} III_{2t},
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} III_{1t} &= \frac{1}{P}\sum_{t=R+1}^{T-1}\frac{1}{T}\sum_{k=1}^{T} H_F F_{0,k} F_{0,k}' H_F' \left(\frac{\sqrt{P}}{R}\sum_{j=t+1-R}^{t}\frac{\Lambda_0' u_j}{N}\epsilon_{j+h}\right) \\
&= O_p\left(\frac{\sqrt{P}}{\sqrt{RN}}\right),
\end{aligned}$$

given AN5'(ii). Also,

$$\begin{aligned}
&\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\|III_{2t}\| \\
&\leq \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\left(\frac{1}{T}\sum_{k=1}^{T}\left\|\widehat{F}_k - H_F F_{0,k}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{k=1}^{T}\left|\frac{1}{R}\sum_{j=t-R+1}^{t}\eta_{kj}\epsilon_{j+h}\right|^2\right)^{1/2} \\
&= O_p\left(\max\left\{\frac{\sqrt{P}}{N},\frac{\sqrt{P}}{T}\right\}\right) o_p(1),
\end{aligned}$$

where the $o_p(1)$ term follows from AN5'(ii). Hence, $\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} III_t = O_p\left(\frac{\sqrt{P}}{\sqrt{RN}}\right) + O_p\left(\max\left\{\frac{\sqrt{P}}{N},\frac{\sqrt{P}}{T}\right\}\right) o_p(1)$. Finally,

$$\begin{aligned}
&\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1} IV_t \\
&= \frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-1}\sum_{j=1}^{t}\frac{1}{RT}\sum_{k=1}^{T} H_F F_{0,k}\xi_{kj}\epsilon_{j+h}
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} \frac{1}{RT} \sum_{j=t+1-R}^{t} \sum_{k=1}^{T} \left(\widehat{F}_k - H_F F_{0,k} \right) \xi_{kj} \epsilon_{j+h} \\
&= \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} IV_{1t} + \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} IV_{2t},
\end{aligned}
$$

where

$$
\begin{aligned}
&\frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} IV_{1t} \\
&= \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} \frac{1}{RT} \sum_{j=t+1-R}^{t} \sum_{k=1}^{T} H_F F_{0,k} \left(F_{0,j}' H_F' \frac{\Lambda_0' u_k}{N} \epsilon_{j+h} \right) \\
&= \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} \frac{1}{T} \sum_{k=1}^{T} \left(H_F F_{0,k} \frac{u_k' \Lambda_t}{N} \right) \frac{1}{R} \sum_{j=t+1-R}^{t} \left(F_{0,j}' H_F' \epsilon_{j+h} \right) \\
&= O_p \left(\frac{\sqrt{P}}{\sqrt{NR}} \right) o_p(1),
\end{aligned}
$$

because of AN3(ii) and because $\sup_{R \leq t \leq T} \left| \frac{1}{R} \sum_{j=t+1-R}^{t} \left(F_{0,j}' H_F' \epsilon_{j+h} \right) \right| = o_p(1)$, by the uniform law of large numbers, given AN1(i)-(ii) and AN4(i)-(ii). Also,

$$
\begin{aligned}
&\frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} \| IV_{2t} \| \\
&\leq \frac{1}{\sqrt{P}} \sum_{t=R+1}^{T-1} \left(\frac{1}{T} \sum_{k=1}^{T} \left\| \widehat{F}_k - H_F F_{0,k} \right\|^2 \right)^{1/2} \left(\frac{1}{T} \sum_{k=1}^{T} \left| \frac{1}{R} \sum_{j=t-R+1}^{t} \xi_{kj} \epsilon_{j+h} \right|^2 \right)^{1/2} \\
&= O_p \left(\max \left\{ \frac{\sqrt{P}}{R}, \frac{\sqrt{P}}{N} \right\} \times \frac{1}{\sqrt{N}} \right),
\end{aligned}
$$

given the statement in (i), and because

$$
\begin{aligned}
&\frac{1}{T} \sum_{k=1}^{T} \left| \frac{1}{R} \sum_{j=t-R+1}^{t-h} \xi_{kj} \epsilon_{j+h} \right|^2 \\
&\leq \frac{1}{N} \left(\frac{1}{T} \sum_{k=1}^{T} \left\| \frac{u_j \Lambda_0}{\sqrt{N}} \right\|^2 \right) \left\| \frac{1}{R} \sum_{j=t-R+1}^{t} H_F F_{0,j} \epsilon_{j+h} \right\|^2 \\
&= \frac{1}{N} O_p(1),
\end{aligned}
$$

because of AN3(iii). Hence, $\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T-h} IV_{2t} = O_p\left(\frac{\sqrt{P}}{\sqrt{NR}}\right) o_p(1) + O_p\left(\max\left\{\frac{\sqrt{P}}{R}, \frac{\sqrt{P}}{N}\right\} \times \frac{1}{\sqrt{N}}\right)$.

Thus,

$$\begin{aligned}
&\frac{1}{\sqrt{P}R}\sum_{t=R+1}^{T-1}\sum_{j=t+1-R}^{t-h}\left(\widehat{F}_j - H_F F_{0,j}\right)\epsilon_{j+h} \\
&= O_p\left(\frac{\sqrt{P}}{R}\right) + O_p\left(\frac{\sqrt{P}}{\sqrt{NR}}\right) + O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{R}\right\}\right) o_p(1) \\
&= O_p\left(\max\left\{\frac{\sqrt{P}}{N}, \frac{\sqrt{P}}{R}\right\}\right).
\end{aligned}$$

Proof of Lemma A3: Given AA1-AA5, the proof follows by Lemma 10 in Han and Inoue (2012).

Proof of Lemma A4: The proof follows, using a similar argument as that used in the proof of Lemma A2.

Proof of Theorem 1: **(i)** Recalling Lemma A1, we have that:

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t \\
&= \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}F_{0,t-1}'H_F'H_F'^{-1}\beta_0 + \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}\epsilon_t \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(H_F F_{0,t-1}F_{0,t-1}'H_F'\right)H_F'^{-1}\beta_0 + \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{F}_{t-1} - H_F F_{0,t-1}\right)F_{0,t-1}'H_F'H_F'^{-1}\beta_0 + \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}\epsilon_t \\
&= \left(H_F\Sigma_{FF}H_F'\right)H_F'^{-1}\beta_0 + \frac{1}{T}\sum_{t=1}^{T}H_F F_{0,t-1}\epsilon_t + o_p\left(\frac{1}{\sqrt{T}}\right) \\
&= H_F'^{-1}\beta_0 + \frac{1}{T}\sum_{t=1}^{T}H_F F_{0,t-1}\epsilon_t + o_p\left(\frac{1}{\sqrt{T}}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t - \beta_0'H_F^{-1}\right) \\
&= \sqrt{\frac{\pi}{1+\pi}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}H_F F_{0,t-1}\epsilon_t + o_p(1).
\end{aligned}$$

Additionally, for all $t \geq R$,

$$\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j$$
$$=\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}\widehat{F}_{j-1}'H_F'^{-1}\beta_0+\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}\left(H_F F_{0,j-1}-\widehat{F}_{j-1}\right)'H_F'^{-1}\beta_0+\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}\epsilon_j.$$

Hence, given Lemma A2(i)-(ii),

$$\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-\beta_0'H_F^{-1}\right)$$
$$=\frac{1}{\sqrt{P}}\sum_{t=1}^{R}\frac{1}{R}\sum_{j=t-R+1}^{t}H_F F_{0,j-1}\epsilon_j+o_p(1).$$

This in turn ensures that:

$$\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)\xrightarrow{d}N\left(0,\Omega_0\right),$$

with

$$\Omega_0=\operatorname{avar}\left(\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-\beta_0'H_F^{-1}\right)\right)+\operatorname{avar}\left(\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t-\beta_0'H_F^{-1}\right)\right)$$
$$-2\operatorname{acov}\left(\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-\beta_0'H_F^{-1}\right),\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t-\beta_0'H_F^{-1}\right)\right)$$
$$=1\left\{\pi\leq 1\right\}\left(\pi-\frac{\pi^2}{3}\right)V_{\beta_0}+1\left\{\pi>1\right\}\left(1-\frac{1}{3\pi}\right)V_{\beta_0}-\frac{\pi}{1+\pi}V_{\beta_0},$$

as

$$\operatorname{avar}\left(\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t-\beta_0'H_F^{-1}\right)\right)=\frac{\pi}{1+\pi}V_{\beta_0}$$

$$V_{\beta_0}=H_F\sum_{j=-\infty}^{\infty}\mathrm{E}\left(F_{0,1}F_{0,1+j}'\epsilon_1\epsilon_{1+j}\right)H_F',$$

and from Lemma 4.1 in West and McCracken (1998),

$$\begin{aligned}&\text{avar}\left(\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-\beta_0'H_F^{-1}\right)\right)\\&=1\left\{\pi\leq 1\right\}\left(\pi-\frac{\pi^2}{3}\right)V_{\beta_0}+1\left\{\pi>1\right\}\left(1-\frac{1}{3\pi}\right)V_{\beta_0}.\end{aligned}$$

Finally,

$$\begin{aligned}&\text{acov}\left(\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}_{j-1}y_j-H^{-1}\beta_0\right),\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t-H^{-1}\beta_0\right)\right)\\&=\frac{P}{T}\text{E}\left(\frac{1}{\sqrt{R}}\sum_{t=1}^{T}F_{0,t-1}\epsilon_t\frac{1}{P}\left(\frac{1}{\sqrt{R}}\sum_{j=1}^{R}F_{0,j-1}\epsilon_j+....+\frac{1}{\sqrt{R}}\sum_{j=T-R+1}^{T}F_{0,j-1}\epsilon_j\right)\right)+o(1)\\&=\frac{P}{T}V_{\beta_0}=\frac{\pi}{1+\pi}V_{\beta_0}.\end{aligned}$$

