Inference for Factor-MIDAS Regression Models

Julia Koh
Department of Econometrics and Operations Research
Tilburg University

January 29, 2026

Abstract

Factor-MIDAS regression models are often used to forecast a target variable using common factors extracted from a large panel of predictors observed at higher frequencies. In the paper, we derive the asymptotic distribution of the factor-MIDAS regression estimator coefficients. We show that there exists an asymptotic bias because the factors are estimated. However, the fact that factors and their lags are aggregated in a MIDAS regression model implies that the asymptotic bias depends on both serial and cross-sectional dependence in the idiosyncratic errors of the factor model. Thus, bias correction is more complicated in this setting. Our second contribution is to propose a bias correction method based on a plug-in version of the analytical formula we derive. This bias correction can be used in conjunction with asymptotic normal critical values to produce asymptotically valid inference. Alternatively, we can use a bootstrap method, which is our third contribution. We show that correcting for bias is important in simulations and in an empirical application to forecasting quarterly U.S. real GDP growth rates using monthly factors.

Keywords: Factor model, Bootstrap, Asymptotic Bias

1 Introduction

MIDAS (Mixed-Data Sampling) regressions are popular tools in forecasting. Originally proposed by Ghysels et al. (2004; 2005; 2006; 2007), these models combine predictors observed at high frequencies by relying on a parametric temporal aggregation function to forecast a target variable sampled at a lower frequency. Originally proposed to handle financial variables, they have become standard tools in macroeconomic forecasting (see e.g., Clements and Galvão (2008; 2009), which relies on MIDAS autoregressions for nowcasting U.S. real output growth).

More recently, standard MIDAS regressions have been generalized to "factor-MIDAS regressions" (or "factor-augmented MIDAS regression models") by including as predictors common factors extracted from a large panel of time series sampled at a higher frequency than the target variable. By combining with the dimension reduction properties of factor models, factor-MIDAS regressions are powerful tools for forecasting and they are often used in empirical applications (see for instance Marcellino and Schumacher (2010), Monteforte and Moretti (2013), Kim and Swanson (2018), and Ferrara and Marsilli (2019)). Estimation of factor-MIDAS regressions is complicated by the fact that some of the predictors are latent common factors. It typically proceeds in two steps: we first extract the common factors using principal component analysis, and then estimate the model using nonlinear least squares, where the estimated factors are aggregated by a temporal aggregation scheme.

Although factor-MIDAS regressions are empirically popular, no formal inference methods have been proposed in the literature. Our paper proposes inference methods for factor-MIDAS regression models and provides the theoretical justification for these methods. The main contributions of this paper are as follows. Firstly, the asymptotic distribution of the factor-MIDAS regression estimators is derived. We show that there is an asymptotic bias in the second step due to the estimation of the factors in the first step. Secondly, we propose two inference methods accounting for this bias: a bias correction method based on the bias formula we derive and a bootstrap method.

Our work is related to the existing literature on factor-augmented regression models (without mixed frequencies). Bai and Ng (2006) first studied the "generated regressor" problem in standard factor-augmented regression models. They showed that inference for the regression coefficients could proceed as if the estimated factors were observed if the cross-sectional dimension N was sufficiently large relative to the time dimension T, more precisely if $\sqrt{T}/N \to 0$. More recently, Gonçalves and Perron (2014) (henceforth, GP (2014)) showed that an asymptotic bias may appear under more relaxed assumption (i.e. if $\sqrt{T}/N \to c,\ 0 < c < \infty$). We extend these results to factor-MIDAS regression models. This is not a trivial extension for two main reasons. First, the estimation problem in a factor-MIDAS regression model is more complicated because the predictors include latent factors (and their lags) sampled at a different frequency than a variable of interest. In addition, the second step is based on nonlinear least squares (rather than OLS) because of a temporal aggregation, and this complicates the asymptotic analysis. In particular, whereas the bias derived in Gonçalves and Perron (2014) depends only on the cross-sectional dependence, the asymptotic bias of a factor-MIDAS regression model depends on both serial and cross-sectional dependence in the idiosyncratic errors. Consequently, different methods of inference are required for factor-MIDAS regressions.

We consider two different methods of inference in this context. The first is an analytical bias correction that can be used along with asymptotic normal critical values. Our plug-in bias correction is robust to both serial and cross-sectional dependence of unknown form in the idiosyncratic errors. It is based on the asymptotic formula of the bias we derive, replacing unknown parameters with consistent estimators. As in Ludvigson and Ng (2009), who also propose a bias correction formula for the standard factor-augmented regression model without mixed frequencies, we rely on the CS-HAC estimator of Bai and Ng (2006) to account for cross-sectional dependence. However, our estimator is more complex since it also requires robustness to serial dependence.

Our second method of inference is based on the bootstrap. The bootstrap has two significant advantages: it can perform better in finite samples, and it avoids the explicit estimation of the bias term which can be complicated in this context. We propose a bootstrap procedure inspired by Gonçalves and Perron (2014), which is a residual-based bootstrap. Although the method is inspired by Gonçalves and Perron (2014), the asymptotic justification is substantially more complicated. More importantly, the need to mimic the asymptotic bias requires the bootstrap to be robust to both serial and cross-sectional dependence. Since none of the existing bootstrap methods in the literature allows for both forms of dependence, we propose a new bootstrap method for factor models that has these properties. Our method is based on an application of the sieve bootstrap to the idiosyncratic residuals of each time series in the panel data model, where the corresponding innovations are resampled using the cross-sectional dependent bootstrap proposed by Gonçalves and Perron (2020). We show that this bootstrap method is asymptotically valid when each idiosyncratic error in the factor model is generated by an $AR(\infty)$ process with innovations that are potentially cross-sectionally correlated across the panel. A special case of this new bootstrap method is considered by Gonçalves, Koh, and Perron (2024) when testing for the number of common factors in group factor models (as proposed by Andreou, Gagliardini, Ghysels, and Rubin (2019)) without theoretical justification.

We illustrate the good finite sample performance of the plug-in bias estimator and the bootstrap using Monte Carlo simulations. In particular, the results show that it is important to correct the bias due to the estimation of the factors in the first step. Although both the plug-in bias correction and the bootstrap methods replicate the bias well, the bootstrap outperforms the plug-in bias estimator by further reducing the coverage rate distortions. Finally, we apply our new inference methods to an empirical application where we nowcast quarterly U.S. real GDP growth rate using monthly macroeconomic factors. The results show that there is a significant bias, thereby indicating the importance of correcting it.

The rest of this paper is organized as follows. In Section 2, we derive the asymptotic distribution of the factor-augmented MIDAS regression model and propose a plug-in bias estimator. In Section 3, we propose and theoretically justify the bootstrap. The simulation results are shown in Section 4, and the empirical application is discussed in Section 5. Section 6 concludes the paper.

For any matrix A, $\|A\|$ denotes its Frobenius norm defined as $\|A\| = (trace(A'A))^{1/2}$. $\rho(A)$ denotes the Euclidean vector norm of the vector Ax: $\rho(A) = \max_{\|x\|=1} \|Ax\|$, where $\|Ax\| = (x'A'Ax)^{1/2}$.

2 Asymptotic Theory

2.1 Factor-augmented MIDAS regression models

The MIDAS regression model projects high-frequency variables onto a target variable, which is denoted as y_t. The regressors are observed at most m times between t and $t-1$. To handle variables sampled at mixed frequency, a MIDAS regression aggregates the high-frequency variables with a lag polynomial function. The basic MIDAS regression model with a single observed regressor x_t can be written as follows:

$$y_t = \beta_0 + \beta_1 W(L^{1/m};\theta)x_t + \varepsilon_t, \quad t = 1, \dots, T, \tag{1}$$

where $W(L^{1/m};\theta) = \sum_{k=1}^{K} w_k(\theta)L^{k/m}$ and $L^{k/m}x_t = x_{t-k/m}$. Here, $w_k(\theta)$ is a weighting function that temporally aggregates the regressor and its lags, and θ is a $p \times 1$ vector of weighting parameters. To identify β_1, we assume that $w_k(\theta) \in (0,1)$ and $\sum_{k=1}^{K} w_k(\theta) = 1$. A common weighting scheme in the MIDAS regression model is the exponential Almon lag with two parameters such that

$$w_k(\theta) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=1}^{K} \exp(\theta_1 k + \theta_2 k^2)}. \tag{2}$$

Other weighting schemes include the beta function and the linear function (see Ghysels, Valkanov, and Serrano (2009) for detail).

In this paper, we consider the factor-MIDAS regression model, which employs unobserved high-frequency factors as regressors. In particular, letting the regressor x_t in (1) be replaced by a latent factor, we write the model as follows.

$$y_t = \beta_0 + \beta_1 W(L^{1/m};\theta)f_t + \varepsilon_t = \beta_0 + \beta_1 \sum_{k=1}^{K} w_k(\theta)f_{t-k/m} + \varepsilon_t, \quad t = 1, \dots, T,$$

where $f_{t-k/m}$ is a (single) factor in the following panel factor model,

$$X_{t-k/m} = \Lambda f_{t-k/m} + e_{t-k/m}, \quad k = m-1, \dots, 0, \text{ and } t = 1, \dots, T. \tag{3}$$

The factor model includes factor loadings denoted by Λ and an idiosyncratic error term, $e_{t-k/m}$. If there are r unobserved factors, represented by a $r \times 1$ vector of common factors denoted by $f_{t-k/m}$ in the factor model (3), then the model can be generalized as follows.

$$y_t = \beta_0 + \beta_1' W(L^{1/m};\theta) f_t + \varepsilon_t = \beta_0 + \beta_1' F_t(\theta) + \varepsilon_t, \quad t = 1,\ldots,T, \tag{4}$$

where $\beta_1 = (\beta_{1,1},\ldots,\beta_{1,r})'$, and $\theta = (\theta_1',\ldots,\theta_r')'$ with $\theta_j = (\theta_{j,1},\ldots,\theta_{j,p})'$, a $p \times 1$ weighting parameter[1] for j-th factor, for $j = 1,\ldots,r$. We define $F_t(\theta) \equiv W(L^{1/m};\theta) f_t$ in the second equality. In fact, the temporal aggregation in this generalized model applies on a vector as

$$F_t(\theta) = \sum_{k=1}^{K} w_k(\theta) L^{k/m} f_t = \sum_{k=1}^{K} w_k(\theta) f_{t-k/m},$$

where $w_k(\theta)$ is a $r \times r$ diagonal matrix such that $w_k(\theta) \equiv \mathrm{diag}(w_{k,1}(\theta_1),\ldots,w_{k,r}(\theta_r))$, where $w_{k,j}(\theta_j)$ is the weight for the k-th lag of the j-th factor.[2] To derive the distribution in the next section, we further simplify the general factor-MIDAS regression model (4) to

$$y_t = g(F_t,\alpha) + \varepsilon_t, \quad t = 1,\ldots,T, \tag{5}$$

where $g(F_t,\alpha) = \beta_0 + \beta_1' F_t(\theta)$, $\alpha = (\beta',\theta')'$ with $\beta = (\beta_0,\beta_1')'$, and $F_t = (1, f_t', f_{t-1/m}',\ldots,f_{t-K/m}')'$. For convenience, we use the high frequency time index denoted by $t_h = 1,\ldots,T_H$, where $T_H = mT$. We derive this by noting that $t_h = m((t-1)+i/m)$ for $i = 1,\ldots,m$, and $t = 1,\ldots,T$.[3] Using this notation, we can write the factor model as $X_{t_h} = \Lambda f_{t_h} + e_{t_h}$, for $t_h = 1,\ldots,T_H$. Using the matrix notation, we write the factor model as $X = f\Lambda' + e$, where X is a $T_H \times N$ matrix of high-frequency time series, $f = (f_1,\ldots,f_{T_H})'$ is a $T_H \times r$ matrix of common factors, and e is a $T_H \times N$ matrix of idiosyncratic errors.[4]

[1] Note that at least one component of β_1 needs to be non-zero to identify the weighting parameters, θ.

[2] Note that when $m = 1$ and $K = 0$, the factor-MIDAS regression model is equivalent to the standard factor-augmented regression model in GP (2014).

[3] With this notation, a high-frequency observation at t_h is equivalent to observing it at the i-th intra-period between $t-1$ and t. Note that the time notation in the factor model (3) can be written as $(t-1)+(m-k)/m$.

[4] One may consider a situation where X includes variables with different frequencies, such as monthly and quarterly, while y_t is observed annually. In this case, the group factor model discussed in Andreou et al. (2019) can be exploited to extract the factors.

2.2 Asymptotic Theory

We denote NLS estimators by $\hat{\alpha}$ when the factors are observed. Then, Andreou, Ghysels, and Kourtellos (2010) show that the limiting distribution of $\hat{\alpha}$ is as following:

$$\sqrt{T}(\hat{\alpha} - \alpha_0) \xrightarrow{d} N(0, \Sigma^{-1}\Omega\Sigma^{-1}), \tag{6}$$

where $\alpha_0 = (\beta', \theta')'$, $\Sigma = E[g_{\alpha,t}g'_{\alpha,t}]$, and $\Omega = E[\varepsilon_t^2 g_{\alpha,t}g'_{\alpha,t}]$ with $g_{\alpha,t} = \partial g(F_t, \alpha)/\partial\alpha$. When the true factors are observed, the estimators are normally distributed with mean zero and a sandwich variance.