(ii) Without loss of generality, and just for notational simplicity, assume that $\Lambda_{0,t}=\begin{cases}\Lambda_{0,1}1\left\{t/T\leq\tau_\lambda\right\}\\ \Lambda_{0,2}1\left\{t/T>\tau_\lambda\right\}\end{cases}$, $\beta_{0,t}=\begin{cases}\beta_{0,1}1\left\{t/T\leq\tau_\beta\right\}\\ \beta_{0,2}1\left\{t/T>\tau_\beta\right\}\end{cases}$ and $\tau_\lambda>\tau_\beta$. Thus,

$$X=\begin{pmatrix}F_{01}&0\\0&F_{02}\end{pmatrix}\begin{pmatrix}\Lambda_{0,1}\\\Lambda_{0,2}\end{pmatrix}+u,$$

where $F_{01}=(F_{0,1},...,F_{0,\tau_\lambda T})'$ is $\tau_\lambda T\times r$, $F_{02}=(F_{0,\tau_\lambda T},...,F_{0,T})'$ is $(T-\tau_\lambda T)\times r$, $\Lambda_{0,1}$ and $\Lambda_{0,2}$ are $r\times N$. Finally, note that

$$y_{t+1}=F_{0,t}'\beta_{0,1}1\left\{t/T\leq\tau_\beta\right\}+F_{0,t}'\beta_{0,2}1\left\{t/T>\tau_\beta\right\}+\epsilon_{t+1},$$

so that the variable to be predicted depends on r factors. As shown in the (unnumbered) Proposition in Han and Inoue (2012), the criterion of Bai and Ng (2002), applied to the full sample, will detect $2r$ factors. Hereafter, let

$$G_t=\begin{pmatrix}F_{0,t}1\left\{t/T\leq\tau_\lambda\right\}&0\times 1\left\{t/T>\tau_\lambda\right\}\\0\times 1\left\{t/T\leq\tau_\lambda\right\}&F_{0,t}1\left\{t/T>\tau_\lambda\right\}\end{pmatrix},$$

and let $\widehat{G}_t$ be the principal component estimator of G_t. By the same argument as in

Lemma 10 in Han and Inoue (2012),

$$\frac{1}{T}\sum_{t=2}^{T}\left\|\widehat{G}_{t-1}-H_G G_t\right\| = o_p\left(\max\left\{\frac{1}{N},\frac{1}{T}\right\}\right).$$

As we would detect $2r$ factors, and thus estimate $2r$ factors, when the alternative is true, the statistic is

$$\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j - \frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t\right).$$

Now, because of Lemma A3,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t \\
&= \frac{1}{T}\sum_{t=1}^{T}H_G G_{t-1}\left(F_{0,t}'\beta_{0,1}1\left\{t/T\le\tau_\beta\right\} + F_{0,t}'\beta_{0,2}1\left\{t/T>\tau_\beta\right\}\right) + \epsilon_t + o_p\left(\frac{1}{\sqrt{T}}\right) \\
&= \left(\frac{1}{T}\sum_{t=1}^{T\tau_\beta}H_G\begin{pmatrix}F_{0,t}\\0\end{pmatrix}F_{0,t}'\beta_{0,1} + \frac{1}{T}\sum_{t=T\tau_\beta+1}^{T\tau_\lambda}H_G\begin{pmatrix}F_{0,t}\\0\end{pmatrix}F_{0,t}'\beta_{0,2}\right. \\
&\left.\frac{1}{T}\sum_{t=T\tau_\lambda+1}^{T}H_G\begin{pmatrix}0\\F_{0,t}\end{pmatrix}F_{0,t}'\beta_{0,2} + \frac{1}{T}\sum_{t=1}^{T\tau_\lambda}H_G\begin{pmatrix}F_{0,t}\\0\end{pmatrix}\epsilon_t + \frac{1}{T}\sum_{t=T\tau_\lambda+1}^{T}H_G\begin{pmatrix}0\\F_{0,t}\end{pmatrix}\epsilon_t\right) \\
&+ o_p\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}$$

and so

$$\begin{aligned}
&\operatorname{p}\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t \\
&= \left(\tau_\beta H_G\Sigma_{F_0,0}\beta_{0,1} + \left(\tau_\lambda-\tau_\beta\right)H_G\Sigma_{F_0,0}\beta_{0,2} + \left(1-\tau_\lambda\right)H_G\Sigma_{0,F_0}\beta_{0,2}\right),
\end{aligned}$$

with

$$\Sigma_{F_{0,0}} = \begin{pmatrix}\operatorname{p}\lim_{T\to\infty}\frac{1}{T\tau_\beta}\sum_{t=2}^{T\tau_\beta}F_{0,t-1}F_{0,t-1}'\\0\end{pmatrix},\ \Sigma_{0,F_0} = \begin{pmatrix}0\\\operatorname{p}\lim_{T\to\infty}\frac{1}{T(1-\tau_\lambda)}\sum_{t=T\tau_\lambda+1}^{T}F_{0,t-1}F_{0,t-1}'\end{pmatrix},$$

and

$$\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{g}_{t-1}y_t - \left(\tau_\beta H_G\Sigma_{F_0,0}\beta_{0,1} + \left(\tau_\lambda-\tau_\beta\right)H_G\Sigma_{F_0,0}\beta_{0,2} + \left(1-\tau_\lambda\right)H_G\Sigma_{0,F_0}\beta_{0,2}\right)\right)$$

$$\xrightarrow{d} N\left(0, \frac{\pi}{1+\pi}\left(\tau_\lambda H_G \begin{pmatrix} V_\beta & 0 \\ 0 & 0 \end{pmatrix} H_G' + (1-\tau_\lambda) H_G \begin{pmatrix} 0 & 0 \\ 0 & V_\beta \end{pmatrix} H_G'\right)\right).$$

We now turn to the rolling window "scheme" estimator. Given that we are considering averaging over rolling windows of length R, the closed expression for the probability limit depends on whether $\tau_\beta T + R$ is smaller, equal to, or larger than T. In the sequel, without loss of generality, assume that $\tau_\beta T \leq R$, $\tau_\lambda T > R$, $\tau_\lambda T - \tau_\beta T < R$, and $\tau_\beta T + R = T$. All other cases follow by exactly the same argument. Recalling Lemma A4, note that:

$$\begin{aligned}
&\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right) \\
&= \frac{1}{PR}\sum_{t=1}^{\tau_\lambda T-R}\left(\sum_{j=t}^{\tau_\beta T}\widehat{G}_{j-1}\left(F_{0,j-1}'\beta_{0,1}+\epsilon_j\right) + \sum_{j=\tau_\beta T+1}^{R+t}\widehat{G}_{j-1}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right)\right) \\
&+ \frac{1}{PR}\sum_{t=1}^{R-(\tau_\lambda-\tau_\beta)T}\left(\sum_{j=\tau_\lambda T-R+t}^{\tau_\beta T}\widehat{G}_{j-1}\left(F_{0,j-1}'\beta_{0,1}+\epsilon_j\right)\right. \\
&\left. + \sum_{j=\tau_\beta T+1}^{\tau_\lambda T}\widehat{G}_{j-1}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right) + \sum_{j=\tau_\lambda T+1}^{\tau_\lambda T+t}\widehat{G}_{j-1}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right)\right) \\
&= \frac{1}{PR}\sum_{t=1}^{\tau_\lambda T-R}\left(\sum_{j=t}^{\tau_\beta T}H_G\begin{pmatrix} F_{0,j-1} \\ 0 \end{pmatrix}\left(F_{0,j-1}'\beta_{0,1}+\epsilon_j\right) + \sum_{j=\tau_\beta T+1}^{R+t}H_G\begin{pmatrix} F_{0,j-1} \\ 0 \end{pmatrix}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right)\right) \\
&+ \frac{1}{PR}\sum_{t=1}^{R-(\tau_\lambda-\tau_\beta)T}\left(\sum_{j=\tau_\lambda T-R+t}^{\tau_\beta T}H_G\begin{pmatrix} F_{0,j-1} \\ 0 \end{pmatrix}\left(F_{0,j-1}'\beta_{0,1}+\epsilon_j\right) + \sum_{j=\tau_\beta T+1}^{\tau_\lambda T}H_G\begin{pmatrix} F_{0,j-1} \\ 0 \end{pmatrix}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right)\right. \\
&\left. + \sum_{j=\tau_\lambda T+1}^{\tau_\lambda T+t}H_G\begin{pmatrix} 0 \\ F_{0,j-1} \end{pmatrix}\left(F_{0,j-1}'\beta_{0,2}+\epsilon_j\right)\right) + o_p\left(\max\left\{\frac{1}{\sqrt{P}}, \frac{1}{\sqrt{R}}\right\}\right).
\end{aligned}$$

Now, let

$\omega_1 = \lim_{P,R,T\to\infty}\frac{\tau_\lambda T-R+1}{P}$, $\omega_2 = \lim_{P,R,T\to\infty}\frac{1}{\tau_\lambda T-R+1}\sum_{t=1}^{\tau_\lambda T-R}\frac{\tau_\beta T-t}{R}$,
$\omega_3 = \lim_{P,R,T\to\infty}\frac{1}{\tau_\lambda T-R}\sum_{t=1}^{\tau_\lambda T-R}\frac{R+t-\tau_\beta T}{R}$,
$\omega_4 = \lim_{P,R,T\to\infty}\frac{R-(\tau_\lambda-\tau_\beta)T}{P}$,
$\omega_5 = \lim_{P,R,T\to\infty}\frac{1}{R-(\tau_\lambda-\tau_\beta)T}\sum_{t=1}^{R-(\tau_\lambda-\tau_\beta)T}\frac{(\tau_\beta-\tau_\lambda)T+R-t}{R}$, and
$\omega_6 = \lim_{P,R,T\to\infty}\frac{(\tau_\lambda-\tau_\beta)T}{R}$ and $\omega_7 = \lim_{P,R,T\to\infty}\frac{1}{R-(\tau_\lambda-\tau_\beta)T}\sum_{t=1}^{R-(\tau_\lambda-\tau_\beta)T}\frac{t}{R}$.

It follows that:

$$\begin{aligned}
&\mathrm{p}\lim_{\mathrm{R,P,T}\to\infty}\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right)\\
&=\left(\omega_1\omega_2+\omega_4\omega_5\right)H_G\Sigma_{F_0,0}\beta_{0,1}+\left(\omega_1\omega_3+\omega_4\omega_6\right)H_G\Sigma_{F_0,0}\beta_{0,2}+\omega_4\omega_7H_G\Sigma_{0,F_0}\beta_{0,2}\\
&\neq\mathrm{p}\lim_{\mathrm{R,P,T}\to\infty}\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t.
\end{aligned}$$

Also,

$$\begin{aligned}
&\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right)-\mathrm{p}\lim_{\mathrm{R,P,T}\to\infty}\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right)\right)\\
&=\omega_1\frac{1}{\tau_\lambda T-R+1}\sum_{t=1}^{\tau_\lambda T-R+1}\left(\omega_2\frac{1}{\tau_\beta T-t}\sum_{j=t}^{\tau_\beta T}H_G\begin{pmatrix}F_{0,j-1}\\0\end{pmatrix}\epsilon_j\right.\\
&\left.+\ \omega_3\frac{1}{R+t-\tau_\beta T}\sum_{j=\tau_\beta T+1}^{R+t}H_G\begin{pmatrix}F_{0,j-1}\\0\end{pmatrix}\epsilon_j\right)\\
&+\omega_4\frac{1}{R-\left(\tau_\lambda-\tau_\beta\right)T}\sum_{t=1}^{R-\left(\tau_\lambda-\tau_\beta\right)T}\left(\omega_5\frac{1}{\left(\tau_\beta-\tau_\lambda\right)T-t+R}\sum_{j=\tau_\lambda T+R-t}^{\tau_\beta T}H_G\begin{pmatrix}F_{0,j-1}\\0\end{pmatrix}\epsilon_j\right.\\
&\left.+\omega_6\frac{1}{\tau_\lambda T-\tau_\beta T}\sum_{j=\tau_\beta T+1}^{\tau_\lambda T}H_G\begin{pmatrix}F_{0,j-1}\\0\end{pmatrix}\epsilon_j+\omega_7\frac{1}{t}\sum_{j=\tau_\lambda T+1}^{\tau_\lambda T+t}H_G\begin{pmatrix}0\\F_{0,j-1}\end{pmatrix}\epsilon_j\right)\\
&\xrightarrow{d}N\left(0,\Omega\right),
\end{aligned}$$

with

$$\begin{aligned}
\Omega&=\left(\omega_1\left(\omega_2+\omega_3\right)+\omega_4\left(\omega_5+\omega_6\right)\right)^2H_G\begin{pmatrix}V_\beta&0\\0&0\end{pmatrix}H_G'\\
&\quad+\left(\omega_4\omega_7\right)^2H_G\begin{pmatrix}0&0\\0&V_\beta\end{pmatrix}H_G'.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}_{j-1}y_j\right)-\frac{1}{T}\sum_{t=1}^{T}\widehat{G}_{t-1}y_t\right)\\
&=\sqrt{P}\left(\left(\left(\omega_1\omega_2+\omega_4\omega_5\right)H_G\Sigma_{F_0,0}-\tau_\beta H_G\Sigma_{F_0,0}\right)\beta_{0,1}\right.
\end{aligned}$$

$$+ \left(\left(\left(\omega_1 \omega_3 + \omega_4 \omega_6 \right) H_G \Sigma_{F_0,0} + \omega_4 \omega_7 H_G \Sigma_{0,F_0} \right) \right.$$
$$\left. - \left(\left(\tau_\lambda - \tau_\beta \right) H_G \Sigma_{F_0,0} + \left(1 - \tau_\lambda \right) H_G \Sigma_{0,F_0} \right) \beta_{0,2} \right) + O_p(1)$$

The statement in (ii) then follows.

Proof of Theorem 2: **(i)** Recalling that $\mathrm{E}^* \left(\widehat{F}^*_{j-1} y^*_j \right) = \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k + O_{p^*} \left(\frac{l}{\sqrt{P}} \right)$,

$$\mathrm{E}^* \left(\sqrt{P} \left(\frac{1}{P} \sum_{t=R+1}^{T} \left(\frac{1}{R} \sum_{j=t-R+1}^{t} \left(\widehat{F}^*_{j-1} y^*_j - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) - \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{F}^*_{t-1} y^*_t - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) \right)$$
$$= O_p \left(\frac{l}{\sqrt{P}} \right) = o_p(1). \tag{7}$$

Now,

$$\mathrm{var}^* \left(\sqrt{P} \left(\frac{1}{P} \sum_{t=R+1}^{T} \left(\frac{1}{R} \sum_{j=t-R+1}^{t} \left(\widehat{F}^*_{j-1} y^*_j - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) - \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{F}^*_{t-1} y^*_t - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) \right)$$
$$= \mathrm{var}^* \left(\sqrt{P} \left(\frac{1}{T} \sum_{t=1}^{T} \left(\widehat{F}^*_{t-1} y^*_t - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) \right)$$
$$+ \mathrm{var}^* \left(\frac{1}{\sqrt{P}} \sum_{t=R+1}^{T} \left(\frac{1}{R} \sum_{j=t-R+1}^{t} \left(\widehat{F}^*_{j-1} y^*_j - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right) \right)$$
$$- 2\mathrm{cov}^* \left(\frac{1}{\sqrt{P}} \sum_{t=R+1}^{T} \left(\frac{1}{R} \sum_{j=t-R+1}^{t} \left(\widehat{F}^*_{j-1} y^*_j - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right), \frac{\sqrt{P}}{T} \sum_{t=1}^{T} \left(\widehat{F}^*_{t-1} y^*_t - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \right)$$

and

$$\mathrm{var}^* \left(\sqrt{P} \left(\frac{1}{T} \sum_{t=1}^{T} \left(\widehat{F}^*_{t-1} y^*_t - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right) \right) \right)$$
$$= \frac{P}{T^2} \sum_{k=1}^{b} \sum_{i=1}^{l} \sum_{i'=1}^{l} \mathrm{E}^* \left(\left(\widehat{F}_{I_k+i} y_{I_k+i} - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right) \left(\widehat{F}_{I_k+i'} y_{I_k+i'} - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right)' \right)$$
$$= \frac{P}{T} \frac{1}{T-l+1} \frac{1}{l} \sum_{t=1}^{T-l} \sum_{i=1}^{l} \sum_{i'=1}^{l} \left(\widehat{F}_{t-1+i} y_{t+i} - \frac{1}{T} \sum_{k=1}^{T} \widehat{F}_{k-1} y_k \right) \left(\widehat{F}_{t-1+i'} y_{t+i'} - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right)' + O_p \left(\frac{l^2}{P} \right)$$
$$= \frac{P}{T} \frac{1}{T} \sum_{t=1}^{T-l} \sum_{i=1}^{l} \left(\widehat{F}_{t-1} y_t - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right) \left(\widehat{F}_{t-1+i} y_{t+i} - \frac{1}{T} \sum_{t=1}^{T} \widehat{F}_{t-1} y_t \right)' + O_p \left(\frac{l^2}{P} \right)$$

$$= \frac{P}{T}\frac{1}{T}\sum_{t=l}^{T-l}\sum_{i=-l}^{l}\left(H_F F_{0,t-1}y_t - \frac{1}{T}\sum_{t=1}^{T} H_F F_{0,t-1}y_t\right)\left(H_F F_{0,t-1+i}y_{t+i} - \frac{1}{T}\sum_{t=1}^{T} H_F F_{0,t-1}y_t\right)' + + o_p(1)$$

$$= \frac{P}{T}\frac{1}{T}\sum_{t=l}^{T-l}\sum_{i=-l}^{l}\left(H_F F_{0,t-1}\epsilon_t\right)\left(H_F F_{0,t-1+i}\epsilon_{t+i}\right)' + o_p(1)$$

$$= \frac{\pi}{\pi+1}V_{\beta_0} + o_p(1).$$

Assuming that $P \leq R$, and following West and McCracken (1998), we have that:

$$\begin{aligned}
&\text{var}^*\left(\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}^*_{j-1}y^*_j\right)\right) \\
&= \text{var}^*\left(\frac{1}{\sqrt{P}R}\sum_{j=1}^{P-1} j\widehat{F}^*_{j-1}y^*_j\right) + \text{var}^*\left(\frac{\sqrt{P}}{R}\sum_{j=P}^{R}\widehat{F}^*_{j-1}y^*_j\right) \\
&+ \text{var}^*\left(\frac{1}{\sqrt{P}R}\sum_{j=R+1}^{T}(T-j)\widehat{F}^*_{j-1}y^*_j\right) + O_p\left(\frac{l^2}{P}\right). \qquad (8)
\end{aligned}$$

Now, let $b = b_1 + b_2 + b_3$, with $b_1 l = P$, $b_2 l = (R-P)$, $b_3 l = (T-R-1)$. Then, following the proof of Proposition 2 in Corradi and Swanson (2006),

$$\begin{aligned}
&\text{var}^*\left(\frac{1}{\sqrt{P}R}\sum_{j=1}^{P-1} j\widehat{F}^*_{j-1}y^*_j\right) \\
&= \text{var}^*\left(\frac{1}{\sqrt{P}R}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\left((k-1)l+i\right)\widehat{F}^*_{I_k+i}y^*_{I_k+i}\right) \\
&= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right) \\
&\times \text{E}^*\left(\left(\widehat{F}^*_{I_k+i}y^*_{I_k+i} - \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)\left(\widehat{F}^*_{I_k+i}y^*_{I_k+i'} - \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)'\right) \\
&= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right) \\
&\times \frac{1}{T}\sum_{t=1}^{T-l}\left(\widehat{F}_{t-1+i}y_{t+i} - \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)\left(\widehat{F}_{t-1+i'}y_{t+i'} - \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)' + O_p\left(\frac{l^2}{P}\right)
\end{aligned}$$

$$= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=-l}^{l}\sum_{i'=-l}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right)$$
$$\times \frac{1}{T}\sum_{t=l}^{T-l}\left(HF_{0,t-1+i}y_{t+i} - \frac{1}{T}\sum_{t=1}^{T}HF_{0,t-1}y_t\right)\left(HF_{0,t-1+i'}y_{t+i'} - \frac{1}{T}\sum_{t=1}^{T}HF_{0,t-1}y_t\right)' + o_p(1)$$

$$= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right)$$
$$\times \frac{1}{T}\sum_{t=1}^{T-l}\left(H_F F_{0,t-1+i}\epsilon_{t+i}\right)\left(H_F F_{0,t-1+i'}\epsilon_{t+i'}\right)' + o_p(1)$$
$$= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right)\mathrm{E}\left(\left(H_F F_{0,i}\epsilon_{1+i}\right)\left(H_F F_{0,i'}\epsilon_{t+i'}\right)'\right) + o_p(1)$$
$$= \frac{1}{T}\sum_{t=l}^{T-l}\sum_{i=-l}^{l}t(t+i)\mathrm{E}\left(\left(H_F F_{0,0}\epsilon_1\right)\left(H_F F_{0,i}\epsilon_{t+i}\right)'\right) + o_p(1)$$
$$\xrightarrow{pr} \frac{\pi^2}{3}V_{\beta_0}.$$

Via a similar argument, it follows that, up to an $o_p(1)$ terms, the second and third terms on the RHS of (8) are equal to $\left(\pi - \pi^2\right)V_{\beta_0}$ and $\frac{\pi^2}{3}V_{\beta_0}$, respectively. The case of $P > R$ can be treated in a similar manner, following the proof of Lemma 4.1 in West and McCracken (1998). Hence,

$$\mathrm{var}^*\left(\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{F}^*_{j-1}y^*_j\right)\right)$$
$$= 1\left\{\pi \leq 1\right\}\left(\pi - \frac{\pi^2}{3}\right)V_{\beta_0} + 1\left\{\pi > 1\right\}\left(1 - \frac{1}{3\pi}\right)V_{\beta_0} + o_p(1).$$

Finally, note that the bootstrap analog of the covariance term mimics the sample covariance, up to a $O_p(l^2/P)$ term, as required.