In factor-MIDAS models, however, the factors are latent, and we have to estimate them. Accordingly, the estimation in the factor-MIDAS regression model proceeds in two steps. First, we estimate the common factors from a panel dataset of high-frequency indicators by principal component analysis (PCA). The estimated factors, $\tilde{f}$, are equivalent to $\sqrt{T_H}$ times the eigenvectors of $XX'/T_H N$ corresponding to the r largest eigenvalues (in decreasing order). The estimated factor loadings are $\tilde{\Lambda} = X'\tilde{f}/T_H$.[5] Second, we estimate the parameters β and θ using nonlinear least squares (NLS) by regressing the low frequency variable on the temporally aggregated estimated factors at high-frequency. In the factor model, the estimated factors $\tilde{f}_t$ are only consistent for Hf_t, where the rotation matrix H is defined as $H = \tilde{V}^{-1}\frac{\tilde{f}'f}{T_H}\frac{\Lambda'\Lambda}{N}$, and $\tilde{V}$ is a $r \times r$ diagonal matrix of eigenvalues of $XX'/T_H N$ in a descending order (for more details, see Bai (2003)). By incorporating the estimated factors in the regression and noting the rotation of the factors, we can rewrite (4) as follows.

$$y_t = \beta_0 + \beta_1' H^{-1}\tilde{F}_t(\theta) + \beta_1' H^{-1}(HF_t(\theta) - \tilde{F}_t(\theta)) + \varepsilon_t = g(\tilde{F}_t, \alpha) + \xi_t, \tag{7}$$

[5]When $T_H > N$, we use normalization such that $\Lambda'\Lambda/N = I_r$ and $f'f$ is a diagonal matrix, which is computationally easier. In this case, $\tilde{\Lambda}$ is the matrix of $\sqrt{N}$ times the eigenvectors of $X'X/T_H N$ corresponding to the r largest eigenvalues and the estimated factors are $\tilde{f} = X\tilde{\Lambda}/N$.

where $g(\tilde{F}_t, \alpha) = \beta_0 + \beta_1' H^{-1} \tilde{F}_t(\theta)$, $\alpha = (\beta_0, \beta_1' H^{-1}, \theta')'$, and $\tilde{F}_t(\theta) = \sum_{k=1}^{K} w_k(\theta) \tilde{f}_{t-k/m}$. The coefficient on the aggregated factors estimates $\beta_1' H^{-1}$. Moreover, the estimation error of the factors implies that the regression error term is $\xi_t = \beta_1' H^{-1}(HF_t(\theta) - \tilde{F}_t(\theta)) + \varepsilon_t$. We denote the NLS estimators of α in (7) by $\tilde{\alpha} = (\tilde{\beta}', \tilde{\theta}')'$ to distinguish from $\hat{\alpha} = (\hat{\beta}', \hat{\theta}')'$, which are the estimators from the regression of y_t on the true factors f_t. Next, we derive the limiting distribution of $\sqrt{T}(\tilde{\alpha} - \alpha)$ under the assumption that $\sqrt{T}/N \to c$, where $0 \leq c < \infty$. Note that although the variable of interest is a linear function of factor estimation error similar to the factor-augmented regression models, there exists a nonlinear weighting function. Furthermore, unlike standard factor-augmented regression models, the lags of the factors are incorporated. As will be demonstrated in the next theorem, the incorporation of the lags of the factors results in the fact that the asymptotic bias relies on the time-series dependence and cross-sectional dependence in the idiosyncratic error term.[6]

The asymptotic distribution of the estimators is derived under Assumptions A.1 - A.6 in Section A in Online Appendix. We also introduce the following notations: $V \equiv \text{plim}\, \tilde{V}$, $Q \equiv \text{plim}\left(\frac{\tilde{f}' f}{T_H}\right)$, $Q_k \equiv \text{plim}\left(\frac{1}{T_H - k} \sum_{t_h=k+1}^{T_H} \tilde{f}_{t_h}' f_{t_h - k}\right)$, and $\Sigma_{\tilde{f}} \equiv V^{-1} Q \Gamma Q' V^{-1}$, which is the asymptotic variance of $\sqrt{N}(\tilde{f}_{t_h} - H f_{t_h})$.[7] The asymptotic variance of the factor estimation error is a function of Γ, which is defined by $\Gamma \equiv \lim_{N \to \infty} Var\left(\frac{\Lambda' e_{t_h}}{\sqrt{N}}\right)$. We assume that the idiosyncratic errors in the factor model, e_{t_h} is stationary in Assumption A.2-(d). Under the stationarity of the idiosyncratic errors, we also denote $\Gamma_k \equiv \lim_{N \to \infty} Cov\left(\frac{\Lambda' e_{t_h - k}}{\sqrt{N}}, \frac{\Lambda' e_{t_h}}{\sqrt{N}}\right)$. Note that by the identification assumption, Assumption A.1-(d) in Online Appendix, we have $Q = H_0$, where $H_0 = \text{plim}\, H$, and H_0 is a diagonal matrix of ± 1, where the sign is determined by the sign of $\tilde{f}' f / T_H$ (for the detail of the proof, see the proof of (2) in Bai and

[6] Note that the time-series dependence in the idiosyncratic error term does not appear in the asymptotic bias in the standard factor augmented regression models. For detail, see GP (2014) (their Theorem 2.1).

[7] For the details, see Bai (2003).

Ng (2013)). Therefore, the asymptotic variance can be also written as $\Sigma_{\tilde{f}} = V^{-1}H_0\Gamma H_0'V^{-1}$.

Theorem 2.1 (Asymptotic distribution of the estimators in the factor-MIDAS models) *If $\sqrt{T}/N \to c$, where $0 \le c < \infty$, and Assumptions A.1 - A.6 in Section A in Online Appendix hold,*

$$\sqrt{T}(\tilde{\alpha} - \alpha) \xrightarrow{d} N(-c\Delta_\alpha, \Sigma_\alpha), \tag{8}$$

where $\Sigma_\alpha \equiv \Phi_0'^{-1}\Sigma^{-1}\Omega\Sigma^{-1}\Phi_0^{-1}$ with $\Phi_0 = diag(1, H_0, I_p)$, and

$$\Delta_\alpha = \begin{bmatrix} \Delta_\beta \\ \Delta_\theta \end{bmatrix} = (\Phi_0\Sigma\Phi_0')^{-1} \begin{bmatrix} B_\beta \\ B_\theta \end{bmatrix}. \tag{9}$$

$B_\beta = (B_{\beta_0}, B_{\beta_1}')'$ and B_θ are such that $B_{\beta_0} = 0$,

$$\begin{aligned} B_{\beta_1} = &\Big[\sum_{k=1}^{K} w_k(\theta)\Big\{\Sigma_{\tilde{f}} + V\Sigma_{\tilde{f}}V^{-1}\Big\}w_k(\theta) \\ &+ \sum_{k=1}^{K}\sum_{l\neq k}^{K} w_k(\theta)\Big\{V^{-1}H_0\Gamma_{k-l}H_0'V^{-1} + Q_{k-l}\Gamma H_0'V^{-2}\Big\}w_l(\theta)\Big] \operatorname{plim}(\tilde{\beta}_1), \end{aligned} \tag{10}$$

and

$$\begin{aligned} B_\theta = \operatorname{plim}(\tilde{\beta}_1) \circ &\Big[\sum_{k=1}^{K} \frac{\partial w_k(\theta)}{\partial\theta}\Big\{\Sigma_{\tilde{f}} + V\Sigma_{\tilde{f}}V^{-1}\Big\}w_k(\theta) \\ &+ \sum_{k=1}^{K}\sum_{l\neq k}^{K} \frac{\partial w_k(\theta)}{\partial\theta}\Big\{V^{-1}H_0\Gamma_{k-l}H_0'V^{-1} + Q_{k-l}\Gamma H_0'V^{-2}\Big\}w_l(\theta)\Big] \operatorname{plim}(\tilde{\beta}_1), \end{aligned} \tag{11}$$

where $\frac{\partial w_k(\theta)}{\partial\theta} \equiv \operatorname{diag}\left(\frac{\partial w_{k,1}(\theta_1)}{\partial\theta_1}, \ldots, \frac{\partial w_{k,r}(\theta_r)}{\partial\theta_r}\right)$ is a block diagonal matrix and the j-th diagonal block is a $p \times 1$ vector given by $\frac{\partial w_{k,j}(\theta_j)}{\partial\theta_j}$ for $j = 1, \ldots, r$.

In (11) in Theorem 2.1, we use the Hadamard product which is equivalent to $(A \circ B)_{ij} = A_{ij}B_{ij}$. More specifically, $\beta \circ \frac{\partial w_k(\theta)}{\partial\theta}$ is a block diagonal matrix where the j-th diagonal block contains $\beta_j \frac{\partial w_{j,k}(\theta_j)}{\partial\theta_j}$ for $j = 1, \ldots, r$. Based on Theorem 2.1, the bias of the estimators is proportional to c, the limiting value of $\sqrt{T}/N$, and also to $\text{plim}(\tilde{\beta}_1) = (H_0^{-1})'\beta_1$. This implies that the estimates are biased unless $\beta_1 = 0$ or $c = 0$. Additionally, the asymptotic variance of the estimated factors, $\Sigma_{\tilde{f}}$, affects the bias. Since the variance of the factor estimation error depends on Γ, which is a variance of the scaled average of the factor loadings and the idiosyncratic errors in the factor model, the cross-sectional dependence of factor errors matters. These findings are similar to the bias in the context of GP (2014).

It is important to highlight two main differences in the asymptotic bias between the factor-MIDAS regression model and standard factor-augmented regression models. Firstly, the bias in the MIDAS regression model depends on the weighting scheme, $w_k(\theta)$, due to a temporal aggregation.[8] Secondly, the bias depends on the covariance of the cross-sectional average of factor loadings and the idiosyncratic error terms between two distinct periods, represented as Γ_{k-l}. This term arises due to the presence of the lags of the estimated factors. To see this, consider a simple factor-augmented regression model with a lag and without mixed-frequency variables as follows.

$$y_t = \beta_1 f_t + \beta_2 f_{t-1} + \varepsilon_t = \beta' F_t + \varepsilon_t,$$

where $\beta = (\beta_1, \beta_2)'$ and $F_t = (f_t, f_{t-1})'$. We assume that the factor is a single factor for simplicity. By the fact that the factors are estimated, we can rewrite it as follows.

$$y_t = \beta' H^{-1} \tilde{F}_t + \beta' H^{-1}(H F_t - \tilde{F}_t) + \varepsilon_t.$$

Note that since we include a lag of the factor, we have a factor estimation error at $t-1$ as well as contemporaneous factor estimation error. Letting $\hat{\beta}$ be OLS estimator from a

[8]When there is no temporal aggregation, the MIDAS regression becomes unrestricted MIDAS (U-MIDAS) proposed by Foroni, Marcellino, and Schumacher (2015). If the estimated factors are used as predictors in U-MIDAS, there will be bias that depends on cross-sectional and serial dependence of the idiosyncratic error term in the factor model, by the fact that lags of the estimated factors are present.

regression of y_t on $\tilde{F}_t$, we can show that

$$\sqrt{T}(\hat{\beta} - H^{-1}\beta) = \left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_t\tilde{F}_t'\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{F}_t\varepsilon_t + \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_t\tilde{F}_t'\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_t(HF_t - \tilde{F}_t)'H^{-1}\beta.$$

In fact, we can show that $\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_t(\tilde{F}_t - HF_t)' \xrightarrow{p} \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}Var\left(\sqrt{N}(\tilde{F}_t - HF_t)\right)\right) = O_p(1/N)$ by Bai (2003) (see their Lemma B.2) and GP (2014). Therefore, the second term is $O_p(\sqrt{T}/N)$ and drives the asymptotic bias under the rate condition such that $\sqrt{T}/N \to c$ for $0 \leq c < \infty$. In GP (2014), since the contemporaneous factor is the sole predictor in their factor-augmented regression model, the variance of contemporaneous factor estimation error appears alone. More specifically, the bias is driven by $\frac{1}{T}\sum_{t=1}^{T}Var\left(\sqrt{N}(\tilde{f}_t - Hf_t)\right)$, which depends on $\frac{1}{T}\sum_{t=1}^{T}Var\left(\frac{\Lambda' e_t}{\sqrt{N}}\right)$. This term implies that the bias depends solely on the cross-sectional dependence of the idiosyncratic error term in the factor model. However, when we incorporate a lag of the factor as a predictor alongside the contemporaneous factor, the covariance between the factor estimation error at t and $t-1$ becomes relevant, which depends on $\frac{1}{T}\sum_{t=1}^{T}Cov\left(\frac{\Lambda' e_t}{\sqrt{N}}, \frac{\Lambda' e_{t-1}}{\sqrt{N}}\right)$. Thus, the inclusion of the lag of the factor indicates that the bias depends not only on the cross-sectional dependence, but also on the time-series dependence of the idiosyncratic error term in the factor model.[9]

In the factor-MIDAS regression model, the inclusion of lagged estimated factors introduces additional complexity. Similar to the previously discussed simple case, we have an extra term such that $\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}Cov(\sqrt{N}(\tilde{f}_{t_h} - Hf_{t_h}), \sqrt{N}(\tilde{f}_{t_h-k} - Hf_{t_h-k}))$ for $k \neq 0$, which depends on $\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}Cov\left(\frac{\Lambda' e_{t_h}}{\sqrt{N}}, \frac{\Lambda' e_{t_h-k}}{\sqrt{N}}\right)$. Therefore, the bias in our context relies on the serial dependence as well as cross-sectional dependence of the idiosyncratic error term in the factor model. This finding holds considerable significance, as the literature surround-

[9]This also explains why the bias in unrestricted MIDAS (U-MIDAS) regression models augmented by the factors depends on cross-sectional as well as serial dependence of the idiosyncratic error term in the factor model.

ing factor-augmented regression models has primarily concentrated on the cross-sectional dependence of the idiosyncratic error term. This focus necessitates the development of novel inference methods that can effectively account for the time-series dependence inherent in the idiosyncratic error term, which appears in our context.

2.3 Plug-in Bias

In this section, we propose an analytical estimator to account for the bias identified in Theorem 2.1. This is inspired by Ludvigson and Ng (2009), where they propose a plug-in bias estimator by replacing the unknown quantities with their consistent estimators and correcting the bias in the context of the factor-augmented regression model. Similarly, we propose a bias-corrected estimator for factor-augmented MIDAS regression models.

In order to do that, we need a consistent estimator for the term Γ_k, which has never been explored previously. Note that it depends on the cross-sectional and the serial dependence of the idiosyncratic error term. When the idiosyncratic error term is serially but not cross-sectionally correlated, we can estimate this term as $\hat{\Gamma}_k = \frac{1}{N(T_H-k)} \sum_{t_h=k+1}^{T_H} \sum_{i=1}^{N} \tilde{\lambda}_i \tilde{\lambda}_i' \tilde{e}_{i,t_h} \tilde{e}_{i,t_h-k}$, where $\hat{\Gamma}_k$ denotes the estimator of Γ_k. However, when the idiosyncratic error term is cross-sectionally and serially dependent, estimating this term is no longer straightforward, as discussed in Bai and Ng (2006). To address this issue, Bai and Ng (2006) propose an estimator for the variance-covariance matrix of the cross-sectional average of factor loadings and the idiosyncratic error term, denoted by Γ. They use the time series observations and truncation with $n < N$ under the covariance stationarity such that $\hat{\Gamma}_{\text{CS-HAC}} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{\lambda}_i \tilde{\lambda}_j \frac{1}{T_H} \sum_{t_h=1}^{T_H} \tilde{e}_{i,t_h} \tilde{e}_{j,t_h}$.