(ii) Recalling that $\mathrm{E}^*\left(\widehat{G}^*_{j-1}y^*_j\right) = \frac{1}{T}\sum_{k=1}^{T}\widehat{G}_{k-1}y_k + O_p\left(\frac{l}{\sqrt{P}}\right)$, we have that:

$$\mathrm{E}^*\left(\sqrt{P}\left(\frac{1}{P}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\left(\widehat{G}^*_{j-1}y^*_j - \frac{1}{T}\sum_{k=1}^{T}\widehat{G}_{k-1}y_k\right)\right) - \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{G}^*_{t-1}y^*_t - \frac{1}{T}\sum_{k=1}^{T}\widehat{G}_{k-1}y_k\right)\right)\right)$$

$$= O_p\left(\frac{l}{\sqrt{P}}\right) = o_p(1).$$

By the same argument as that used in the proof of part (i), note that:

$$\begin{aligned}
&\text{var}^*\left(\sqrt{P}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{F}^*_{t-1}y^*_t - \frac{1}{T}\sum_{t=1}^{T}\widehat{F}_{t-1}y_t\right)\right)\right) \\
&= \frac{P}{T}\frac{1}{T}\sum_{t=l}^{T-l}\sum_{i=-l}^{l}\left(H_G G_{0,t-1}\epsilon_t\right)\left(H_G G_{0,t-1+i}\epsilon_{t+i}\right)' + O_p\left(\frac{l^2}{P}\right) + o_p(1) \\
&= \frac{P}{T}\frac{1}{T}\sum_{t=l}^{T\tau_\lambda}\sum_{i=-l}^{l}\left(H_G\begin{pmatrix} F_{0,t-1} \\ 0 \end{pmatrix}\epsilon_t\right)\left(H_G\begin{pmatrix} F_{0,t-1+i} \\ 0 \end{pmatrix}\epsilon_{t+i}\right)' \\
&+ \frac{P}{T}\frac{1}{T}\sum_{t=T\tau_\lambda+1}^{T-\lambda}\sum_{i=-l}^{l}\left(H_G\begin{pmatrix} 0 \\ F_{0,t-1} \end{pmatrix}\epsilon_t\right)\left(H_G\begin{pmatrix} 0 \\ F_{0,t-1+i} \end{pmatrix}\epsilon_{t+i}\right)' + O_p\left(\frac{l^2}{P}\right) + o_p(1) \\
&= \frac{\pi}{\pi+1}\left(\tau_\lambda H_G\begin{pmatrix} V_\beta & 0 \\ 0 & 0 \end{pmatrix}H_G' + (1-\tau_\lambda)H_G\begin{pmatrix} 0 & 0 \\ 0 & V_\beta \end{pmatrix}H_G'\right) + o_p(1).
\end{aligned}$$

Now, by the same argument used in part (i), and recalling (8),

$$\begin{aligned}
&\text{var}^*\left(\frac{1}{\sqrt{P}R}\sum_{j=1}^{P-1} j\widehat{G}^*_{j-1}y^*_j\right) \\
&= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right)\frac{1}{T}\sum_{t=1}^{T-l}\left(H_G G_{0,t-1+i}\epsilon_{t+i}\right)\left(H_G G_{0,t-1+i'}\epsilon_{t+i'}\right)' + o_p(1) \\
&= \frac{1}{PR^2}\sum_{k=1}^{b_1}\sum_{i=1}^{l}\sum_{i'=1}^{l}\left((k-1)l+i\right)\left((k-1)l+i'\right)\left(\frac{1}{T}\sum_{t=1}^{T\tau_\lambda-l}\left(H_G\begin{pmatrix} F_{0,t-1+i} \\ 0 \end{pmatrix}\epsilon_{t+i}\right)\left(H_G\begin{pmatrix} F_{0,t-1+i} \\ 0 \end{pmatrix}\epsilon_{t+i}\right)'\right. \\
&\left.\frac{1}{T}\sum_{t=T\tau_\lambda+1}^{T-l}\left(H_G\begin{pmatrix} 0 \\ F_{0,t-1+i} \end{pmatrix}\epsilon_{t+i}\right)\left(H_G\begin{pmatrix} 0 \\ F_{0,t-1+i} \end{pmatrix}\epsilon_{t+i}\right)'\right) + o_p(1) \\
&= \frac{\pi^2}{3}\left(\tau_\lambda H_G\begin{pmatrix} V_\beta & 0 \\ 0 & 0 \end{pmatrix}H_G' + (1-\tau_\lambda)H_G\begin{pmatrix} 0 & 0 \\ 0 & V_\beta \end{pmatrix}H_G'\right).
\end{aligned}$$

It further follows that,

$$\begin{aligned}
&\text{var}^*\left(\frac{1}{\sqrt{P}}\sum_{t=R+1}^{T}\left(\frac{1}{R}\sum_{j=t-R+1}^{t}\widehat{G}^*_{j-1}y^*_j\right)\right) \\
&= 1\left\{\pi \leq 1\right\}\left(\pi - \frac{\pi^2}{3}\right)\left(\tau_\lambda H_G\begin{pmatrix} V_\beta & 0 \\ 0 & 0 \end{pmatrix}H_G' + (1-\tau_\lambda)H_G\begin{pmatrix} 0 & 0 \\ 0 & V_\beta \end{pmatrix}H_G'\right)
\end{aligned}$$

$$+ 1\{\pi > 1\}\left(1 - \frac{1}{3\pi}\right)\left(\tau_\lambda H_G \begin{pmatrix} V_\beta & 0 \\ 0 & 0 \end{pmatrix} H_G' + (1-\tau_\lambda) H_G \begin{pmatrix} 0 & 0 \\ 0 & V_\beta \end{pmatrix} H_G'\right) + o_p(1).$$

Hence, the bootstrap statistic has a well defined limiting distribution, under the alternative. The statement then follows immediately.

8 References

Armah, N.A. and N. R. Swanson, (2010). Seeing Inside the Black Box: Using Diffusion Index Methodology to Construct Factor Proxies in Largescale Macroeconomic Time Series Environments, *Econometric Reviews*, 29, 476-510.

Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions, *Econometrica*, 71, 135-171.

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

Bai, J. and S. Ng (2006). Confidence Intervals for Diffusion Index Forecasts and Inference with Factor Augmented Regressions. *Econometrica*, 74, 1133-1150.

Banerjee, A., M. Marcellino and I. Marsten (2009). Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Changes. forthcoming in *Forecasting in the Presence of Structural Breaks and Model Uncertainty,* edited by D. Rapach and M. Wohar, Elsevier.

Banerjee, A., M. Marcellino and I. Marsten (2010). Forecasting with Factor-Augmented Error Correction Models. Working Paper, European University Institute.

Breitung, J. and S. Eickmeier (2011). Testing for Structural Breaks in Dynamic Factor Models. *Journal of Econometrics,* 163, 71-84.

Castle, J.L., J.A. Doornik and D.F. Hendry (2010). Model Selection when there are Multiple Breaks. Working Paper, University of Oxford.

Castle, J.A., M.P. Clements and D.F. Hendry (2012). Forecasting by Factors, by Variables or Both? *Journal of Econometrics,* forthcoming.

Chen, L., J.J. Dolado and J. Gonzalo (2011). Detecting Big Structural Breaks in Large Factor Models. MPRA W.P. 31344.

Clements, M.P. and D.F. Hendry (2002). Modeling Methodology and Forecast Failure. *Econometrics Journal,* 5, 319-344.

Clements, M.P. and D.F. Hendry (2006). Forecasting with Breaks, in *Handbook of Economic Forecasting,* edited by G. Elliott, C.W.J. Granger and A. Timmermann, North-

Holland Elsevier.

Corradi, V. and N.R. Swanson (2006). Predictive Density and Conditional Confidence Interval Accuracy Tests. *Journal of Econometrics,* 135-228.

Dufour, J.-M. and D. Stevanovic (2011). Factor Augmented VARMA Models: Identification, Estimation, Forecasting and Impulse Responses. Working Paper, McGill University.

Giacomini R. and B. Rossi (2009). Detecting and Predicting Forecast Breakdowns. *Review of Economic Studies,* 76, 669-705.

Goncalves S. and B. Perron (2013). Bootstrapping Factor-Augmented Regression Models. Working Paper, *Journal of Econometrics,* forthcoming.

Goncalves S. and H. White (2004). Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models. *Journal of Econometrics,* 119, 199-219.

Han, X. and A. Inoue (2012). Tests for Parameter Instability in Dynamic Factor Models. Working Paper, North Carolina State University.

Hansen, P.R. and A. Timmermann (2012). Choice of Sample Split in Out-of-Sample Forecast Evaluation. Working Paper, UCSD, Rady School of Management.

Hendry D.F. and M.P. Clements (2002). Pooling of Forecasts. *Econometrics Journal.* 5, 1-26.

Hendry D.F. and G. Mizon (2005). Forecasting in the Presence of Structural Breaks and Policy Regime Shifts in *Identification and Inference for Econometric Models: Essays in Honour of Thomas Rothemberg,* edited by D.W.K. Andrews and J.H. Stock, Cambridge University Press.

Kim H.H. and N.R. Swanson (2012). Forecasting Financial and Macroeconomic Variables Using Data Reduction Methods: New Empirical Evidence. *Journal of Econometrics,* forthcoming.

Kim H.H. and N.R. Swanson (2013). Large Dataset Mining Using Parsimonious Factor and Shrinkage Methods. Working Paper, Rutgers University.

Rossi, B. and A. Inoue (2012). Out of Sample Forecast Tests Robust to Window Size Choice. *Journal of Business and Economic Statistics,* 30, 432-453.

Stock, J.H. and M.W. Watson (2002a). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics,* 20, 147-162.

Stock, J.H. and M.W. Watson (2002b). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association,* 97, 1167-1179.

Stock, J.H. and M.W. Watson (2004). Combination Forecasts of Output Growth in a Seven-Countries Data-Set. *Journal of Forecasting,* 23, 405-430.

Stock, J.H. and M.W. Watson (2009). Forecasting in Dynamic Factor Models Subject to Structural Instability. *The Methodology and Practice of Econometrics: Festschrift in Honour of D.F. Hendry,* edited by J. Castle and N. Shephard. Oxford University Press.

West K.D. and M.W. McCracken (1998). Regression-Based Tests for Predictive Ability. *International Economic Review,* 39, 817-840.

Table 1: Target Forecasting Variables *

Series	Abbreviation	Y_{t+h}
Unemployment Rate	UR	$Z_{t+1}-Z_t$
Personal Income Less transfer payments	PI	$\ln(Z_{t+1}/Z_t)$
10-Year Treasury Bond	TB	$Z_{t+1}-Z_t$
Consumer Price Index	CPI	$\ln(Z_{t+1}/Z_t)$
Producer Price Index	PPI	$\ln(Z_{t+1}/Z_t)$
Nonfarm Payroll Employment	NPE	$\ln(Z_{t+1}/Z_t)$
Housing Starts	HS	$\ln(Z_t)$
Industrial Production	IPX	$\ln(Z_{t+1}/Z_t)$
M2	M2	$\ln(Z_{t+1}/Z_t)$
S&P 500 Index	SNP	$\ln(Z_{t+1}/Z_t)$
Gross Domestic Product	GNP	$\ln(Z_{t+1}/Z_t)$

* Notes: Data used in model estimation and prediction construction are monthly U.S. figures for the period 1962:10-2009:5. Data transformations used in prediction experiments are given in the last column of the table.

Table 2: Factor Test Statistic and Its Critical Values for $h = 1$[§]

Panel A. Block length, $l = 2$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.7253	0.6456	0.6683	0.5703	0.0242	0.0220	1.0412	1.6649	0.0397	0.0976	0.0781
	95%	1.0599	0.0244	1.1617	0.0233	0.0285	0.0172	1.8973	0.0437	0.0237	0.1344	0.0214
	90%	0.9376	0.0216	0.9921	0.0183	0.0255	0.0159	1.8062	0.0403	0.0186	0.1160	0.0159
	50%	0.7584	0.0153	0.7195	0.0141	0.0177	0.0127	1.5093	0.0324	0.0087	0.0707	0.0101
	p-value	0.5867	0.0000	0.6533	0.0000	0.1300	0.0100	1.0000	0.0000	0.0033	0.1833	0.0000
$R = 240$	Statistics	0.6126	0.5249	0.6159	0.4243	0.0197	0.0187	0.8780	1.3643	0.0330	0.0873	0.0678
	95%	1.0071	0.0201	1.0399	0.0176	0.0256	0.0149	1.7067	0.0418	0.0178	0.1017	0.0160
	90%	0.8919	0.0185	0.8936	0.0159	0.0215	0.0141	1.6376	0.0383	0.0149	0.0905	0.0138
	50%	0.7038	0.0147	0.6660	0.0132	0.0163	0.0120	1.3904	0.0300	0.0079	0.0659	0.0094
	p-value	0.7533	0.0000	0.6700	0.0000	0.2100	0.0000	1.0000	0.0000	0.0067	0.1300	0.0000
$R = 280$	Statistics	0.5338	0.4708	0.5710	0.3550	0.0172	0.0165	0.7689	1.2119	0.0297	0.0805	0.0622
	95%	0.8502	0.0174	0.8882	0.0167	0.0223	0.0140	1.6409	0.0373	0.0156	0.1012	0.0152
	90%	0.7959	0.0163	0.7896	0.0151	0.0201	0.0130	1.5652	0.0341	0.0133	0.0882	0.0125
	50%	0.6444	0.0137	0.6098	0.0124	0.0151	0.0110	1.2998	0.0280	0.0070	0.0625	0.0088
	p-value	0.8467	0.0000	0.5967	0.0000	0.2600	0.0000	1.0000	0.0000	0.0033	0.1500	0.0000
$R = 320$	Statistics	0.4945	0.4439	0.5531	0.3189	0.0159	0.0154	0.7207	1.1344	0.0280	0.0774	0.0602
	95%	0.8054	0.0158	0.8242	0.0151	0.0198	0.0124	1.4874	0.0340	0.0121	0.0825	0.0122
	90%	0.7486	0.0147	0.7316	0.0137	0.0175	0.0119	1.4137	0.0323	0.0107	0.0752	0.0104
	50%	0.6025	0.0120	0.5635	0.0113	0.0138	0.0102	1.1795	0.0259	0.0064	0.0558	0.0080
	p-value	0.8700	0.0000	0.5433	0.0000	0.2200	0.0033	1.0000	0.0000	0.0000	0.0833	0.0000
$R = 360$	Statistics	0.4837	0.4402	0.5604	0.3043	0.0156	0.0151	0.7169	1.1217	0.0278	0.0774	0.0604
	95%	0.7119	0.0142	0.6910	0.0135	0.0167	0.0116	1.3497	0.0299	0.0092	0.0733	0.0102
	90%	0.6471	0.0136	0.6550	0.0125	0.0158	0.0110	1.3124	0.0287	0.0079	0.0677	0.0095
	50%	0.5409	0.0109	0.5139	0.0104	0.0125	0.0092	1.1028	0.0232	0.0056	0.0492	0.0073
	p-value	0.7167	0.0000	0.3133	0.0000	0.1133	0.0000	1.0000	0.0000	0.0000	0.0300	0.0000
$R = 400$	Statistics	0.4977	0.4587	0.5918	0.3065	0.0160	0.0156	0.7543	1.1780	0.0288	0.0806	0.0629
	95%	0.6389	0.0132	0.6243	0.0116	0.0148	0.0104	1.2226	0.0274	0.0080	0.0649	0.0092
	90%	0.5981	0.0124	0.5786	0.0109	0.0140	0.0099	1.1708	0.0257	0.0071	0.0594	0.0084
	50%	0.4978	0.0100	0.4590	0.0092	0.0110	0.0084	0.9823	0.0207	0.0049	0.0450	0.0065
	p-value	0.5033	0.0000	0.0967	0.0000	0.0200	0.0000	0.9833	0.0000	0.0000	0.0100	0.0000

Panel B. Block length, $l = 5$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.7253	0.6456	0.6683	0.5703	0.0242	0.0220	1.0412	1.6649	0.0397	0.0976	0.0781
	95%	1.2480	0.0264	1.5676	0.0289	0.0382	0.0191	2.0679	0.0503	0.0341	0.1730	0.0278
	90%	1.0600	0.0224	1.1566	0.0234	0.0302	0.0166	1.9805	0.0463	0.0246	0.1306	0.0221
	50%	0.7763	0.0159	0.7480	0.0151	0.0187	0.0128	1.5063	0.0320	0.0101	0.0730	0.0108
	p-value	0.6233	0.0000	0.6833	0.0000	0.2067	0.0233	0.9800	0.0000	0.0133	0.2167	0.0000
$R = 240$	Statistics	0.6126	0.5249	0.6159	0.4243	0.0197	0.0187	0.8780	1.3643	0.0330	0.0873	0.0678
	95%	1.1521	0.0257	1.3014	0.0232	0.0321	0.0169	1.9756	0.0494	0.0255	0.1574	0.0219
	90%	0.9786	0.0221	1.0236	0.0193	0.0252	0.0148	1.8837	0.0428	0.0200	0.1247	0.0176
	50%	0.7030	0.0143	0.6839	0.0138	0.0174	0.0119	1.4294	0.0310	0.0091	0.0689	0.0102
	p-value	0.7300	0.0000	0.6633	0.0000	0.2767	0.0233	0.9900	0.0000	0.0200	0.2467	0.0000
$R = 280$	Statistics	0.5338	0.4708	0.5710	0.3550	0.0172	0.0165	0.7689	1.2119	0.0297	0.0805	0.0622
	95%	0.9557	0.0209	1.0626	0.0240	0.0269	0.0149	1.7663	0.0413	0.0218	0.1047	0.0175
	90%	0.8858	0.0186	0.8928	0.0186	0.0223	0.0142	1.6637	0.0381	0.0172	0.0855	0.0144
	50%	0.6719	0.0136	0.6177	0.0129	0.0155	0.0114	1.3262	0.0280	0.0079	0.0606	0.0091
	p-value	0.8000	0.0000	0.6367	0.0000	0.3500	0.0067	0.9967	0.0000	0.0167	0.1200	0.0000
$R = 320$	Statistics	0.4945	0.4439	0.5531	0.3189	0.0159	0.0154	0.7207	1.1344	0.0280	0.0774	0.0602
	95%	0.8021	0.0191	0.9591	0.0163	0.0219	0.0135	1.6000	0.0407	0.0168	0.0952	0.0149
	90%	0.7580	0.0166	0.8022	0.0154	0.0194	0.0129	1.4979	0.0360	0.0131	0.0778	0.0126
	50%	0.6200	0.0121	0.5772	0.0118	0.0145	0.0105	1.2047	0.0260	0.0071	0.0548	0.0084
	p-value	0.8400	0.0000	0.5500	0.0000	0.3800	0.0133	1.0000	0.0000	0.0100	0.1033	0.0000
$R = 360$	Statistics	0.4837	0.4402	0.5604	0.3043	0.0156	0.0151	0.7169	1.1217	0.0278	0.0774	0.0604
	95%	0.7901	0.0166	0.7886	0.0157	0.0204	0.0124	1.4903	0.0340	0.0138	0.0831	0.0118
	90%	0.7330	0.0146	0.6898	0.0139	0.0187	0.0118	1.3749	0.0309	0.0108	0.0730	0.0104
	50%	0.5551	0.0110	0.5100	0.0106	0.0128	0.0093	1.0679	0.0233	0.0061	0.0510	0.0077
	p-value	0.7067	0.0000	0.3367	0.0000	0.2333	0.0033	0.9967	0.0000	0.0000	0.0767	0.0000
$R = 400$	Statistics	0.4977	0.4587	0.5918	0.3065	0.0160	0.0156	0.7543	1.1780	0.0288	0.0806	0.0629
	95%	0.6686	0.0137	0.6537	0.0131	0.0177	0.0109	1.3064	0.0308	0.0104	0.0672	0.0100
	90%	0.6260	0.0130	0.5902	0.0120	0.0163	0.0103	1.2552	0.0279	0.0083	0.0587	0.0088
	50%	0.4909	0.0100	0.4639	0.0093	0.0113	0.0082	0.9924	0.0208	0.0053	0.0451	0.0066
	p-value	0.4800	0.0000	0.0967	0.0000	0.1100	0.0000	0.9700	0.0000	0.0000	0.0333	0.0000