To propose a method to estimate Γ_k that takes into account cross-sectional and serial dependence, we take an approach, similar to the one used in Bai and Ng (2006). We use the time series observations and a truncation method, that limits $n < N$ observations. We denote the estimator for Γ_k by $\hat{\Gamma}_k$, which is defined as follows.

$$\hat{\Gamma}_{k,\text{CS-HAC}} = \frac{1}{T_H - k} \sum_{t_h=k+1}^{T_H} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{\lambda}_i \tilde{\lambda}_j' \tilde{e}_{i,t_h} \tilde{e}_{j,t_h-k}, \tag{12}$$

where $n = \min(\sqrt{N}, \sqrt{T_H})$. Note that by Assumption A.2-(d), Γ_k does not depend on time.

Theorem 2.2 *Suppose the Assumptions A.1 - A.4 in Section A in Online Appendix hold. Then, for any fixed* $k = 0, 1, 2, \ldots, K-1$

$$\|\hat{\Gamma}_k - H_0^{-1\prime} \Gamma_k H_0^{-1}\| \xrightarrow{p} 0 \quad \textit{if} \quad \frac{n}{\min(N, T_H)} \to 0,$$

Here, in Theorem 2.2, $\hat{\Gamma}_k$ depends on the assumption on the serial and cross-sectional dependence in the idiosyncratic errors of the factor model. If there is only serial dependence, $\hat{\Gamma}_k = \frac{1}{N} \sum_{i=1}^{N} \tilde{\lambda}_i \tilde{\lambda}_i' \frac{1}{T_H - k} \sum_{t_h=k+1}^{T_H} \tilde{e}_{i,t_h} \tilde{e}_{i,t_h-k}$. If we allow for cross-sectionally dependence additionally, $\hat{\Gamma}_k = \tilde{\Gamma}_{k,\text{CS-HAC}}$ defined in (12). Note that if $k = 0$, our estimators are equivalent to the estimators proposed in Bai and Ng (2006). Theorem 2.2 enables us to construct consistent estimators for (10) and (11) as follows.

$$\hat{B}_{\beta_1} = \left[2 \sum_{k=1}^{K} w_k(\tilde{\theta}) \tilde{\Sigma}_{\tilde{f}} w_k(\tilde{\theta}) + \sum_{k=1}^{K} \sum_{l \neq k}^{K} w_k(\tilde{\theta}) \left\{ \tilde{V}^{-1} \hat{\Gamma}_{k-l,\text{CS-HAC}} \tilde{V}'^{-1} + \tilde{Q}_{k-l} \hat{\Gamma}_{\text{CS-HAC}} \tilde{V}^{-2} \right\} w_l(\tilde{\theta}) \right] \tilde{\beta}_1, \text{ and}$$

$$\hat{B}_{\theta} = \tilde{\beta}_1 \circ \left[2 \sum_{k=1}^{K} \frac{\partial w_k(\tilde{\theta})}{\partial \theta} \tilde{\Sigma}_{\tilde{f}} w_k(\tilde{\theta}) + \sum_{k=1}^{K} \sum_{l \neq k}^{K} \frac{\partial w_k(\tilde{\theta})}{\partial \theta} \left\{ \tilde{V}^{-1} \hat{\Gamma}_{k-l,\text{CS-HAC}} \tilde{V}^{-1} + \tilde{Q}_{k-l} \hat{\Gamma}_{\text{CS-HAC}} \tilde{V}^{-2} \right\} w_l(\tilde{\theta}) \right] \tilde{\beta}_1,$$

where $\tilde{\Sigma}_{\tilde{f}} = \tilde{V}^{-1} \tilde{Q} \hat{\Gamma}_{\text{CS-HAC}} \tilde{Q} \tilde{V}^{-1}$ with $\tilde{Q} = \tilde{f}' \tilde{f} / T_H$, and $\tilde{Q}_{k-l} = \sum_{t_h=k+1}^{T_H} \tilde{f}_{t_h}' \tilde{f}_{t_h-k}$. Note that the bias estimates can be simpler under the restriction on either cross-sectional or serial dependence, or both. We denote the bias-corrected estimator by $\hat{\alpha}_{\text{BC}}$ such that $\hat{\alpha}_{\text{BC}} \equiv \tilde{\alpha} - (-\frac{1}{N} \hat{\Delta}_{\alpha})$. Here, $-\hat{\Delta}_{\alpha}$ is the estimate of the bias in $\tilde{\alpha}$, where $\hat{\Delta}_{\alpha} = \hat{\Sigma}^{-1} (\hat{B}_{\beta}', \hat{B}_{\theta}')'$ with $\hat{\Sigma}$ a consistent estimator of Σ, $\hat{B}_{\beta} = (\hat{B}_{\beta_0}, \hat{B}_{\beta_1}')'$, and $\hat{B}_{\beta_0} = 0$.

Proposition 2.1 *Suppose the Assumptions A.1 - A.6 in Section A in Online Appendix hold and* $\sqrt{T}/N \to c$, *where* $0 \leq c < \infty$, *then*

$$\sqrt{T}(\hat{\alpha}_{BC} - \alpha) \xrightarrow{d} N(0, \Sigma_\alpha). \tag{13}$$

Based on Proposition 2.1, the bias corrected estimator no longer contains an asymptotic bias. However, it is well known that an approach based on asymptotic theory does not perform well in finite samples. Additionally, the bias takes a very complicated form in our context, which makes it difficult to implement. Therefore, we discuss an alternative approach, a bootstrap method in the next section.

3 Bootstrap method: AR-sieve+CSD bootstrap

In this section, we propose a bootstrap method and show its validity by proving that our method satisfies bootstrap high level conditions under which any general residual-based bootstrap is satisfied. We leave the bootstrap high level conditions in the appendix (see Section C in the Online Appendix).

In particular, we propose a bootstrap procedure, where we resample the factor model and the MIDAS regression model, and then obtain the bootstrap estimates. Note that in Theorem 2.1, we show that the asymptotic bias in our context relies on the cross-sectional and serial dependence in the idiosyncratic error term in the factor model, therefore, it is crucial that the bootstrap resampled idiosyncratic error term in the factor model mimics these dependences. To the best of our knowledge, replicating the time-series dependence in the error term in the factor model has not been studied in the literature. GP (2014) propose a wild bootstrap and prove its validity in the context of the factor-augmented regression models under no cross-sectional dependence in the error term in the factor model.[10] To allow for cross-sectional dependence, Gonçalves and Perron (2020) propose a bootstrap method that utilizes a thresholding technique to allow for the cross-sectional dependence, so-called CSD (cross-sectional dependent) bootstrap. However, these methods cannot be used in our context as it destroys the serial dependence in the idiosyncratic error terms.

On the other hand, to resample the error term in the MIDAS regression model, GP (2014) propose a wild bootstrap under the assumption that the regression error terms follow martingale difference sequence. Djogbenou, Gonçalves, and Perron (2015) propose a block wild bootstrap and a dependent wild bootstrap to resample the regression error terms to account for serially correlated regression error terms. Depending on the assumption a researcher is willing to make, either the approach proposed by GP (2014) or by Djogbenou et al. (2015) can be similarly applied to resample the regression error terms in our context. In this paper, for simplicity, we rely on the assumption that the regression error terms follow martingale difference sequence and use the wild bootstrap.

The key finding in our paper is that the bias within our framework is influenced by both serial and cross-sectional dependence in the idiosyncratic error term in the factor model. To address this, we propose a novel bootstrap method that can replicate both dependences. Specifically, we combine autoregressive sieve bootstrap and the CSD bootstrap to resample the residuals in the factor model.[11] The autoregressive sieve bootstrap, initially introduced by Bühlmann (1997) and further explored by Kreiss, Paparoditis, and Politis (2011) and Meyer and Kreiss (2015), has been effectively applied to the estimated factors by Bi, Shang, Yang,

[10] Note that the asymptotic bias in the factor augmented regression models studied in GP (2014) only depends on the cross-sectional dependence. For detail, see GP (2014).

[11] Note that we cannot use block-based bootstrap or dependent wild bootstrap to account for serial dependence, because these bootstrap methods induce a zero cross-sectional dependence. (For detail, see Gonçalves and Perron (2020).)

and Zhu (2021). In our paper, we combine this method with the CSD bootstrap method and apply it to the residuals in the factor model, which we refer to as the AR-sieve+CSD bootstrap method. A more restricted version of our approach is recently considered by Gonçalves et al. (2024), where they substitute the autoregressive sieve bootstrap with an autoregressive parametric bootstrap of order one, albeit without theoretical justification. Also, as addressed in Bühlmann (1997), the autoregressive sieve bootstrap method offers more flexibility than a parametric autoregressive model, which is highly subject to model misspecification. The AR-sieve+CSD bootstrap method resamples each time series residual in the factor model through an autoregressive sieve process, while the corresponding innovations are resampled by the CSD bootstrap method. This approach effectively captures cross-sectional dependence in the innovation terms through the CSD bootstrap method and the serial dependence through the autoregressive process. The detailed algorithm to use the AR-sieve+CSD bootstrap to resample the residuals in the factor model can be found in Algorithm 1.[12] In Algorithm 1, we resample the residuals in the factor model similar to the bootstrap procedure in Kreiss et al. (2011) and Bühlmann (1997). The difference is that we resample the innovation terms in the autoregressive process using CSD bootstrap proposed by Gonçalves and Perron (2020).

One might consider utilizing high-dimensional vector autoregressive (VAR) models to resample the idiosyncratic error term in the factor model. Recent studies, such as those by Kock and Callot (2015) and Krampe, Kreiss, and Paparoditis (2021), have explored this high-dimensional VAR model. Kock and Callot (2015) establishes oracle inequalities for both LASSO and adaptive LASSO estimators in the context of high-dimensional VAR models. Meanwhile, Krampe et al. (2021) develops a bootstrap method applicable to this framework.

[12]The full bootstrap procedure to obtain the bootstrap estimators can be found in the Online Appendix.

Algorithm 1 : AR-sieve + CSD Bootstrap for the factor model

For $t_h = 1, \ldots, T_H$, let

$$X^*_{i,t_h} = \tilde{\lambda}_i' \tilde{f}_{t_h} + e^*_{i,t_h} \quad \text{and} \quad X^*_{t_h} = \tilde{\Lambda} \tilde{f}_{t_h} + e^*_{t_h},$$

where e^*_{i,t_h} is obtained as follows.

For each $i = 1, \ldots, N$, select an order $p_i = p_i(T_H)$, $p_i << T_H$, for example, by an information criterion such as the Akaike information criterion (AIC), and fit a p_i-th order autoregressive model to $\tilde{e}_{i,1}, \ldots, \tilde{e}_{i,T_H}$, where $\tilde{e}_{i,t_h} = X_{i,t_h} - \tilde{\lambda}_i \tilde{f}_{t_h}$. We denote $\tilde{\phi}_i(p_i) = (\tilde{\phi}_{i,j}(p_i), j = 1, \ldots, p_i)$, the Yule-Walker autoregressive parameter estimators, such that $\tilde{\phi}_i(p_i) = \tilde{\Gamma}(p_i)^{-1} \tilde{\gamma}_{p_i}$, with $\tilde{\gamma}_{p_i} = (\tilde{\gamma}_e(1), \tilde{\gamma}_e(2), \ldots, \tilde{\gamma}_e(p_i))'$ and $\tilde{\Gamma}(p_i) = (\tilde{\gamma}_e(r-s))_{r,s=1,2,\ldots,p_i}$ such that

$$\tilde{\gamma}_e(\tau) = \frac{1}{T_H} \sum_{t_h=1}^{T_H - |\tau|} (\tilde{e}_{i,t_h} - \bar{e}_i)(\tilde{e}_{i,t_h+|\tau|} - \bar{e}_i), \tag{14}$$

for $\tau = 0, \ldots, p_i$ and $\bar{e}_i = T_H^{-1} \sum_{t_h=1}^{T_H} \tilde{e}_{i,t_h}$.

With chosen lag length $p_i = p_i(T_H)$,

$$e^*_{i,t_h} = \sum_{j=1}^{p_i} \tilde{\phi}_{i,j}(p_i) e^*_{i,t_h-j} + u^*_{i,t_h}, \text{ for } t_h = 1, \ldots, T_H, \tag{15}$$

where $u^*_{t_h} = (u^*_{1,t_h}, \ldots, u^*_{N,t_h}) = \tilde{\Sigma}_u^{1/2} \eta_{t_h}$ with $\eta_{t_h} \sim$ i.i.d $(0, I_N)$. The initial conditions are $e^*_{i,0}, \ldots, e^*_{i,1-p_i} = 0$, for $i = 1, \ldots, N$, which is equivalent to the stationary mean of e^*_{i,t_h} in the bootstrap world. Following Gonçalves and Perron (2020), we choose $\tilde{\Sigma}_u$ by a thresholding technique such that

$$\tilde{\Sigma}_u = (\hat{\sigma}_{u,ij})_{i,j=1,\ldots,N},$$

with

$$\hat{\sigma}_{u,ij} = \begin{cases} \tilde{\sigma}_{u,ij} & i = j \\ \tilde{\sigma}_{u,ij} 1\left(|\tilde{\sigma}_{u,ij}| > \omega\right) & i \neq j, \end{cases} \quad \text{with } \tilde{\sigma}_{u,ij} = \frac{1}{T_H} \sum_{t_h=1}^{T_H} \tilde{u}_{i,t_h} \tilde{u}_{j,t_h},$$

where ω is a threshold and $\tilde{u}_{i,t_h} = \tilde{e}_{i,t_h} - \sum_{j=1}^{p_i} \tilde{\phi}_{i,j}(p_i) \tilde{e}_{i,t_h-j}$ for $i = 1, \ldots, N$ and $t_h = 1 + p_i, \ldots, T_H$.

In our paper, we do not address the high-dimensional VAR model due to the complexities involved in its theoretical justification in our framework, opting instead to reserve this for future research.

In order to prove our bootstrap method is valid, we assume that $\{e_{i,t_h}\}_{t_h=1}^{T_H}$ is an infinite order moving average process that can be represented as an AR(∞) process such that $e_{i,t_h} = \sum_{j=1}^{\infty} \phi_{i,j} e_{i,t_h-j} + u_{i,t_h}$, for $t_h = 1, \ldots, T_H$ and $i = 1, \ldots, N$. The innovation terms in AR(∞) process, $u_{t_h} = (u_{1,t_h}, \ldots, u_{N,t_h})'$, are identically and independently distributed from a distribution with mean zero and finite variance, Σ_u. Here, Σ_u is assumed to be non-diagonal to account for cross-sectional dependence in the idiosyncratic error term. More formal representation of the assumptions on our bootstrap method is provided below.