Panel C. Block length, $l = 10$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.7253	0.6456	0.6683	0.5703	0.0242	0.0220	1.0412	1.6649	0.0397	0.0976	0.0781
	95%	1.2776	0.0326	1.7352	0.0320	0.0459	0.0209	2.2660	0.0662	0.0399	0.2211	0.0367
	90%	1.1570	0.0244	1.4092	0.0253	0.0326	0.0192	2.0950	0.0500	0.0324	0.1862	0.0252
	50%	0.7968	0.0162	0.7798	0.0154	0.0189	0.0132	1.5115	0.0343	0.0115	0.0763	0.0114
	p-value	0.6367	0.0000	0.7233	0.0000	0.2667	0.0400	0.9167	0.0000	0.0500	0.3000	0.0033
$R = 240$	Statistics	0.6126	0.5249	0.6159	0.4243	0.0197	0.0187	0.8780	1.3643	0.0330	0.0873	0.0678
	95%	1.4478	0.0270	1.6100	0.0249	0.0362	0.0181	2.1170	0.0482	0.0325	0.1839	0.0294
	90%	1.0965	0.0210	1.2684	0.0192	0.0278	0.0162	1.8663	0.0440	0.0250	0.1364	0.0215
	50%	0.7442	0.0149	0.7206	0.0137	0.0174	0.0124	1.4030	0.0319	0.0109	0.0690	0.0101
	p-value	0.7633	0.0000	0.7033	0.0000	0.3467	0.0400	0.9533	0.0000	0.0433	0.2767	0.0067
$R = 280$	Statistics	0.5338	0.4708	0.5710	0.3550	0.0172	0.0165	0.7689	1.2119	0.0297	0.0805	0.0622
	95%	1.0591	0.0235	1.3153	0.0207	0.0321	0.0167	2.0439	0.0465	0.0231	0.1474	0.0231
	90%	0.9523	0.0202	1.0145	0.0176	0.0256	0.0150	1.8224	0.0406	0.0199	0.1104	0.0165
	50%	0.6990	0.0138	0.6593	0.0130	0.0160	0.0115	1.2710	0.0296	0.0089	0.0625	0.0095
	p-value	0.8067	0.0000	0.7200	0.0000	0.4000	0.0533	0.9900	0.0000	0.0300	0.2000	0.0000
$R = 320$	Statistics	0.4945	0.4439	0.5531	0.3189	0.0159	0.0154	0.7207	1.1344	0.0280	0.0774	0.0602
	95%	0.8998	0.0184	1.0600	0.0174	0.0249	0.0141	1.7473	0.0396	0.0228	0.1186	0.0208
	90%	0.8550	0.0166	0.8624	0.0158	0.0219	0.0135	1.6330	0.0368	0.0166	0.0894	0.0153
	50%	0.5940	0.0127	0.5821	0.0117	0.0144	0.0104	1.2221	0.0264	0.0075	0.0568	0.0086
	p-value	0.7300	0.0000	0.5867	0.0000	0.3367	0.0133	0.9833	0.0000	0.0167	0.1667	0.0000
$R = 360$	Statistics	0.4837	0.4402	0.5604	0.3043	0.0156	0.0151	0.7169	1.1217	0.0278	0.0774	0.0604
	95%	0.8491	0.0162	0.9731	0.0145	0.0204	0.0134	1.5684	0.0359	0.0141	0.0849	0.0140
	90%	0.7682	0.0148	0.7776	0.0136	0.0173	0.0124	1.4256	0.0332	0.0122	0.0710	0.0120
	50%	0.5474	0.0114	0.5388	0.0104	0.0131	0.0098	1.1005	0.0243	0.0062	0.0513	0.0075
	p-value	0.6667	0.0000	0.4400	0.0000	0.2333	0.0100	0.9700	0.0000	0.0033	0.0733	0.0000
$R = 400$	Statistics	0.4977	0.4587	0.5918	0.3065	0.0160	0.0156	0.7543	1.1780	0.0288	0.0806	0.0629
	95%	0.7108	0.0139	0.8282	0.0142	0.0166	0.0116	1.4185	0.0294	0.0126	0.0766	0.0109
	90%	0.6601	0.0131	0.6546	0.0126	0.0153	0.0107	1.2934	0.0277	0.0095	0.0634	0.0098
	50%	0.4966	0.0100	0.4790	0.0094	0.0112	0.0081	0.9511	0.0212	0.0056	0.0448	0.0066
	p-value	0.4867	0.0000	0.1467	0.0000	0.0567	0.0033	0.8500	0.0000	0.0000	0.0467	0.0000

§ Notes: See notes to Table 1. Entries denoted "Statistics" are of the proposed test statistic, $\|Z_{P,R}\|$, and are reported for the 11 macroeconomic variables summarized in Table 1. Prediction models are assumed to be constructed according to the generic specification given in equations (1) and (2). In total, 155 monthly variables for the period 1962:10 - 2009:5 are used in factor construction, so that $N = 155$ and $T = 560$. In implementation of the bootstrap, critical values based on block lengths of $bT = l =$ 2 (Panel A), 5 (Panel B), and 10 (Panel C) are reported, for $95th$, $90th$, and $50th$ percentiles of the bootstrap empirical distribution (with $B = 300$). Additionally, standard p-values based on the empirical bootstrap distribution are reported. Finally, note that a variety of different values of R (and hence P, given that $T = R + P$) are reported on, including $R = 200$, 240, 280, 320, 360, 400 , corresponding to values of P/R including 9/5, 4/3, 1, 3/4, 5/9, 2/5 . See Section 5 for further details.

Table 3: Factor Test Statistic and Its Critical Values for $h = 3$[§]

Panel A. Block length, $l = 2$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.5913	0.6374	0.3155	0.2362	0.0204	0.0169	1.0157	1.5315	0.0314	0.0703	0.0532
	95%	0.9264	0.0214	1.2640	0.0176	0.0262	0.0160	1.8351	0.0402	0.0235	0.1120	0.0191
	90%	0.8098	0.0193	0.7547	0.0153	0.0209	0.0145	1.7060	0.0364	0.0186	0.0867	0.0142
	50%	0.6189	0.0134	0.3789	0.0114	0.0149	0.0118	1.4038	0.0270	0.0090	0.0436	0.0081
	p-value	0.5800	0.0000	0.7100	0.0000	0.1100	0.0333	1.0000	0.0000	0.0300	0.1533	0.0000
$R = 240$	Statistics	0.5079	0.5233	0.2801	0.1871	0.0168	0.0148	0.8728	1.2602	0.0262	0.0598	0.0467
	95%	0.7614	0.0183	0.7808	0.0148	0.0228	0.0138	1.6339	0.0346	0.0178	0.0943	0.0156
	90%	0.7188	0.0162	0.5612	0.0135	0.0185	0.0130	1.5356	0.0318	0.0141	0.0666	0.0129
	50%	0.5778	0.0127	0.3353	0.0105	0.0140	0.0111	1.2832	0.0250	0.0077	0.0385	0.0073
	p-value	0.7667	0.0000	0.7100	0.0000	0.1800	0.0333	0.9967	0.0000	0.0167	0.1467	0.0000
$R = 280$	Statistics	0.4527	0.4689	0.2538	0.1682	0.0147	0.0132	0.7809	1.1224	0.0236	0.0528	0.0421
	95%	0.6918	0.0162	0.7812	0.0140	0.0191	0.0129	1.4953	0.0297	0.0160	0.0765	0.0127
	90%	0.6504	0.0149	0.5297	0.0121	0.0173	0.0123	1.4268	0.0283	0.0123	0.0576	0.0098
	50%	0.5294	0.0115	0.3122	0.0099	0.0129	0.0105	1.2066	0.0230	0.0072	0.0343	0.0067
	p-value	0.8100	0.0000	0.7133	0.0000	0.2167	0.0367	1.0000	0.0000	0.0133	0.1167	0.0000
$R = 320$	Statistics	0.4260	0.4429	0.2452	0.1557	0.0137	0.0123	0.7448	1.0537	0.0222	0.0491	0.0400
	95%	0.6720	0.0148	0.5967	0.0122	0.0165	0.0121	1.3922	0.0281	0.0118	0.0588	0.0100
	90%	0.6127	0.0134	0.4441	0.0111	0.0151	0.0111	1.3166	0.0267	0.0101	0.0533	0.0089
	50%	0.4906	0.0103	0.2839	0.0089	0.0118	0.0094	1.1060	0.0219	0.0064	0.0325	0.0061
	p-value	0.7933	0.0000	0.6833	0.0000	0.1967	0.0333	1.0000	0.0000	0.0033	0.1467	0.0000
$R = 360$	Statistics	0.4230	0.4391	0.2491	0.1529	0.0136	0.0121	0.7519	1.0441	0.0219	0.0478	0.0395
	95%	0.5531	0.0127	0.4828	0.0106	0.0141	0.0106	1.2606	0.0249	0.0090	0.0600	0.0085
	90%	0.5348	0.0119	0.3854	0.0101	0.0132	0.0102	1.1983	0.0238	0.0076	0.0470	0.0077
	50%	0.4443	0.0096	0.2594	0.0082	0.0107	0.0086	1.0256	0.0197	0.0054	0.0286	0.0055
	p-value	0.6500	0.0000	0.5500	0.0000	0.0767	0.0033	0.9967	0.0000	0.0000	0.0933	0.0000
$R = 400$	Statistics	0.4408	0.4558	0.2601	0.1580	0.0141	0.0125	0.7999	1.0995	0.0226	0.0482	0.0403
	95%	0.5104	0.0118	0.4919	0.0095	0.0128	0.0094	1.1226	0.0229	0.0072	0.0459	0.0078
	90%	0.4831	0.0110	0.3383	0.0088	0.0120	0.0089	1.0866	0.0212	0.0066	0.0379	0.0069
	50%	0.4048	0.0086	0.2284	0.0074	0.0095	0.0078	0.9036	0.0175	0.0049	0.0250	0.0049
	p-value	0.2500	0.0000	0.3533	0.0000	0.0200	0.0000	0.8167	0.0000	0.0000	0.0433	0.0000