Assumption 1 *λ_i are either deterministic such that $\|\lambda_i\| \leq M \leq \infty$, or stochastic such that $E\|\lambda_i\|^{24} \leq M < \infty$ for all i: $E\|f_{t_h}\|^{24} \leq M < \infty$; $E|e_{i,t_h}|^{24} \leq M < \infty$, for all (i, t_h); and for some $q > 1$, $E|\varepsilon_t|^{4q} \leq M < \infty$, for all t.*

Assumption 2 *$E(\varepsilon_t | y_t, F_t, y_{t-1}, F_{t-1}, \ldots) = 0$, and $F_t = (f_{t-1/m}, \ldots, f_{t-k/m})'$ and ε_t are independent of the idiosyncratic errors e_{i,s_h} for all (i, s_h, t).*

Assumption 3 *$e_{i,t_h} = \sum_{j=1}^{\infty} \phi_{i,j} e_{i,t_h-j} + u_{i,t_h}$, with $\sum_{j=1}^{\infty} (1+|j|)^r |\phi_{i,j}|^8 < \infty$ for some $r \geq 0$, for $i = 1, \ldots, N$.*

Assumption 4 *$\Sigma_u \equiv E(u_{t_h} u_{t_h}') = (\sigma_{u,ij})_{i,j=1,\ldots,N}$, with $u_{t_h} = (u_{1,t_h}, \ldots, u_{N,t_h})'$, for all t_h, i, j and is such that $\lambda_{\min}(\Sigma_u) > c_1$ and $\lambda_{\max}(\Sigma_u) < c_2$ for some positive constants c_1 and c_2.*

Assumption 5 *As $N, T_H \to \infty$ such that $\log N / T_H \to 0$,*

(a) $\max_{i,j \leq N} \left| \frac{1}{T_H} \sum_{t_h=1}^{T_H} u_{i,t_h} u_{j,t_h} - \sigma_{u,ij} \right| = O_p\left(\sqrt{\frac{\log N}{T_H}} \right)$.

(b) $\max_{i \leq N} \left\| \frac{1}{T_H} \sum_{t_h=1}^{T_H} f_{t_h} u_{i,t_h} \right\| = O_p\left(\sqrt{\frac{\log N}{T_H}} \right)$.

Assumptions 1 and 2 are similar to the Assumptions 6 and 7 in GP (2014), except that we need higher moments in Assumption 1. We require a large number of moments because our proof relies on repeated applications of Cauchy-Schwarz's inequality to prove the validity of our bootstrap method under cross-sectional and serial dependence. If we further assume that the factors, factor loadings, and idiosyncratic error terms are mutually independent, then having $E\|\lambda_i\|^8 \leq M$, $E\|f_{t_h}\|^8 \leq M$, and $E|e_{i,t_h}|^{16} \leq M$ are sufficient. Assumption 2 justifies that we use wild bootstrap in the second step as the regression error term is a martingale difference sequence. This assumption can be relaxed to allow for serial correlation in the regression error term and block-based bootstrap can be applied as explained in Djogbenou et al. (2015). Furthermore, in Assumption 3, we assume that idiosyncratic error term is a stationary autoregressive (AR) process of infinite order with polynomial decaying coefficients. In the proof of Section 3 (see Section C in Online Appendix), we show that $r = 4$ is sufficient. Finally, Assumption 4 and Assumption 5 are similar to the CS and TS assumptions in Gonçalves and Perron (2020) (on the idiosyncratic error terms) and Gonçalves et al. (2024) (on the innovations of the idiosyncratic error terms). We assume that the variance-covariance matrix of the innovation terms is time-invariant and the innovation terms are weakly dependent in cross-sectional dimension. Under these additional assumptions, we show the validity of the AR-sieve +CSD bootstrap method in the following theorem.

Theorem 3.1 *Suppose that autoregressive sieve with CSD (AR-sieve + CSD) bootstrap and wild bootstrap are used to generate $\{e^*_{i,t_h}\}$ and $\{\varepsilon^*_t\}$, respectively with $E^*|\eta_{i,t_h}|^4 < C$ for all (i, t_h) and $E^*|\nu_t|^{4q} < C$ for all t, for some $q > 1$. If Assumptions A.1 - A.6 in Section A in Online Appendix and Assumptions 1 - 5 hold,*

$$\sup_{x\in\mathbb{R}^{r+p}} |P^*(\sqrt{T}(\Phi^*_0\tilde{\alpha}^* - \tilde{\alpha}) \leq x) - P(\sqrt{T}(\tilde{\alpha} - \alpha) \leq x)| \xrightarrow{p} 0,$$

where $\Phi^*_0 = diag(1, H^*_0, I_p)$ with $H^*_0 = \operatorname{plim} H^*$ and $H^* = \tilde{V}^{*-1}\frac{\tilde{f}^{*\prime}\tilde{f}}{T_H}\frac{\tilde{\Lambda}'\tilde{\Lambda}}{N}$, which is a bootstrap analogue of rotation matrix, H.

4 Monte Carlo Simulation

In this section, we confirm the presence of bias in the factor-MIDAS regression models, and show the finite sample performance of both inference methods we propose. The data generating process (DGP) is similar to GP (2014) and Aastveit, Foroni, and Ravazzolo (2017). We consider the factor-MIDAS regression model with a single factor model as follows.

$$y_t = \beta_0 + \beta_1 \sum_{k=1}^{K} w_k(\theta) f_{t-k/m} + \varepsilon_t, \tag{16}$$

$$X_{i,t-k/m} = \lambda_i f_{t-k/m} + e_{i,t-k/m}, \; k = m-1, \ldots, 0. \tag{17}$$

For a weighting function, $w_k(\theta)$, for $k = 1, \ldots, K$, we use the exponential Almon lag with two parameters, (2).

The factors and factor loadings are generated similarly to GP (2014). The single factor f_t is randomly drawn from a standard normal distribution independently over time. The factor loading, λ_i is randomly drawn from a uniform distribution of the interval $[0, 1]$ independently across indicators, i. We consider that the high-frequency variable is observed at most 3 times between $t-1$ and t (equivalent to low-frequency data being quarterly and high-frequency data being monthly), which implies $m = 3$. The parameters are $\beta_0 = 0$, $\beta_1 = 2.5$, $\theta_1 = 0.007$, and $\theta_2 = -0.01$. We choose the weighting parameters similar to Aastveit et al. (2017) to induce fast-decaying weights.

Table 1 shows six different scenarios to generate the idiosyncratic error terms and MIDAS regression error terms. We consider the error term in the regression model to be either homoskedastic or heteroskedastic. In DGP 1, we consider homoskedastic error term and in the rest of the DGPs, the error terms are conditionally heteroskedastic. When they are homoskedastic, the errors are drawn independently and identically from a standard nor-

Table 1: Data generating process

DGP	ε_t	e_{i,t_h}
1	$N(0,1)$	$N(0,1)$
2	$\varepsilon_t = \sqrt{h_t} v_t$	$N(0,1)$
3	$\varepsilon_t = \sqrt{h_t} v_t$	$N(0,\sigma_i^2)$
4	$\varepsilon_t = \sqrt{h_t} v_t$	AR + $N(0,\sigma_i^2)$
5	$\varepsilon_t = \sqrt{h_t} v_t$	CS + $N(0,1)$
6	$\varepsilon_t = \sqrt{h_t} v_t$	CS + AR

where $h_t = 0.1 + 0.3\varepsilon_{t-1}^2 + 0.6h_{t-1}$ and $\nu_t \sim$ i.i.d.$N(0,1)$ for $t = 1,\dots,T$ and $t_h = 1,\dots,T_H$.

mal distribution. To allow for heteroskedasticity, we assume that the error terms follow a GARCH model, which implies that they are conditionally heteroskedastic but unconditionally homoskedastic. Particularly, we use the same process as in Aastveit et al. (2017): $\varepsilon_t = \sqrt{h_t} v_t$, where $h_t = 0.1 + 0.3\varepsilon_{t-1}^2 + 0.6h_{t-1}$ and $v_t \sim$ i.i.d.$N(0,1)$.

For the idiosyncratic term in the factor model, we use the same data-generating process in GP (2014). In DGP 1 and DGP 2, the idiosyncratic error terms are homoskedastic by randomly generating them from a standard normal distribution. DGP 3 induces heteroskedasticity in the idiosyncratic term, where the variance for each indicator is drawn from $U[0.5, 1.5]$. DGP 4 introduces the serial correlation by generating the idiosyncratic term from an autoregressive model of order one such that $e_{i,t_h} = \rho_i e_{i,t_h-1} + u_{i,t_h}$, where $u_{i,t_h} \sim$ i.i.d.$N(0,1)$. For simplicity, we let $\rho_i = \rho$ for all $i = 1,\dots,N$, and $\rho = 0.5$. The idiosyncratic terms are re-scaled by $(1-\rho^2)^{1/2}$ so that the variance of the idiosyncratic error terms is 1. DGP 5 allows for cross-sectional dependence in the homoskedastic idiosyncratic terms as in GP (2014) and Bai and Ng (2006). Precisely, we let the correlation between e_{i,t_h} and e_{j,t_h} be $0.5^{|i-j|}$ for $|i-j| \leq 5$ and 0 for otherwise. In DGP 6, the idiosyncratic error terms have both serial and cross-sectional dependence. The idiosyncratic error terms follow the autoregressive process of order 1 with the innovation term being cross-sectionally correlated. The idiosyncratic terms in DGP 5 and 6 are also re-scaled to have the variance 1, the same as in other designs.

To focus on the bias, which arises by the fact that the factors are estimated, we do not estimate the number of the factors in the estimation process. Instead, we assume that we know that there is a single factor. We report the size of the bias in a slope coefficient for the single factor, β_1. Mainly, we report two sets of results: based on asymptotic theory and based on the bootstrap method. The bias based on asymptotic theory is reported when we use the true factor, the estimated factor, and the plug-in bias estimator. We also impose that we know $Cov(e_{i,t_h}, e_{i,t_h-k}) = 0$ for $k > 1$, and therefore we only compute the plug-in bias estimator up to the first degree covariance term. The other set of results includes the bias based on two different bootstrap methods: wild bootstrap (WB) and AR-sieve+CSD bootstrap. For AR-sieve+CSD bootstrap, we choose a lag order for each series by AIC. Note that the wild bootstrap is only valid when the idiosyncratic error terms do not have serial and cross-sectional dependence, DGP 1 - 3. For the rest of the designs, the wild bootstrap is not valid. Therefore, under more general settings (DGP 4 - 6), we can quantify the cost of not accounting for either time-series or cross-sectional dependence or both in the idiosyncratic error term by comparing two bootstrap methods.

To compute the size of bias, we use the approach described in GP (2014). The bias in the original sample is calculated as the average of $H\tilde{\beta}_1 - \beta_1$. This guarantees each estimator in the replication to be consistent for β_1. In the bootstrap world, similarly, we compute the bias of the bootstrap estimator as the average of $HH^*\tilde{\beta}_1^* - H\tilde{\beta}_1$. We also report the 95% coverage rate for the associated estimators: estimated factors, plug-in bias and two bootstrap methods. The coverage rates associated with the bootstrap methods are reported by using the bootstrap equal-tailed percentile-t method.

All our simulation results are based on 5000 replications and 399 bootstraps. We consider $T = 50, 100, 200$ and $N = 50, 100, 200$. Since the high frequency variable is observed $m = 3$ times more, the time-series dimensions in the factor model as 150, 300, and 600, respectively. We choose $K = 11$, which implies that a low-frequency variable can be explained by 11 lagged monthly factors.

Since the results of DGP 1 - 3 are very similar, we leave the results of DGP 1 - 2 in the Online Appendix. The results of DGP 3 and 4 are presented in Table 2. In both scenarios, the MIDAS regression error terms are now heteroskedastic for both DGPs. The idiosyncratic error terms are heteroskedastic in DGP 3. We find that there exists a bias when we use the estimated factor and the plug-in estimator overestimates the magnitude of the bias, especially in small samples. Both bootstrap methods outperform the plug-in estimator in terms of replicating the bias size and correcting the distortion. In DGP 4, the idiosyncratic error terms exhibit not only heteroskedasticity but also display serially dependence. In contrast to DGP 3, the bias size increases as we introduce serial dependence in the error term of the factor model, and it is about twice as large as that in DGP 3. This is consistent with the asymptotic bias result in Theorem 2.1, where time-series dependence contributes to the bias. The plug-in bias is no longer overestimating the bias size.[13]

Comparing the two bootstrap methods, it is evident that AR-sieve+CSD bootstrap method performs better than the wild bootstrap method in DGP 4 - 6. Note that the wild bootstrap is no longer valid under serial dependence. In fact, for some sample sizes, the

[13] It is important to note that since the bias depends on the serial dependence, the persistence in the idiosyncratic error term may also have an impact. We have observed that with an increase in persistence, the bias also increases (documented in Table 1 in Section D in Online Appendix).

Table 2: DGP 3 & DGP 4 - Bias and coverage rate of 95% CIs for β

		$N=50$			$N=100$			$N=200$		
		$T=50$	100	200	50	100	200	50	100	200
		$T_H=150$	300	600	150	300	600	150	300	600
		bias								
DGP 3: hetero & hetero	True Factor	0.00	-0.01	0.00	-0.01	0.00	0.00	0.01	0.00	0.00
	Estimated Factor	-0.37	-0.34	-0.32	-0.22	-0.19	-0.17	-0.12	-0.11	-0.10
	Plug-in	-0.41	-0.36	-0.35	-0.22	-0.20	-0.19	-0.11	-0.11	-0.10
	WB	-0.27	-0.26	-0.26	-0.17	-0.16	-0.15	-0.11	-0.10	-0.09
	AR-sieve+CSD	-0.26	-0.26	-0.25	-0.17	-0.16	-0.15	-0.11	-0.10	-0.09
		95% coverage rate								
	Estimated Factor	75.0	72.6	63.9	85.0	85.5	84.4	88.5	90.3	91.0
	Plug-in	80.9	87.9	88.9	86.8	89.3	92.1	88.9	91.1	92.5
	WB	91.7	94.2	92.7	92.6	93.5	94.1	91.3	93.9	93.8
	AR-sieve+CSD	93.7	92.1	90.4	93.6	94.3	94.1	94.1	95.1	93.6
		bias								
DGP 4: hetero & AR	True Factor	0.00	0.00	0.00	-0.01	0.00	0.00	-0.01	0.00	0.00
	Estimated Factor	-0.64	-0.57	-0.54	-0.41	-0.35	-0.31	-0.28	-0.21	-0.18
	Plug-in	-0.45	-0.42	-0.41	-0.26	-0.26	-0.25	-0.14	-0.14	-0.14
	WB	-0.22	-0.22	-0.22	-0.15	-0.14	-0.14	-0.10	-0.09	-0.08
	AR-sieve+CSD	-0.38	-0.37	-0.36	-0.29	-0.26	-0.25	-0.22	-0.18	-0.16
		95% coverage rate								
	Estimated Factor	52.2	44.5	29.2	72.3	71.8	67.3	81.5	85.0	84.1
	Plug-in	72.0	77.1	77.1	81.1	86.0	87.9	85.0	90.1	91.3
	WB	82.8	79.4	68.7	89.0	88.8	86.1	89.6	92.4	91.3
	AR-sieve+CSD	88.7	87.4	81.4	91.9	91.9	91.3	93.6	94.9	93.5

In DGP 3, both error terms are heteroskedastic. In DGP 4, the idiosyncratic error term is generated as the autoregressive process of lag 1 for each variable and with heteroskedastic. For coverage rates, the results for estimated factors and plug-ins are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

wild bootstrap even performs worse than the plug-in bias, when it comes to compare the size of the bias. We can also confirm that the AR-sieve+CSD bootstrap procedure outperforms the plug-in bias and wild bootstrap procedure by comparing the results of coverage rates, particularly in small sample sizes.

Table 3: DGP 5 & DGP 6 - Bias and coverage rate of 95% CIs for β

		$N=50$			$N=100$			$N=200$		
		$T=50$	100	200	50	100	200	50	100	200
		$T_H=150$	300	600	150	300	600	150	300	600
	bias									
DGP 5: hetero & CSD	True Factor	0.00	-0.01	0.00	-0.01	0.00	0.00	0.01	0.00	0.00
	Estimated Factor	-0.37	-0.34	-0.32	-0.22	-0.19	-0.17	-0.12	-0.11	-0.10
	Plug-in	-0.41	-0.36	-0.35	-0.22	-0.20	-0.19	-0.11	-0.11	-0.10
	WB	-0.10	-0.10	-0.10	-0.06	-0.06	-0.04	-0.04	-0.04	-0.03
	AR-sieve+CSD	-0.16	-0.16	-0.16	-0.10	-0.10	-0.10	-0.06	-0.06	-0.06
	95% coverage rate									
	Estimated Factor	75.0	72.6	63.9	85.0	85.5	84.4	88.5	90.3	91.0
	Plug-in	80.9	87.9	88.9	86.8	89.3	92.1	88.9	91.1	92.5
	WB	88.7	86.2	79.5	92.7	92.6	90.0	94.2	93.5	93.5
	AR-sieve+CSD	90.9	90.0	87.0	93.3	94.1	92.3	94.3	93.9	93.7
	bias									
DGP 6: hetero & CSD+AR	True Factor	0.00	0.00	0.00	-0.01	0.00	0.00	-0.01	0.00	0.00
	Estimated Factor	-0.64	-0.57	-0.54	-0.41	-0.35	-0.31	-0.28	-0.21	-0.18
	Plug-in	-0.45	-0.42	-0.41	-0.26	-0.26	-0.25	-0.14	-0.14	-0.14
	WB	-0.08	-0.09	-0.08	-0.06	-0.06	-0.05	-0.04	-0.03	-0.03
	AR-sieve+CSD	-0.23	-0.23	-0.24	-0.17	-0.16	-0.16	-0.12	-0.10	-0.10
	95% coverage rate									
	Estimated Factor	52.2	44.5	29.2	72.3	71.8	67.3	81.5	85.0	84.1
	Plug-in	72.0	77.1	77.1	81.1	86.0	87.9	85.0	90.1	91.3
	WB	76.5	66.2	47.4	87.5	84.2	77.6	91.1	91.5	89.3
	AR-sieve+CSD	86.3	80.0	73.5	91.0	89.8	87.1	93.2	93.2	92.6

In DGP 5 and 6, both error terms are heteroskedastic. In DGP 5, the idiosyncratic error term contains the cross-sectional dependence. In DGP 6, we impose the dependence in both dimensions for the idiosyncratic error terms. For coverage rates, the results for estimated factors and plug-in are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

Finally, we present the results of DGP 5 and 6, which are shown in Table 3. In DGP 5, the idiosyncratic error term is only cross-sectionally correlated. Comparing the size of the bias, the AR-sieve+CSD bootstrap performs better than the wild bootstrap method but worse than the plug-in bias method. The AR-sieve+CSD bootstrap method recovers the size distortion better than the plug-in method in most of the cases. The plug-in estimation method performs better than the AR-sieve+CSD bootstrap method when $N = 50$ and $T = 200$. In DGP 6, we allow for cross-sectional dependence as well as serial dependence in the idiosyncratic error terms. The results follow a similar pattern to the findings of DGP 5. The plug-in bias method replicates the bias better than bootstrap methods. However, it does worse than AR-sieve+CSD bootstrap in terms of recovering the size distortion in the coverage rates except when $T = 200$. Furthermore, when the time series dimension is as small as 50, the plug-in bias method performs even worse than the wild bootstrap method, which is not valid in this design. Overall, the AR-sieve+CSD bootstrap works well in correcting the distortion.[14]

5 Empirical Application

In this section, we apply the factor-MIDAS regression model to validate the presence of bias in an empirical example. It is well documented that incorporating high-frequency indicators to forecast a quarterly variable using the MIDAS regression model improves the forecast performance (e.g., see Clements and Galvão (2008; 2009), Aastveit et al. (2017), Marcellino and Schumacher (2010), Andreou, Ghysels, and Kourtellos (2013), and Beyhum and Striaukas (2024)).

In this paper, we focus on nowcasting quarterly U.S. real GDP growth using monthly

[14] Similar findings can be found when the AR-sieve+CSD bootstrap is used in the context of the unrestricted MIDAS regression model. The performance of AR-sieve+CSD bootstrap dominates the plug-in bias estimation method in all DGPs. See Table 4 - 6 in Section D in Online Appendix.

macroeconomic factors from 1984 Q1 to 2022 Q4 including great moderation period. We have divided this period into two: the long period (1984 Q1 - 2022 Q4), which includes the COVID pandemic period, and the short period (1984 Q1 to 2019 Q4). Although we look into two different periods, the results are very similar; therefore, we present the results for the shorter period in the Online Appendix. Our nowcasting model is similar to the model in Beyhum and Striaukas (2024). Given the number of leading months, $l = 1, 2, 3$, we write our model as follows.

$$y_t = \beta_0 + \sum_{i=1}^{p_y} \rho_i y_{t-i} + \beta_1' \sum_{k=1-l}^{K-l} w_{(k-1)+l}(\theta) f_{t-1-(j-1)/m} + \varepsilon_t, \tag{18}$$

where y_t is quarterly U.S. GDP growth rate. We denote common factors containing timely information about monthly macroeconomic predictors by $f_{t-k/m}$. The number of leading months represents a nowcasting horizon, denoted by h. For instance, $l = 1$ indicates that we exploit information of one leading month; hence, we nowcast two months away ($h = 2$). We use the exponential Almon lag with two parameters defined in (2) for the lag polynomial function. The quarterly U.S. output is obtained from a FRED-QD dataset (for detail, see M. McCracken and Ng (2020)). As U.S. real output is available in level in the dataset, we compute the growth rate in percentage, by $\{\ln(\text{GDP})_t - \ln(\text{GDP})_{t-1}\} \times 100$. We also include the lags of the growth rate in the regression. The number of lags of the dependent variable is chosen by BIC, before we apply MIDAS regression. BIC selects one lag in the long period and three lags in the short period.

To estimate the monthly factors, we utilize the FRED-MD dataset[15] (for detail, see M. W. McCracken and Ng (2016)). We consider 74 macroeconomic variables available for the entire period and exclude all financial variables. Using PCA, we extract two common factors

[15] We use the 'current' version downloaded on October 3rd, 2023.

in both periods. The information criterion proposed by Bai and Ng (2002) (particularly, IC_p) chooses eight factors in the long period and five factors in the short period. Although the information criterion chooses more than 2 factors, the two factors we extract explain more than 60% of the variability explained by all the factors chosen by the information criterion proposed by Bai and Ng (2002).

Our primary goal is to verify the existence of bias in the estimators. Instead of focusing solely on the forecasting performance of the factor-MIDAS regression model, we aim to examine the behaviour of the estimators, particularly their 90% confidence interval. We present three sets of confidence intervals, one based on asymptotic theory and the other two based on the bootstrap method. We use two different bootstrap methods for resampling the idiosyncratic error terms in the factor model: wild bootstrap and AR-sieve + CSD bootstrap, described in Section 3. We also rotate the bootstrap estimators, $\tilde{\beta}_1^*$, with the rotation matrix H^* as in GP (2014) and Gonçalves and Perron (2020).

In Table 4, we present the confidence interval for the point estimates in the long period, 1984 Q1 - 2022 Q4 for each nowcasting horizon, $h = 2, 1$, and 0. We also report the estimate associated with each parameter on the top of the three confidence intervals. We can find that there exists a bias in the estimators associated with the factors. For example, the point estimate associated with the first factor for horizon $h = 2$ is 2.54. The confidence interval of this estimate is centered around 2.54, but the bootstrap interval shifts to the right, suggesting a negative bias. The results are similar for the other horizons, $h = 1$ and 0. Although the second factor is not significant at $h = 2$, we can confirm that there exists a bias in the estimator associated with the second factor at $h = 1$ and $h = 0$. When $h = 1$, the result implies a negative bias, whereas when $h = 0$, there exists a positive bias, shifting the interval to the left. Comparing the two bootstrap methods, there is a small change in the

Table 4: Estimates in the long period (1984 Q1 - 2022 Q4)

		$h=2$		$h=1$		$h=0$	
constant		0.90		0.83		0.99	
	Asymptotic	0.67	1.01	0.67	0.99	0.78	1.21
	WB	0.71	0.98	0.69	0.95	0.73	1.28
	AR sieve+CSD	0.71	0.98	0.69	0.94	0.75	1.26
first factor		2.54		3.79		1.87	
	Asymptotic	1.64	3.44	2.97	4.61	0.31	3.44
	WB	2.01	3.56	3.29	4.72	0.91	3.93
	AR sieve+CSD	2.13	3.54	3.34	4.80	0.90	3.39
second factor		0.04		0.36		-0.95	
	Asymptotic	-0.22	0.30	0.08	0.65	-1.47	-0.43
	WB	-0.17	0.37	0.14	0.75	-1.62	-0.01
	AR sieve+CSD	-0.12	0.38	0.16	0.77	-1.63	-0.21
y_{t-1}		-0.30		-0.30		-0.58	
	Asymptotic	-0.54	-0.06	-0.52	-0.09	-0.87	-0.28
	WB	-0.49	-0.12	-0.44	-0.14	-1.25	-0.26
	AR sieve+CSD	-0.49	-0.12	-0.43	-0.14	-1.22	-0.25

The interval based on the asymptotic theory is obtained by adding and subtracting 1.645 times the heteroskedasticity robust standard errors. The confidence intervals based on bootstrap methods are obtained with equal-tailed bootstrap intervals with a bootstrap number 4999. WB indicates that we use wild bootstrap and AR sieve + CSD indicates that we use the bootstrap algorithm described in Section 3.

bootstrap confidence intervals of the estimators associated with the two factors. However, the difference is not huge, indicating that the serial and cross-sectional dependence in this example may be small.

6 Conclusion

In this paper, we derive the asymptotic distribution of the estimators in the factor-augmented MIDAS regression models. We find that there exists an asymptotic bias arising from the fact that the factors are latent and must be estimated. We show that the bias depends on the serial dependence as well as the cross-sectional dependence of the idiosyncratic error term in the factor model, because MIDAS temporally aggregates the factors and their lags. We propose two inference methods that account for this bias: an analytical bias estimator based on the bias formula derived and a bootstrap method. Both inference methods are robust to serial and cross-sectional dependence.

Although our simulation results support the theoretical findings, the bootstrap method more effectively corrects the size distortion in the coverage rates, while the plug-in method outperforms the bootstrap method in estimating the size of the bias, especially in small samples. We further apply the factor-MIDAS regression model to nowcast quarterly U.S. GDP growth rate using monthly macroeconomic factors. Our empirical results indicate that there exists a bias in the estimates associated with the estimated factors.

Our results can be extended to the context of forecasting, such as to construct forecast intervals, similar to Gonçalves, Perron, and Djogbenou (2017), where they construct it in the context of the factor-augmented regression models without mixed frequency datasets. By letting $\hat{y}_{T+1} = g(\tilde{F}_T, \tilde{\alpha})$ be the forecast of y_{T+1} based on information up to time T, we can decompose the forecast error as

$$\hat{y}_{T+1} - y_{T+1} = -\varepsilon_{T+1} + \frac{1}{\sqrt{T}} \frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha'} \sqrt{T}(\tilde{\alpha} - \alpha) + \frac{1}{\sqrt{N}} \beta' H^{-1} \sqrt{N}(\tilde{F}_t(\theta) - H F_t(\theta)) + o_p(1).$$

This underscores the importance of the asymptotic distribution of the estimators derived in this paper in constructing to construct the forecast interval. We leave this for future research.

An interesting extension involves the use of machine-learning techniques. Machine learning techniques are popularly used to handle high-dimensional data. Along the same lines, Babii, Ghysels, and Striaukas (2022) propose a machine learning regression by applying the sparse-group LASSO technique for mixed-frequency data.

Acknowledgement

We would like to thank John Galbraith, Eric Ghysels, Sílvia Gonçalves, and Benoit Perron for the advice. We also want to thank participants at the Canadian Economic Association (2022), International Symposium on Econometrics Theory and Applications (2022), and Canadian Econometrics Study Group (2023).

Computations were made on the supercomputer Beluga from McGill University, managed by Calcul Québec and the Digital Research Alliance of Canada.

References

Aastveit, K. A., Foroni, C., & Ravazzolo, F. (2017). Density forecasts with midas models. *Journal of Applied Econometrics*, *32*(4), 783-801.

Andreou, E., Gagliardini, P., Ghysels, E., & Rubin, M. (2019). Inference in group factor models with an application to mixed-frequency data. *Econometrica*, *87*(4), 1267–1305.

Andreou, E., Ghysels, E., & Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, *158*(2), 246-261.

Andreou, E., Ghysels, E., & Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, *31*(2), 240–251.

Babii, A., Ghysels, E., & Striaukas, J. (2022). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics*, *40*(3), 1094–1106.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.