Panel B. Block length, $l = 5$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.5913	0.6374	0.3155	0.2362	0.0204	0.0169	1.0157	1.5315	0.0314	0.0703	0.0532
	95%	1.2046	0.0265	1.2989	0.0290	0.0342	0.0181	1.9957	0.0455	0.0307	0.1552	0.0230
	90%	0.9429	0.0215	1.0159	0.0229	0.0279	0.0155	1.8502	0.0394	0.0250	0.1094	0.0178
	50%	0.6507	0.0137	0.3970	0.0126	0.0160	0.0118	1.3987	0.0267	0.0099	0.0444	0.0085
	p-value	0.6400	0.0000	0.7300	0.0000	0.2100	0.0700	0.9533	0.0000	0.0433	0.1900	0.0033
$R = 240$	Statistics	0.5079	0.5233	0.2801	0.1871	0.0168	0.0148	0.8728	1.2602	0.0262	0.0598	0.0467
	95%	1.0432	0.0228	1.1510	0.0207	0.0287	0.0149	1.8692	0.0436	0.0235	0.1324	0.0190
	90%	0.8765	0.0191	0.7879	0.0173	0.0233	0.0140	1.7493	0.0386	0.0181	0.1116	0.0151
	50%	0.5837	0.0124	0.3582	0.0114	0.0147	0.0112	1.3131	0.0262	0.0090	0.0425	0.0079
	p-value	0.7267	0.0000	0.7367	0.0000	0.2933	0.0633	0.9900	0.0000	0.0267	0.2700	0.0100
$R = 280$	Statistics	0.4527	0.4689	0.2538	0.1682	0.0147	0.0132	0.7809	1.1224	0.0236	0.0528	0.0421
	95%	0.8284	0.0191	0.8136	0.0164	0.0209	0.0140	1.6536	0.0358	0.0208	0.0828	0.0183
	90%	0.7496	0.0169	0.6066	0.0144	0.0186	0.0130	1.5454	0.0324	0.0157	0.0634	0.0129
	50%	0.5454	0.0117	0.3316	0.0099	0.0133	0.0106	1.2294	0.0238	0.0076	0.0347	0.0072
	p-value	0.7667	0.0000	0.7733	0.0000	0.3133	0.0867	0.9867	0.0000	0.0367	0.1700	0.0000
$R = 320$	Statistics	0.4260	0.4429	0.2452	0.1557	0.0137	0.0123	0.7448	1.0537	0.0222	0.0491	0.0400
	95%	0.6954	0.0169	0.7038	0.0142	0.0188	0.0127	1.4520	0.0342	0.0152	0.0669	0.0130
	90%	0.6365	0.0150	0.5439	0.0129	0.0165	0.0120	1.4070	0.0301	0.0124	0.0535	0.0106
	50%	0.5088	0.0107	0.2871	0.0093	0.0124	0.0098	1.1050	0.0221	0.0071	0.0319	0.0064
	p-value	0.8000	0.0000	0.6967	0.0000	0.3033	0.0767	0.9933	0.0000	0.0167	0.1300	0.0000
$R = 360$	Statistics	0.4230	0.4391	0.2491	0.1529	0.0136	0.0121	0.7519	1.0441	0.0219	0.0478	0.0395
	95%	0.6380	0.0149	0.5801	0.0122	0.0186	0.0116	1.3701	0.0294	0.0135	0.0717	0.0108
	90%	0.5978	0.0134	0.4730	0.0109	0.0151	0.0109	1.2589	0.0263	0.0112	0.0529	0.0085
	50%	0.4509	0.0097	0.2658	0.0083	0.0109	0.0086	1.0001	0.0198	0.0063	0.0284	0.0058
	p-value	0.6133	0.0000	0.5967	0.0000	0.2067	0.0333	0.9667	0.0000	0.0067	0.1400	0.0000
$R = 400$	Statistics	0.4408	0.4558	0.2601	0.1580	0.0141	0.0125	0.7999	1.0995	0.0226	0.0482	0.0403
	95%	0.5644	0.0123	0.4661	0.0104	0.0146	0.0101	1.2230	0.0255	0.0101	0.0448	0.0089
	90%	0.5103	0.0112	0.3728	0.0095	0.0133	0.0096	1.1671	0.0237	0.0083	0.0406	0.0074
	50%	0.4042	0.0087	0.2328	0.0074	0.0098	0.0077	0.9143	0.0178	0.0056	0.0253	0.0050
	p-value	0.3233	0.0000	0.3533	0.0000	0.0667	0.0033	0.7800	0.0000	0.0000	0.0367	0.0000

Panel C. Block length, $l = 10$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.5913	0.6374	0.3155	0.2362	0.0204	0.0169	1.0157	1.5315	0.0314	0.0703	0.0532
	95%	1.1870	0.0297	1.7113	0.0295	0.0366	0.0203	2.1858	0.0635	0.0431	0.2251	0.0325
	90%	0.9800	0.0229	1.1700	0.0211	0.0276	0.0174	1.9497	0.0443	0.0327	0.1595	0.0256
	50%	0.6504	0.0138	0.4067	0.0126	0.0167	0.0122	1.4181	0.0284	0.0121	0.0492	0.0091
	p-value	0.6367	0.0000	0.7467	0.0000	0.2567	0.1300	0.8867	0.0000	0.1033	0.3133	0.0067
$R = 240$	Statistics	0.5079	0.5233	0.2801	0.1871	0.0168	0.0148	0.8728	1.2602	0.0262	0.0598	0.0467
	95%	1.2164	0.0255	1.3869	0.0198	0.0340	0.0183	1.9305	0.0407	0.0309	0.1901	0.0254
	90%	0.9361	0.0200	1.1035	0.0163	0.0262	0.0158	1.7448	0.0365	0.0252	0.1217	0.0201
	50%	0.6145	0.0132	0.4006	0.0110	0.0154	0.0115	1.2772	0.0266	0.0102	0.0430	0.0083
	p-value	0.7333	0.0000	0.7867	0.0000	0.3933	0.1533	0.9433	0.0000	0.0867	0.2900	0.0033
$R = 280$	Statistics	0.4527	0.4689	0.2538	0.1682	0.0147	0.0132	0.7809	1.1224	0.0236	0.0528	0.0421
	95%	0.9139	0.0203	1.2521	0.0193	0.0286	0.0157	1.8313	0.0422	0.0247	0.1156	0.0208
	90%	0.8212	0.0169	0.7706	0.0156	0.0231	0.0139	1.6795	0.0374	0.0197	0.0797	0.0162
	50%	0.5613	0.0117	0.3451	0.0105	0.0140	0.0106	1.1992	0.0241	0.0090	0.0384	0.0073
	p-value	0.7400	0.0000	0.7933	0.0000	0.4267	0.1567	0.9567	0.0000	0.0600	0.2700	0.0033
$R = 320$	Statistics	0.4260	0.4429	0.2452	0.1557	0.0137	0.0123	0.7448	1.0537	0.0222	0.0491	0.0400
	95%	0.8373	0.0183	0.9350	0.0143	0.0244	0.0137	1.6510	0.0340	0.0212	0.1103	0.0188
	90%	0.7387	0.0151	0.6764	0.0131	0.0181	0.0125	1.5642	0.0302	0.0167	0.0713	0.0144
	50%	0.5034	0.0112	0.3031	0.0094	0.0122	0.0098	1.1265	0.0218	0.0077	0.0344	0.0070
	p-value	0.6800	0.0000	0.7533	0.0000	0.3233	0.1267	0.9567	0.0000	0.0433	0.2167	0.0033
$R = 360$	Statistics	0.4230	0.4391	0.2491	0.1529	0.0136	0.0121	0.7519	1.0441	0.0219	0.0478	0.0395
	95%	0.7371	0.0146	0.7124	0.0129	0.0170	0.0124	1.4830	0.0321	0.0143	0.0656	0.0138
	90%	0.6712	0.0135	0.5569	0.0115	0.0153	0.0115	1.3245	0.0287	0.0116	0.0524	0.0102
	50%	0.4510	0.0098	0.2616	0.0084	0.0112	0.0090	1.0297	0.0204	0.0063	0.0293	0.0057
	p-value	0.5833	0.0000	0.5433	0.0000	0.2067	0.0600	0.8900	0.0000	0.0033	0.1300	0.0000
$R = 400$	Statistics	0.4408	0.4558	0.2601	0.1580	0.0141	0.0125	0.7999	1.0995	0.0226	0.0482	0.0403
	95%	0.6338	0.0123	0.5667	0.0109	0.0154	0.0106	1.3362	0.0259	0.0110	0.0537	0.0107
	90%	0.5769	0.0116	0.4202	0.0100	0.0133	0.0099	1.2179	0.0241	0.0093	0.0427	0.0089
	50%	0.4108	0.0084	0.2294	0.0074	0.0100	0.0077	0.8748	0.0175	0.0058	0.0246	0.0054
	p-value	0.3633	0.0000	0.3567	0.0000	0.0767	0.0033	0.6333	0.0000	0.0000	0.0667	0.0000

§ Notes: See notes to Table 2.

Table 4: Factor Test Statistic and Its Critical Values for $h = 12$[§]

Panel A. Block length, $l = 2$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.3874	0.1918	0.1870	0.1424	0.0183	0.0135	0.6727	0.7601	0.0156	0.0329	0.0337
	95%	0.6504	0.0196	0.8830	0.0188	0.0250	0.0094	1.2573	0.0294	0.0224	0.1098	0.0178
	90%	0.5279	0.0153	0.6580	0.0152	0.0194	0.0084	1.1279	0.0222	0.0170	0.0774	0.0139
	50%	0.3313	0.0076	0.3127	0.0095	0.0118	0.0061	0.8661	0.0141	0.0077	0.0382	0.0054
	p-value	0.2600	0.0000	0.9400	0.0000	0.1100	0.0033	0.9667	0.0000	0.1067	0.6733	0.0033
$R = 240$	Statistics	0.3415	0.1565	0.1586	0.1004	0.0151	0.0118	0.6042	0.5730	0.0137	0.0292	0.0285
	95%	0.5208	0.0155	0.5901	0.0140	0.0188	0.0088	1.0108	0.0231	0.0186	0.1011	0.0145
	90%	0.4173	0.0118	0.4601	0.0123	0.0160	0.0076	0.9603	0.0192	0.0145	0.0703	0.0100
	50%	0.2987	0.0072	0.2973	0.0087	0.0108	0.0058	0.7972	0.0132	0.0071	0.0361	0.0050
	p-value	0.2933	0.0000	0.9800	0.0000	0.1167	0.0133	0.9900	0.0000	0.1200	0.7133	0.0067
$R = 280$	Statistics	0.3100	0.1405	0.1439	0.0755	0.0132	0.0101	0.5575	0.4661	0.0122	0.0261	0.0246
	95%	0.4747	0.0114	0.5831	0.0126	0.0157	0.0075	0.9219	0.0192	0.0141	0.0644	0.0106
	90%	0.4015	0.0097	0.4676	0.0106	0.0135	0.0069	0.8758	0.0170	0.0118	0.0522	0.0086
	50%	0.2886	0.0065	0.2684	0.0082	0.0100	0.0053	0.7416	0.0118	0.0065	0.0309	0.0045
	p-value	0.3700	0.0000	0.9767	0.0000	0.1200	0.0133	0.9867	0.0000	0.0967	0.7033	0.0067
$R = 320$	Statistics	0.2945	0.1320	0.1431	0.0555	0.0123	0.0091	0.5448	0.4009	0.0112	0.0248	0.0224
	95%	0.3937	0.0099	0.4696	0.0109	0.0130	0.0068	0.8250	0.0169	0.0104	0.0481	0.0089
	90%	0.3498	0.0082	0.3926	0.0096	0.0118	0.0062	0.7891	0.0150	0.0085	0.0418	0.0070
	50%	0.2624	0.0059	0.2363	0.0076	0.0093	0.0049	0.6808	0.0110	0.0056	0.0287	0.0040
	p-value	0.2967	0.0000	0.9700	0.0000	0.0700	0.0033	0.9667	0.0000	0.0400	0.6967	0.0067
$R = 360$	Statistics	0.2947	0.1316	0.1514	0.0453	0.0120	0.0085	0.5632	0.3675	0.0108	0.0246	0.0213
	95%	0.3260	0.0083	0.3590	0.0089	0.0118	0.0059	0.7574	0.0143	0.0084	0.0478	0.0080
	90%	0.3009	0.0074	0.3132	0.0082	0.0108	0.0056	0.7217	0.0130	0.0075	0.0384	0.0062
	50%	0.2264	0.0055	0.2212	0.0067	0.0084	0.0045	0.6211	0.0098	0.0049	0.0254	0.0036
	p-value	0.1100	0.0000	0.9100	0.0000	0.0433	0.0033	0.8000	0.0000	0.0133	0.5500	0.0000
$R = 400$	Statistics	0.3082	0.1379	0.1644	0.0473	0.0123	0.0083	0.6138	0.3726	0.0107	0.0252	0.0211
	95%	0.3005	0.0066	0.3241	0.0078	0.0104	0.0054	0.6753	0.0124	0.0076	0.0388	0.0060
	90%	0.2732	0.0059	0.2823	0.0074	0.0096	0.0050	0.6429	0.0118	0.0066	0.0319	0.0049
	50%	0.2072	0.0047	0.1996	0.0061	0.0076	0.0040	0.5566	0.0086	0.0044	0.0231	0.0033
	p-value	0.0233	0.0000	0.7600	0.0000	0.0167	0.0033	0.1700	0.0000	0.0100	0.3633	0.0000