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

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

Beyhum, J., & Striaukas, J. (2024). Testing for sparse idiosyncratic components in factor-augmented regression models. *Journal of Econometrics*, *244*(1), 105845.

Bi, D., Shang, H. L., Yang, Y., & Zhu, H. (2021). *Ar-sieve bootstrap for high-dimensional time series.* Retrieved from https://arxiv.org/abs/2112.00414

Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli*, 123–148.

Clements, M. P., & Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the united states. *Journal of Business & Economic Statistics*, *26*(4), 546–554.

Clements, M. P., & Galvão, A. B. (2009). Forecasting us output growth using leading indicators: An appraisal using midas models. *Journal of Applied Econometrics*, *24*(7), 1187–1206.

Djogbenou, A., Gonçalves, S., & Perron, B. (2015). Bootstrap inference in regressions with estimated factors and serial correlation. *Journal of Time Series Analysis*, *36*(3), 481–502.

Ferrara, L., & Marsilli, C. (2019). Nowcasting global economic growth: A factor-augmented mixed-frequency approach. *The World Economy*, *42*(3), 846-875.

Foroni, C., Marcellino, M., & Schumacher, C. (2015). Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *178*(1), 57–82.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004, 06). The midas touch: mixed data sampling regression.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, *76*(3), 509–548.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, *131*(1-2), 59–95.

Ghysels, E., Sinko, A., & Valkanov, R. (2007, 02). Midas regressions: Further results and new directions. *Econometric Reviews*, *26*, 53-90.

Ghysels, E., Valkanov, R. I., & Serrano, A. R. (2009). Multi-period forecasts of volatility: Direct, iterated, and mixed-data approaches. In *Efa 2009 bergen meetings paper.*

Gonçalves, S., & Perron, B. (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics*, *182*(1), 156–173.

Gonçalves, S., & Perron, B. (2020). Bootstrapping factor models with cross sectional dependence. *Journal of Econometrics*, *218*(2), 476–495.

Gonçalves, S., Perron, B., & Djogbenou, A. (2017). Bootstrap prediction intervals for factor models. *Journal of Business & Economic Statistics*, *35*(1), 53–69.

Gonçalves, S., Koh, J., & Perron, B. (2024, 11). Bootstrap inference for group factor models.

Journal of Financial Econometrics, nbae020.

Kim, H. H., & Swanson, N. R. (2018). Methods for backcasting, nowcasting and forecasting using factor-midas: With an application to korean gdp. *Journal of Forecasting*, *37*(3), 281–302.

Kock, A. B., & Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, *186*(2), 325-344. (High Dimensional Problems in Econometrics)

Krampe, J., Kreiss, J.-P., & Paparoditis, E. (2021). Bootstrap based inference for sparse high-dimensional time series models. *Bernoulli*, *27*(3), 1441 – 1466.

Kreiss, J.-P., Paparoditis, E., & Politis, D. N. (2011). On the range of validity of the autoregressive sieve bootstrap. *The Annals of Statistics*, *39*(4), 2103–2130.

Ludvigson, S. C., & Ng, S. (2009). *A factor analysis of bond risk premia* (Tech. Rep.). National Bureau of Economic Research.

Marcellino, M., & Schumacher, C. (2010). Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics*, *72*(4), 518–550.

McCracken, M., & Ng, S. (2020). *Fred-qd: A quarterly database for macroeconomic research* (Tech. Rep.). National Bureau of Economic Research.

McCracken, M. W., & Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, *34*(4), 574–589.

Meyer, M., & Kreiss, J.-P. (2015). On the vector autoregressive sieve bootstrap. *Journal of Time Series Analysis*, *36*(3), 377-397.

Monteforte, L., & Moretti, G. (2013). Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, *32*(1), 51–61.

Online Appendix for "Inference for Factor-MIDAS Regression Models"

Abstract

Appendix A presents the primitive assumptions necessary for proving the results in the main text. Appendices B and D provide the proofs of the results in the main text. Appendix C presents the bootstrap procedure for the factor-MIDAS regression model. Appendix E contains additional simulation results. Finally, in Appendix F, we include an additional empirical result, which is omitted from the main text.

A Primitive assumptions

This section delivers the primitive assumption for asymptotic theory. The factor-augmented MIDAS regression involves two frequencies, thus we use two time indices: $t_h = 1, \ldots, T_H$ denotes the high-frequency time index and $t = 1, \ldots, T$ denotes the low-frequency time index. Particularly, we use a subscript h to denote high-frequency time index (e.g. s_h also denotes the high-frequency time index).

Assumption A.1 (Factors and Factor Loadings)

(a) f_{t_h} *are stationary with* $E\left\|f_{t_h}\right\|^4 \leq M$ *and* $\frac{1}{T_H}\sum_{t_h=1}^{T_H} f_{t_h} f'_{t_h} \xrightarrow{p} \Sigma_f > 0$*, where* Σ_f *is a non-random* $r \times r$ *matrix.*

(b) The factor loadings λ_i *are either deterministic such that* $\|\lambda_i\| \leq M$*, or stochastic such that* $E\|\lambda_i\|^4 \leq M$*. In either case,* $\Lambda'\Lambda/N \xrightarrow{p} \Sigma_\Lambda > 0$*, where* Σ_Λ *is a non-random matrix.*

(c) The eigenvalues of the $r \times r$ *matrix* $(\Sigma_\Lambda \Sigma_f)$ *are distinct.*

(d) $f'f/T_H = I_r$ *and* $\Lambda'\Lambda$ *is a diagonal matrix with distinct entries, where* $f = (f_1, \ldots, f_{T_H})'$*.*

Assumption A.2 (Time and Cross Section Dependence and Heteroskedasticity)

(a) $E(e_{i,t_h}) = 0, E|e_{i,t_h}|^8 \leq M$*.*

(b) $E(e_{i,t_h} e_{j,s_h}) = \sigma_{ij,t_h s_h}, |\sigma_{ij,t_h s_h}| \leq \bar{\sigma}_{ij}$ *for all* (t_h, s_h) *and* $|\sigma_{ij,t_h s_h}| \leq \tau_{t_h s_h}$ *for all* (i,j) *such that* $\frac{1}{N}\sum_{i,j=1}^{N} \bar{\sigma}_{ij} \leq M$*,* $\frac{1}{T_H}\sum_{t_h,s_h=1}^{T_H} \tau_{t_h s_h} \leq M$*, and* $\frac{1}{NT_H}\sum_{t_h,s_h,i,j} |\sigma_{ij,t_h s_h}| \leq M$*.*

(c) For every (t_h, s_h)*,* $E\left|N^{-1/2}\sum_{i=1}^{N}(e_{i,t_h} e_{i,s_h} - E(e_{i,t_h} e_{i,s_h}))\right|^4 \leq M$*.*

(d) $E(e_{i,t_h}e_{j,t_h}) = \sigma_{ij}$ *and* $E(e_{i,t_h}e_{j,t_h-k}) = \sigma_{ij,k}$ *for all* t *and* k.

Assumption A.3 (Moments and Weak Dependence Among $\{f_{t_h}\}, \{\lambda_i\}$ and $\{e_{i,t_h}\}$)

(a) $E\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T_H}}\sum_{t_h=1}^{T_H} f_{t_h}e_{i,t_h}\right\|^2\right) \le M$, *where* $E(f_{t_h}e_{i,t_h}) = 0$ *for all* (i, t_h).

(b) *For each* t_h, $E\left\|\frac{1}{\sqrt{T_H N}}\sum_{s_h=1}^{T_H}\sum_{i=1}^{N} f_{s_h}(e_{i,t_h}e_{i,s_h} - E(e_{i,t_h}e_{i,s_h}))\right\|^2 \le M$.

(c) $E\left\|\frac{1}{\sqrt{T_H N}}\sum_{t_h=1}^{T_H} f_{t_h}e'_{t_h}\Lambda\right\|^2 \le M$, *where* $E(f_{t_h}\lambda'_i e_{i,t_h}) = 0$ *for all* (i, t_h).

(d) $E\left(\frac{1}{T_H}\sum_{t_h=1}^{T_H}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{i,t_h}\right\|^2\right) \le M$, *where* $E(\lambda_i e_{i,t_h}) = 0$ *for all* (i, t_h).

(e) *As* $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda'_j e_{i,t_h}e_{j,t_h} - \Gamma \xrightarrow{p} 0$ *and* $\Gamma \equiv \lim_{N\to\infty} Var\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{i,t_h}\right)$.

Assumption A.4 (Serial Dependence between $\{f_{t_h}\}$, $\{\lambda_i\}$ and $\{e_{i,t_h}\}$)

(a) $\frac{1}{T_H}\sum_{t_h=1}^{T_H} f_{t_h}f'_{t_h-k} \xrightarrow{p} \Sigma_{f,k}$, *where* $\Sigma_{f,k}$ *is a non-random* $r \times r$ *matrix.*

(b) *For each* t_h *and all* k, $E\left\|\frac{1}{\sqrt{T_H N}}\sum_{s_h=1}^{T_H}\sum_{i=1}^{N} f_{s_h}(e_{i,t_h}e_{i,s_h-k} - E(e_{i,t_h}e_{i,s_h-k}))\right\|^2 \le M$.

(c) $E\left\|\frac{1}{\sqrt{NT_H}}\sum_{t_h=1}^{T_H} f_{t_h}e'_{t_h-k}\Lambda\right\|^2 \le M$, *where* $E(f_{t_h}\lambda'_i e_{i,t_h-k}) = 0$ *for all* (i, t_h) *and all* k.

(d) *As* $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda'_j e_{i,t_h}e_{j,t_h-k} - \Gamma_k \xrightarrow{p} 0$ *and* $\Gamma_k \equiv \lim_{N\to\infty} Cov\left(\frac{\Lambda' e_{t_h}}{\sqrt{N}}, \frac{\Lambda' e_{t_h-k}}{\sqrt{N}}\right)$.

Assumption A.5 (Weak Dependence Between Idiosyncratic Errors and Regression Errors)

(a) *For each* t, $E\left|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\varepsilon_s(e_{i,t-j/m}e_{i,s-j/m} - E(e_{i,t-j/m}e_{i,s-j/m}))\right|^2 \le M$ *for* $j = 0, \ldots, m-1$.

(b) $E\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}\sum_{i=1}^{N}\lambda_i e_{i,t-j/m}\varepsilon_t\right\|^2 \le M$, *where* $E(\lambda_i e_{i,t-j/m}\varepsilon_t) = 0$ *for all* (i, t) *and* $j = 0, \ldots, m-1$.

Assumption A.6 (Moments and CLT for the Score Vector)

(a) $E(\varepsilon_t) = 0$ *and* $E\left|\varepsilon_t\right|^2 < M$.

(b) $E\left\|g_{\alpha,t}\right\|^4 \leq M$ *and* $\frac{1}{T}\sum_{t=1}^{T} g_{\alpha,t}g'_{\alpha,t} \xrightarrow{p} \Sigma > 0$ *where* $g_{\alpha,t} = \partial g(F_t, \alpha)/\partial\alpha$.

(c) *As* $T \to \infty$, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{\alpha,t}\varepsilon_t \xrightarrow{d} N(0, \Omega)$, *where* $E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{\alpha,t}\varepsilon_t\right\|^2 < M$ *and* $\Omega \equiv \lim_{T\to\infty} Var\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{\alpha,t}\varepsilon_t\right) > 0$.

Assumption A.1 are standard assumptions on the factors and the factor loadings in the factor analysis. Additionally, we assume that the factors are stationary. This is to allow $\Sigma_f = \text{plim}\,\frac{1}{T_H}\sum_{t_h=1}^{T_H} f_{t_h}f'_{t_h} = \text{plim}\,\frac{1}{T}\sum_{t=1}^{T} f_{t-j/m}f'_{t-j/m}$, for all j. Assumption A.1-(d) is one of the identifying restrictions from Bai and Ng (2013). By imposing this assumption, the rotation matrix H_0 is a diagonal matrix of ± 1, where the sign is determined by $\tilde{f}'f/T_H$. However, since the true factors are unknown, we still do not know the sign of the rotation matrix.

Assumption A.2 and Assumption A.3 can be found equivalently in Gonçalves and Perron (2014) (henceforth, GP (2014)) (their Assumptions 2 and 3, respectively). In Assumption A.2, we allow weak cross-sectional and serial dependence in the idiosyncratic error terms. In Assumption A.3, we impose some moment condition between the factors, idiosyncratic error terms, and the factor loadings. We also allow some weak dependence among them. Due to the MIDAS structure, where the lags of the factors are used, we also allow some serial dependence between them in Assumption A.4. This set of assumptions is new in the context of the factor-augmented regression models. In particular, Assumption A.4-(d) allows for the serial dependence in the scaled average over cross-sectional dimension of factor loadings and idiosyncratic error term in the factor model.

We impose some weak dependence between idiosyncratic error terms and the regression errors in Assumption A.5. This Assumption is equivalent to the Assumption 4 in GP (2014). Assumption A.6 imposes some moment condition on $\{\varepsilon_t\}$ and the score vector $g_{\alpha,t}$. Assumption A.6-(b) requires that we can apply a law of large numbers on $\{g_{\alpha,t}g'_{\alpha,t}\}$. By introducing Assumption A.6-(c), we can apply a central limit theorem on $\{g_{\alpha,t}\varepsilon_t\}$. Similar assumptions to Assumption A.5 and A.6 can be found in GP (2014).

B Proof of results in Section 2

In this section, we prove the asymptotic distribution of NLS estimators in Theorem 2.1 and Theorem 2.2, the consistency of the variance-covariance of the cross-sectional average of the factor loadings and idiosyncratic error term across time for the plug-in bias estimator. To prove the asymptotic distribution, we use the following lemmas. The proof for the following lemmas Lemma B.1 to Lemma B.3 can be found at the end of proof of Theorem 2.1.

Lemma B.1 $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(\tilde{F}_t(\theta) - HF_t(\theta)) = o_p(1)$.

Lemma B.2 *For* $j, l = 0, \ldots, m-1$, *if* $\sqrt{T}/N \to c$, *where* $0 \leq c < \infty$,

(a) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m} - HF_{t-j/m})(\tilde{f}_{t-j/m} - Hf_{t-j/m})' = cV^{-1}H\Gamma HV^{-1} + o_p(1)$,

(b) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m} - Hf_{t-j/m})(\tilde{f}_{t-l/m} - Hf_{t-l/m})' = cV^{-1}H\Gamma_{j-l}HV^{-1} + o_p(1)$ *for* $j \neq l$,

(c) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Hf_{t-j/m}(\tilde{f}_{t-j/m} - Hf_{t-j/m})' = cH\Gamma Q'V^{-2} + o_p(1)$,

(d) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Hf_{t-l/m}(\tilde{f}_{t-j/m} - Hf_{t-j/m})' = cQ_{j-l}\Gamma Q'V^{-2} + o_p(1)$ *for* $j \neq l$.

Lemma B.3 *If* $\sqrt{T}/N \to c$, *where* $0 \leq c < \infty$,

(a) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{F}_t(\theta)-HF_t(\theta))(\tilde{F}_t(\theta)-HF_t(\theta))'$

$= cV^{-1}Q\left\{\sum_{k=1}^{K}w_k(\theta)\Gamma w_k(\theta)+\sum_{k=1}^{K}\sum_{l\neq k}^{K}w_k(\theta)\Gamma_{k-l}w_l(\theta)\right\}Q'V^{-1}+o_p(1),$

(b) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{F}_t(\theta)-HF_t(\theta))(HF_t(\theta))'$

$= c\left\{\sum_{k=1}^{K}w_k^2(\theta)H+\sum_{k=1}^{K}\sum_{l\neq k}^{K}w_k(\theta)Q_{k-l}w_l(\theta)\right\}\Gamma Q'V^{-2}+o_p(1).$

Note that we write $F_t(\theta)=\sum_{k=1}^{K}w_k(\theta)f_{t-k/m}$, where $w_k(\theta)\equiv \mathrm{diag}(w_{k,1}(\theta_1),\ldots,w_{k,r}(\theta_r))$ is a $r\times r$ diagonal matrix. We also define $\delta_{NT_H}=\min(\sqrt{N},\sqrt{T_H})$. We first prove Theorem 2.1 and then we prove Lemmas B.1 - B.3.

Proof of Theorem 2.1. As the NLS estimators $\tilde{\alpha}$ maximizes the objective function $\tilde{Q}_T(\alpha)=-\frac{1}{T}\sum_{t=1}^{T}[y_t-g(\tilde{F}_t,\alpha)]^2$, we have

$$\sqrt{T}(\tilde{\alpha}-\alpha)=-\Big[\frac{1}{T}\sum_{t=1}^{T}H(\tilde{F}_t,\alpha_T)\Big]^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s(\tilde{F}_t,\alpha), \tag{1}$$

where α_T is the intermediate between $\tilde{\alpha}$ and α and $H(\tilde{F}_t,\alpha)$ is a hessian matrix and $s(\tilde{F}_t,\alpha)$ is a score vector. For deriving the asymptotic distribution, we analyse the convergence of each term. Let $g_\alpha(\cdot)=\partial g(\cdot)/\partial\alpha$. We write the term with a score vector as follows.

$$\begin{aligned}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}s(\tilde{F}_t,\alpha)&=2\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\varepsilon_t+\beta'H^{-1}(HF_t(\theta)-\tilde{F}_t(\theta))]g_\alpha(\tilde{F}_t,\alpha)\\&=2\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\varepsilon_t+\beta'H^{-1}(HF_t(\theta)-\tilde{F}_t(\theta))](\Phi_0 g_\alpha(F_t,\alpha)+P_t),\end{aligned}$$

where where $\Phi_0=\mathrm{diag}(1,H_0,I_p)$ and $H_0=\mathrm{plim}\,H$ and P_t is a $(1+r+p)\times 1$ vector such that

$$P_t=\begin{bmatrix}0\\ \tilde{F}_t(\theta)-HF_t(\theta)\\ \left(\frac{\partial\tilde{F}_t(\theta)}{\partial\theta}H^{-1}-\frac{\partial F_t(\theta)}{\partial\theta}\right)'\beta\end{bmatrix},$$

with $\frac{\partial \tilde{F}_t(\theta)'}{\partial\theta} = \text{diag}\left(\frac{\partial \tilde{F}_{1,t}(\theta_1)}{\partial\theta_1}, \ldots, \frac{\partial \tilde{F}_{r,t}(\theta_r)}{\partial\theta_r}\right)$ is a $r \times r$ block-diagonal matrix. k-th block is $\partial \tilde{F}_{k,t}(\theta_k)/\partial\theta_k$, which is a $p_j \times 1$ column vector, for $j = 1, \ldots, r$. Under Assumption A.6 and Lemma B.1, we have $\frac{1}{\sqrt{T}}\sum_{t=1}^T \varepsilon_t g_\alpha(\tilde{F}_t, \alpha) \xrightarrow{d} N(0, \Phi_0\Omega\Phi_0')$. The remaining term drives the bias in Theorem 2.1. Note that the bias exists in the slope coefficients β_1 and the weighting parameters θ. With respect to β_1, the remaining term is as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{F}_t(\theta)[HF_t(\theta) - \tilde{F}_t(\theta)]'H^{-1'}\beta_1 \\
&= -\left[\frac{1}{\sqrt{T}}\sum_{t=1}^T (\tilde{F}_t(\theta) - HF_t(\theta))(\tilde{F}_t(\theta) - HF_t(\theta))' + \frac{1}{\sqrt{T}}\sum_{t=1}^T HF_t(\theta)(\tilde{F}_t(\theta) - HF_t(\theta))'\right] H^{-1'}\beta_1 \\
&= -c\left[V^{-1}H\left\{\sum_{k=1}^K w_k(\theta)\Gamma w_k(\theta) + \sum_{k=1}^K\sum_{l\neq k}^K w_k(\theta)\Gamma_{k-l}w_l(\theta)\right\}HV^{-1}\right. \\
&\qquad \left. + \left\{\sum_{k=1}^K w_k(\theta)Hw_k(\theta) + \sum_{k=1}^K\sum_{l\neq k}^K w_k(\theta)Q_{k-l}w_l(\theta)\right\}\Gamma Q'V^{-2}\right] \text{plim}(\tilde{\beta}_1) \\
&= -cB_{\beta_1} + o_p(1), \qquad (2)
\end{aligned}
$$

where $\text{plim}(\tilde{\beta}_1) = H^{-1'}\beta_1$. The second equality follows by applying Lemma B.3. Similarly, with respect to θ, we have

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^T \frac{\partial \tilde{F}_t(\theta)'}{\partial\theta}H^{-1'}\beta_1\beta_1'H^{-1}[HF_t(\theta) - \tilde{F}_t(\theta)] \\
&= -H^{-1'}\beta_1 \circ \frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{F}_{t,\theta}(\theta)[\tilde{F}_t(\theta) - HF_t(\theta)]'H^{-1'}\beta_1 \\
&= -c\,\text{plim}(\tilde{\beta}_1) \circ \left[V^{-1}H\left\{\sum_{k=1}^K \frac{\partial w_k(\theta)}{\partial\theta}\Gamma w_k(\theta) + \sum_{k=1}^K\sum_{l\neq k}^K \frac{\partial w_k(\theta)}{\partial\theta}\Gamma_{k-l}w_l(\theta)\right\}HV^{-1}\right. \\
&\qquad \left. + \left\{\sum_{k=1}^K \frac{\partial w_k(\theta)}{\partial\theta}Hw_k(\theta) + \sum_{k=1}^K\sum_{l\neq k}^K \frac{\partial w_k(\theta)}{\partial\theta}Q_{k-l}w_l(\theta)\right\}\Gamma Q'V^{-2}\right] \text{plim}(\tilde{\beta}_1) \\
&= -cB_\theta + o_p(1), \qquad (3)
\end{aligned}
$$

where $\tilde{F}_{t,\theta}(\theta) = \left(\frac{\partial \tilde{F}_{1,t}(\theta_1)}{\partial \theta_1}, \ldots, \frac{\partial \tilde{F}_{r,t}(\theta_r)}{\theta_r}\right)'$. To apply the lemmas, we use the Hadamard product such that $(A \circ B)_{ij} = A_{ij}B_{ij}$. By applying Hadamard product, we have $\frac{\partial \tilde{F}_t(\theta)'}{\partial \theta} H^{-1'}\beta = H^{-1'}\beta \circ \tilde{F}_{t,\theta}(\theta)$ to obtain the first equality. Then, we apply Lemma B.3 for the second equality. Finally, we have $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} s(\tilde{F}_t, \alpha) \xrightarrow{d} N(-cB_\alpha, \Phi_0 \Omega \Phi_0')$. Next, we derive the term with Hessian matrix. First, we rewrite the first term in (1) as follows.

$$\frac{1}{T}\sum_{t=1}^{T} H(\tilde{F}_t, \alpha) = \frac{1}{T}\sum_{t=1}^{T}\left[\varepsilon_t + \beta' H^{-1}(HF_t(\theta) - \tilde{F}_t(\theta))\right] \frac{\partial^2 g(\tilde{F}_t, \alpha)}{\partial \alpha \partial \alpha'} + \frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha} \frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha'}.$$

Under Assumption A.6 and Lemma B.1, $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \frac{\partial^2 g(\tilde{F}_t, \alpha)}{\partial \alpha \partial \alpha'} = o_p(1)$. We can also show that $-\frac{1}{T}\sum_{t=1}^{T} \beta' H^{-1}(\tilde{F}_t(\theta) - HF_t(\theta)) \frac{\partial^2 g(\tilde{F}_t, \alpha)}{\partial \alpha \partial \alpha'} = o_p(1)$. Finally, for the second term, we have

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha} \frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha'} = \Phi_0 \Sigma \Phi_0' + o_p(1) \tag{4}$$

where $\Sigma \equiv E\left[\frac{\partial g(F_t, \alpha)}{\partial \alpha} \frac{\partial g(F_t, \alpha)}{\partial \alpha'}\right]$ by replacing $\frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha}$ with $\Phi_0 \frac{\partial g(F_t, \alpha)}{\partial \alpha} + P_t$. Then, by Lemma B.2, we have $\frac{1}{T}\sum_{t=1}^{T} g_\alpha(F_t, \alpha) P_t' = o_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T} P_t P_t' = o_p(1)$. By plugging the terms, (2), (3), and (4) into (1), we have $\sqrt{T}(\tilde{\alpha} - \alpha) \xrightarrow{d} N(-c(\Phi_0 \Sigma \Phi_0')^{-1} B_\alpha, \Phi_0'^{-1} \Sigma^{-1} \Omega \Sigma^{-1} \Phi_0^{-1})$. ■

Next, we prove Lemma B.1-B.3, which we used to prove Theorem 2.1. We can obtain Lemma B.1 by applying the arguments in the proof of Lemma 1.1 in GP (2014). The proofs for (a) and (c) in Lemma B.2 are also similar to the proof of Lemma A.2 - (a) and (b) in GP (2014). Therefore, here, we show the proof for (b) and (d) in Lemma B.2. While we employ similar arguments to those in GP (2014) to prove (b) and (d), our proof relies on a new set of assumption, specifically Assumption A.4. This highlights the importance to account for serial dependence in the idiosyncratic error term in our framework, representing a novel contribution to the literature.

Proof of Lemma B.2 - (b). First, we use the identity for the factor estimation error in GP (2014) such that $\tilde{f}_{t_h} - Hf_{t_h} = \tilde{V}^{-1}(A_{1,t_h} + A_{2,t_h} + A_{3,t_h} + A_{4,t_h})$, where $A_{1,t_h} = \frac{1}{T_H}\sum_{s_h=1}^{T_H} \tilde{f}_{s_h}\gamma_{s_h t_h}$, $A_{2,t_h} = \frac{1}{T_H}\sum_{s_h}^{T_H} \tilde{f}_{s_h}\zeta_{s_h t_h}$, $A_{3,t_h} = \frac{1}{T_H}\sum_{s_h}^{T_H} \tilde{f}_{s_h}\eta_{s_h t_h}$, and $A_{4,t_h} = \frac{1}{T_H}\sum_{s_h}^{T_H} \tilde{f}_{s_h}\xi_{s_h t_h}$. Each term in A_{i,t_h} for $i = 1,2,3,4$ denotes the following: $\gamma_{s_h t_h} = E\left(\frac{1}{N}\sum_{i=1}^{N} e_{i,s_h}e_{i,t_h}\right)$, $\zeta_{s_h t_h} = \frac{1}{N}\sum_{i=1}^{N}(e_{i,s_h}e_{i,t_h} - E(e_{i,s_h}e_{i,t_h}))$, $\eta_{s_h t_h} = f_{s_h}'\frac{\Lambda' e_{t_h}}{N}$, and $\xi_{s_h t_h} = f_{t_h}'\frac{\Lambda' e_{s_h}}{N} = \eta_{t_h s_h}$. Under this identity and using the low-frequency notation, we have

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m} - Hf_{t-j/m})(\tilde{f}_{t-l/m} - Hf_{t-l/m})' \\
&\quad = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big[\tilde{V}^{-1}(A_{1,t-j/m} + A_{2,t-j/m} + A_{3,t-j/m} + A_{4,t-j/m}) \\
&\qquad \times(A_{1,t-l/m} + A_{2,t-l/m} + A_{3,t-l/m} + A_{4,t-l/m})'\tilde{V}^{-1}\Big],
\end{aligned}$$

for $j = 1,\ldots,m-1$. We analyse the convergence limit of each term, respectively. The proof is similar to the proof of Lemma A.2 - (a) in GP (2014). By applying the Cauchy-Schwarz inequality, we have $\left\|\frac{1}{T}\sum_{t=1}^{T} A_{1,t-j/m}A_{1,t-l/m}'\right\| \leq \left(\frac{1}{T}\sum_{t=1}^{T}\|A_{1,t-j/m}\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\|A_{1,t-l/m}\|^2\right)^{1/2} = O_p(1/T)$, by Assumptions A.1 and A.2. This implies $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} A_{1,t-j/m}A_{1,t-l/m}' = o_p(1)$. We can also show that $\left\|\frac{1}{T}\sum_{t=1}^{T} A_{2,t-j/m}A_{2,t-l/m}'\right\| \leq \left(\frac{1}{T}\sum_{t=1}^{T}\left\|A_{2,t-j/m}\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|A_{2,t-l/m}\right\|^2\right)^{1/2} = O_p(N^{-1}\delta_{NT_H}^{-2})$ by Cauchy-Schwarz. We also use $\frac{1}{T}\sum_{t=1}^{T}\|A_{2,t-j/m}\|^2 = O_p(N^{-1}\delta_{NT_H}^{-2})$ by Assumption A.2 and $\frac{1}{T_H}\sum_{s_h=1}^{T_H}\|\tilde{f}_s - Hf_s\|^2 = O_p(\delta_{NT_H}^{-2})$ in Bai and Ng (2006). Again, this implies $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} A_{2,t-j/m}A_{2,t-l/m}' = o_p(1)$. Similarly, we can show all the terms are negligible, except the term $\frac{1}{T}\sum_{t=1}^{T} A_{3,t-j/m}A_{3,t-l/m}'$. In fact, this term is $O_p(1/N)$, which is non-negligible when it is multiplied by $\sqrt{T}$ under our assumption, $\sqrt{T}/N \to c$. To see this, we first rewrite the term as follows.