Panel B. Block length, $l = 5$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.3874	0.1918	0.1870	0.1424	0.0183	0.0135	0.6727	0.7601	0.0156	0.0329	0.0337
	95%	0.8698	0.0255	1.1263	0.0257	0.0319	0.0121	1.3203	0.0396	0.0288	0.1634	0.0220
	90%	0.7128	0.0176	0.8417	0.0192	0.0231	0.0098	1.1671	0.0279	0.0235	0.0937	0.0173
	50%	0.3673	0.0082	0.3212	0.0101	0.0128	0.0066	0.8691	0.0143	0.0100	0.0389	0.0061
	p-value	0.4500	0.0000	0.9600	0.0000	0.1700	0.0400	0.9233	0.0000	0.2367	0.6600	0.0133
$R = 240$	Statistics	0.3415	0.1565	0.1586	0.1004	0.0151	0.0118	0.6042	0.5730	0.0137	0.0292	0.0285
	95%	0.6825	0.0181	0.9744	0.0178	0.0220	0.0104	1.2654	0.0277	0.0221	0.1048	0.0217
	90%	0.5307	0.0134	0.6955	0.0137	0.0183	0.0090	1.0550	0.0238	0.0167	0.0854	0.0133
	50%	0.3271	0.0073	0.2956	0.0092	0.0116	0.0061	0.7956	0.0140	0.0076	0.0381	0.0054
	p-value	0.4500	0.0000	0.9800	0.0000	0.2100	0.0367	0.9400	0.0000	0.1900	0.7800	0.0200
$R = 280$	Statistics	0.3100	0.1405	0.1439	0.0755	0.0132	0.0101	0.5575	0.4661	0.0122	0.0261	0.0246
	95%	0.5803	0.0149	0.7776	0.0165	0.0197	0.0085	1.0536	0.0250	0.0164	0.0953	0.0139
	90%	0.4583	0.0116	0.5323	0.0126	0.0170	0.0076	0.9536	0.0193	0.0146	0.0731	0.0105
	50%	0.3059	0.0070	0.2579	0.0083	0.0108	0.0055	0.7555	0.0129	0.0071	0.0315	0.0048
	p-value	0.4733	0.0000	0.9733	0.0000	0.2500	0.0200	0.9500	0.0000	0.1733	0.7467	0.0133
$R = 320$	Statistics	0.2945	0.1320	0.1431	0.0555	0.0123	0.0091	0.5448	0.4009	0.0112	0.0248	0.0224
	95%	0.4505	0.0114	0.7214	0.0111	0.0173	0.0079	0.9822	0.0221	0.0143	0.0643	0.0130
	90%	0.4137	0.0092	0.4960	0.0098	0.0136	0.0072	0.9037	0.0179	0.0112	0.0506	0.0101
	50%	0.2715	0.0062	0.2433	0.0076	0.0096	0.0051	0.7023	0.0113	0.0063	0.0285	0.0043
	p-value	0.3667	0.0000	0.9300	0.0000	0.1733	0.0167	0.9267	0.0000	0.1000	0.6967	0.0033
$R = 360$	Statistics	0.2947	0.1316	0.1514	0.0453	0.0120	0.0085	0.5632	0.3675	0.0108	0.0246	0.0213
	95%	0.3932	0.0107	0.5485	0.0116	0.0158	0.0066	0.8537	0.0185	0.0117	0.0699	0.0085
	90%	0.3495	0.0081	0.3928	0.0095	0.0128	0.0060	0.8038	0.0151	0.0095	0.0485	0.0070
	50%	0.2366	0.0057	0.2197	0.0069	0.0086	0.0044	0.6378	0.0105	0.0055	0.0251	0.0040
	p-value	0.2300	0.0000	0.8533	0.0000	0.1233	0.0000	0.7633	0.0000	0.0667	0.5233	0.0000
$R = 400$	Statistics	0.3082	0.1379	0.1644	0.0473	0.0123	0.0083	0.6138	0.3726	0.0107	0.0252	0.0211
	95%	0.3364	0.0084	0.3398	0.0097	0.0114	0.0053	0.7388	0.0146	0.0090	0.0399	0.0071
	90%	0.3005	0.0068	0.2794	0.0084	0.0103	0.0051	0.6845	0.0129	0.0077	0.0325	0.0055
	50%	0.2155	0.0049	0.1892	0.0061	0.0074	0.0041	0.5656	0.0089	0.0050	0.0218	0.0033
	p-value	0.0833	0.0000	0.6800	0.0000	0.0400	0.0000	0.2900	0.0000	0.0200	0.3333	0.0000

Panel C. Block length, $l = 10$

		UR	PI	TB10Y	CPI	PPI	NPE	HS	IPX	M2	SNP	GDP
$R = 200$	Statistics	0.3874	0.1918	0.1870	0.1424	0.0183	0.0135	0.6727	0.7601	0.0156	0.0329	0.0337
	95%	1.0285	0.0245	1.9698	0.0254	0.0321	0.0158	1.4809	0.0392	0.0366	0.1920	0.0260
	90%	0.7968	0.0162	1.2291	0.0207	0.0251	0.0118	1.2832	0.0314	0.0276	0.1281	0.0194
	50%	0.4050	0.0081	0.3357	0.0106	0.0134	0.0068	0.8663	0.0153	0.0102	0.0428	0.0063
	p-value	0.5800	0.0000	0.9400	0.0000	0.2400	0.0800	0.8367	0.0000	0.2800	0.7467	0.0333
$R = 240$	Statistics	0.3415	0.1565	0.1586	0.1004	0.0151	0.0118	0.6042	0.5730	0.0137	0.0292	0.0285
	95%	0.8109	0.0191	1.1394	0.0194	0.0250	0.0114	1.2251	0.0410	0.0266	0.1503	0.0223
	90%	0.6553	0.0147	0.7739	0.0145	0.0204	0.0094	1.1351	0.0284	0.0202	0.1145	0.0144
	50%	0.3508	0.0074	0.3202	0.0096	0.0118	0.0061	0.8294	0.0142	0.0091	0.0402	0.0054
	p-value	0.5200	0.0000	0.9767	0.0000	0.2367	0.0400	0.8967	0.0000	0.2533	0.7500	0.0167
$R = 280$	Statistics	0.3100	0.1405	0.1439	0.0755	0.0132	0.0101	0.5575	0.4661	0.0122	0.0261	0.0246
	95%	0.6743	0.0160	1.1178	0.0190	0.0212	0.0097	1.1536	0.0278	0.0218	0.1194	0.0185
	90%	0.5538	0.0125	0.8284	0.0134	0.0167	0.0085	1.0586	0.0229	0.0184	0.0812	0.0138
	50%	0.3305	0.0071	0.2942	0.0087	0.0111	0.0057	0.7472	0.0134	0.0079	0.0330	0.0052
	p-value	0.5600	0.0000	0.9500	0.0000	0.2567	0.0333	0.8800	0.0000	0.2067	0.7333	0.0333
$R = 320$	Statistics	0.2945	0.1320	0.1431	0.0555	0.0123	0.0091	0.5448	0.4009	0.0112	0.0248	0.0224
	95%	0.6238	0.0140	1.0127	0.0122	0.0201	0.0091	1.0235	0.0227	0.0161	0.1043	0.0116
	90%	0.5140	0.0106	0.6466	0.0110	0.0165	0.0078	0.9333	0.0189	0.0127	0.0769	0.0097
	50%	0.2871	0.0064	0.2458	0.0078	0.0097	0.0053	0.7193	0.0119	0.0067	0.0320	0.0045
	p-value	0.4633	0.0000	0.9367	0.0000	0.2533	0.0500	0.8800	0.0000	0.1300	0.7500	0.0067
$R = 360$	Statistics	0.2947	0.1316	0.1514	0.0453	0.0120	0.0085	0.5632	0.3675	0.0108	0.0246	0.0213
	95%	0.4473	0.0099	0.5954	0.0118	0.0147	0.0075	0.9043	0.0189	0.0136	0.0711	0.0099
	90%	0.3912	0.0079	0.4876	0.0099	0.0129	0.0068	0.8177	0.0159	0.0106	0.0511	0.0082
	50%	0.2493	0.0057	0.2165	0.0071	0.0089	0.0047	0.6424	0.0102	0.0057	0.0268	0.0041
	p-value	0.3000	0.0000	0.8933	0.0000	0.1600	0.0233	0.7100	0.0000	0.0933	0.6033	0.0067
$R = 400$	Statistics	0.3082	0.1379	0.1644	0.0473	0.0123	0.0083	0.6138	0.3726	0.0107	0.0252	0.0211
	95%	0.3699	0.0085	0.4499	0.0097	0.0132	0.0061	0.8082	0.0150	0.0104	0.0537	0.0092
	90%	0.3026	0.0075	0.3654	0.0089	0.0114	0.0056	0.7490	0.0135	0.0093	0.0402	0.0077
	50%	0.2173	0.0052	0.1942	0.0063	0.0076	0.0040	0.5584	0.0094	0.0051	0.0221	0.0035
	p-value	0.0967	0.0000	0.7167	0.0000	0.0700	0.0000	0.3233	0.0000	0.0400	0.3000	0.0033

§ Notes: See notes to Table 2.