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} A_{3,t-j/m}A_{3,t-l/m}' &= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T_H}\sum_{t=1}^{T_H}(\tilde{f}_s - Hf_s + Hf_s)\eta_{s,t-j/m}\right)\left(\frac{1}{T_H}\sum_{s=1}^{T_H}(\tilde{f}_s - Hf_s + Hf_s)\eta_{s,t-l/m}\right)' \\
&= b_{33.1} + b_{33.2} + b_{33.2}' + b_{33.3}
\end{aligned}$$

The first term $b_{33.1}$ is bounded by $\left(\frac{1}{T_H}\sum_{s=1}^{T_H}\|\tilde{f}_s - Hf_s\|^2\right)\left(\frac{1}{TT_H}\sum_{t=1}^{T}\sum_{s=1}^{T_H}|\eta_{s,t-j/m}\eta_{s,t-l/m}|\right)$ by applying Cauchy-Schwarz inequality. This is $O_p(N^{-1}\delta_{NT_H}^{-2})$ by $\frac{1}{TT_H}\sum_{t=1}^{T}\sum_{s_h=1}^{T_H}|\eta_{s_h,t-j/m}|^2 = O_p(N^{-1})$ under Assumption A.3. Similarly, the second term is bounded by Cauchy-Schwarz such that $b_{33.2} \leq \left(\frac{1}{T_H}\sum_{s=1}^{T_H}\left\|Hf_s(\tilde{f}_s - Hf_s)\right\|\right)\left(\frac{1}{TT_H}\sum_{t=1}^{T}\sum_{s=1}^{T_H}|\eta_{s,t-j/m}\eta_{s,t-l/m}|\right) = O_p(N^{-1}\delta_{NT_H}^{-1})$. Then, the final term is $b_{33.3} = H\left(\frac{f'f}{T_H}\right)\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\Lambda' e_{t-j/m}}{N}\right)\left(\frac{e'_{t-l/m}\Lambda}{N}\right)\right]\left(\frac{f'f}{T_H}\right)H' = O_p(N^{-1})$ by Assumption A.3. Thus,

$$\sqrt{T}b_{33.3} = \frac{\sqrt{T}}{N}H\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\Lambda' e_{t-j/m}}{\sqrt{N}}\right)\left(\frac{e'_{t-l/m}\Lambda}{\sqrt{N}}\right)\right]H = cH\Gamma_{j-l}H + o_p(1),$$

where we use $\frac{f'f}{T_H} = I_r$ by Assumptions A.1-(d) and A.4-(d). Finally, we have $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m} - Hf_{t-j/m})(\tilde{f}_{t-l/m} - Hf^{(m)}_{t-l/m})' = cV^{-1}H\Gamma_{j-l}HV^{-1} + o_p(1)$. ■

Proof of Lemma B.2 - (d). The proof is similar to the proof of Lemma A.2 - (b) in GP (2014). By using the identity we use in the proof of B.2-(b), we have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Hf_{t-l/m}(\tilde{f}_{t-j/m} - Hf_{t-j/m})' = H\frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{t-l/m}(A_{1,t-j/m} + A_{2,t-j/m} + A_{3,t-j/m} + A_{4,t-j/m})'\tilde{V}^{-1}$$
$$\equiv \sqrt{T}H(d_{f1} + d_{f2} + d_{f3} + d_{f4})\tilde{V}^{-1}.$$

We show the convergence limit for d_{fi}, for $i = 1, 2, 3, 4$. We can show that all the terms except d_{f4} is negligible. For example, $d_{f1} = O_p(\delta_{NT_H}^{-1}T^{-1/2}) + O_p(T_H^{-1})$. To show this, we first rewrite d_{f1} as $\frac{1}{T}\sum_{t=1}^{T}f_{t-l/m}\left(\frac{1}{T_H}\sum_{s=1}^{T_H}(\tilde{f}_s - Hf_s)'\gamma_{s,t-j/m}\right) + \frac{1}{T}\sum_{t=1}^{T}f_{t-l/m}\left(\frac{1}{T_H}\sum_{s=1}^{T_H}f'_s\gamma_{s,t-j/m}\right)H'$. The first term of d_{f1} is $O_p(\delta_{NT_H}^{-1}T^{-1/2})$ by applying Assumptions A.1-A.2 and $\frac{1}{T_H}\sum_{s_h=1}^{T_H}\|\tilde{f}_s - Hf_s\|^2 = O_p(\delta_{NT_H}^{-2})$. The second term is $O_p(T_H^{-1})$ by Cauchy-Schwarz inequality and Assumptions A.1 and A.2. We can also show that $\|d_{f2}\| = O_p((TN)^{-1/2})$ by showing $\frac{1}{T_H}\sum_{s=1}^{T_H}\left\|\frac{1}{T}\sum_{t=1}^{T}f_{t-l/m}\zeta_{s,t-j/m}\right\|^2 = O_p((TN)^{-1})$ under Assumption A.4-(b). The third term is also bounded by Cauchy-Schwarz inequality such that $\|d_{f3}\| = O_p((NT)^{-1/2})$ and by applying Assumption A.4-(c). Finally, we decompose the last term into two parts as follows.

$$d_{f4} = \frac{1}{T}\sum_{t=1}^{T} f_{t-l/m}\left(\frac{1}{T_H}\sum_{s=1}^{T_H}(\tilde{f}_s - Hf_s)'\xi_{s,t-j/m}\right) + \frac{1}{T}\sum_{t=1}^{T} f_{t-l/m}\left(\frac{1}{T_H}\sum_{s=1}^{T_H} f_s'\xi_{s,t-j/m}\right)H'$$
$$\equiv d_{f4.1} + d_{f4.2}.$$

By rearranging the second term, we have $d_{f4.2} = \frac{1}{\sqrt{T_H N}}\left(\frac{1}{T}\sum_{s=1}^{T} f_{t-l/m}f_{t-j/m}'\right)\left(\frac{1}{\sqrt{T_H N}}\sum_{s=1}^{T_H}\Lambda' e_s f_s'\right) = O_p(1/(\sqrt{T_H N}))$ by Assumptions A.4-(1) and A.3-(c). We can also rearrange the terms in the first term and write it as follows.

$$\begin{aligned} d_{f4.1} &= \frac{1}{T}\sum_{t=1}^{T} f_{t-l/m}\left[\frac{1}{T_H}\sum_{s=1}^{T_H}(\tilde{f}_s - Hf_s)'\left(f_{t-j/m}'\frac{\Lambda' e_s}{N}\right)\right] \\ &= \left(\frac{1}{T}\sum_{t=1}^{T} f_{t-l/m}f_{t-j/m}'\right)\left(\frac{1}{T_H}\sum_{s=1}^{T_H}\frac{\Lambda' e_s}{N}(\tilde{f}_s - Hf_s)'\right). \end{aligned}$$

This is $O_p(1/N)$ under our assumptions. By using $\frac{1}{T_H}\sum_{s=1}^{T_H}\frac{\Lambda' e_s}{N}(\tilde{f}_s - Hf_s)' = \frac{1}{N}(\Gamma + o_p(1))Q'V^{-1}$, from the proof in GP (2014), we have

$$\sqrt{T}Hd_{f4.1} = H\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-l/m}f_{t-j/m}'\right)\left(\frac{\sqrt{T}}{N}(\Gamma + o_p(1))Q'V^{-1}\right) = cQ_{j-l}\Gamma Q'V^{-1} + o_p(1)$$

Thus, $\sqrt{T}d_{f4.1}\tilde{V}^{-1} = cQ_{j-l}\Gamma Q'V^{-2} + o_p(1)$, where $Q_{j-l} = \frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-j/m}f_{t-l/m} = \frac{1}{T_H}\sum_{t=1}^{T_H}\tilde{f}_t f_{t-(j-l)}$.

■

Proof of Lemma B.3 - (a). We write the equation as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{F}_t(\theta)-F_t(\theta))(\tilde{F}_t(\theta)-HF_t(\theta))' \\
&=\frac{1}{\sqrt{T}}\left[\sum_{j=1}^{K}w_j(\theta)(\tilde{f}_{t-j/m}-Hf_{t-j/m})\right]\left[\sum_{j=1}^{K}w_j(\theta)(\tilde{f}_{t-j/m}-Hf_{t-j/m})\right]' \\
&=\sum_{j=1}^{K}w_j(\theta)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m}-Hf_{t-j/m})(\tilde{f}_{t-j/m}-Hf_{t-j/m})'\right]w_j(\theta) \\
&\quad+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\theta)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}_{t-j/m}-Hf_{t-j/m})(\tilde{f}_{t-l/m}-Hf_{t-l/m})'\right]w_l(\theta) \\
&=cV^{-1}Q\left\{\sum_{j=1}^{K}w_j^2(\theta)\Gamma+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\theta)\Gamma_{j-l}w_l(\theta)\right\}Q'V^{-1}+o_p(1).
\end{aligned}
$$

By applying Lemmas B.2-(a) and (b), the result follows immediately. ■

Proof of Lemma B.3 - (b). Similar to previous proof, we rewrite the equation as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}HF_t(\theta)(\tilde{F}_t(\theta)-HF_t(\theta))' \\
&=\frac{1}{\sqrt{T}}\left[\sum_{j=1}^{K}w_j(\theta)Hf_{t-j/m})\right]\left[\sum_{j=1}^{K}w_j(\theta)(\tilde{f}_{t-j/m}-Hf_{t-j/m})\right]' \\
&=\sum_{j=1}^{K}w_j(\theta)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Hf_{t-j/m}(\tilde{f}_{t-j/m}-Hf_{t-j/m})'\right]w_j(\theta) \\
&\quad+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\theta)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Hf_{t-l/m}(\tilde{f}_{t-j/m}-Hf_{t-j/m})'\right]w_l(\theta) \\
&=c\left\{\sum_{j=1}^{K}w_j^2(\theta)H+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\theta)Q_{j-l}w_l(\theta)\right\}\Gamma Q'V^{-2}+o_p(1).
\end{aligned}
$$

By applying Lemmas B.2-(c) and (d), the result follows. ■

Next, we prove Theorem 2.2 and Proposition 2.1. To prove Theorem 2.2, we first prove the case when there is no cross-sectional dependence (only serial correlation) in the idiosyncratic term in the factor model, and then we prove when the cross-sectional dependence is allowed.

Proof of Theorem 2.2.

If the idiosyncratic terms are serially correlated, but not cross-sectionally correlated, note that $\Gamma_k = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \lambda_i \lambda_i' E(e_{i,t_h} e_{i,t_h-k})$. Recall that the estimator for Γ_k under serial dependence without cross-sectional dependence is $\hat{\Gamma}_k = \frac{1}{N(T_H-k)} \sum_{t_h=k+1}^{T_H} \sum_{i=1}^{N} \tilde{\lambda}_i \tilde{\lambda}_i' \tilde{e}_{i,t_h} \tilde{e}_{i,t_h-k}$. To show that $\hat{\Gamma}_k - H_0^{-1\prime}\Gamma_k H_0^{-1} \to 0$, we can use the arguments in the proof of Theorem 6 in Bai (2003). In fact, we can use the fact that $\tilde{e}_{i,t_h} = e_{i,t_h} + O_p(\delta_{NT_H}^{-1})$ and $\tilde{\lambda}_i = H^{-1\prime}\lambda_i + O_p(\delta_{NT_H}^{-1})$, and rewrite $\hat{\Gamma}_k$ as follows.

$$\hat{\Gamma}_k = H^{-1\prime} \frac{1}{N(T_H-k)} \sum_{t_h=k+1}^{T_H} \sum_{i=1}^{N} \lambda_i \lambda_i' e_{i,t_h} e_{i,t_h-k} H^{-1} + o_p(1).$$

Since we have $\frac{1}{T_H-k}\sum_{t_h=k+1}^{T_H} e_{i,t_h} e_{i,t_h-k} \to E(e_{i,t_h} e_{i,t_h-k})$ and $H \to H_0$, we can show that $\hat{\Gamma}_k - H_0^{-1\prime}\Gamma_k H_0^{-1} \xrightarrow{p} 0$.

Next, we prove the case when the idiosyncratic terms are serially and cross-setionally correlated, we can use the arguments in the proof of Theorem 4 in Bai and Ng (2006). Under Assumption A.2 - (d), we have $\sigma_{ij,k} = E(e_{i,t_h} e_{j,t_h-k})$. Let $\tilde{\sigma}_{ij,k} = \frac{1}{T_H-k}\sum_{t_h=k+1}^{T_H} \tilde{e}_{i,t_h} \tilde{e}_{i,t_h-k}$ and $\Gamma_{n,k} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij,k}\lambda_i\lambda_j'$. By the definition, $\Gamma_k = \lim_{n\to\infty}\Gamma_{n,k}$. Let $\bar{\Gamma}_{n,k} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{\sigma}_{ij,k}\lambda_i\lambda_j'$. Then, we can write

$$\hat{\Gamma}_k - H^{-1'}\Gamma_k H^{-1} = \hat{\Gamma}_k - H^{-1'}\bar{\Gamma}_{n,k}H^{-1} + H^{-1'}(\bar{\Gamma}_{n,k} - \Gamma_{n,k})H^{-1} + H^{-1'}(\Gamma_{n,k} - \Gamma_k)H^{-1}.$$

Since Γ_k is the limit of $\Gamma_{n,k}$, we have $\Gamma_{n,k} - \Gamma_k \to 0$. The remaining parts to show are $\bar{\Gamma}_{n,k} - \Gamma_{n,k} \xrightarrow{p} 0$ if $n/N \to 0$ and $n/T_H \to 0$ and $\hat{\Gamma}_k - H^{-1'}\bar{\Gamma}_{n,k}H^{-1} \xrightarrow{p} 0$. We first rewrite

$\bar{\Gamma}_{n,k} - \Gamma_{n,k}$ as follows.

$$\begin{aligned}
\bar{\Gamma}_{n,k} - \Gamma_{n,k} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(\tilde{\sigma}_{ij,k} - \sigma_{ij,k})\lambda_i\lambda_j' \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}(e_{i,t_h}e_{j,t_h-k} - \sigma_{ij,k})\lambda_i\lambda_j' \\
&\quad - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H} e_{i,t_h}(c_{j,t_h-k} - \tilde{c}_{j,t_h-k})\lambda_i\lambda_j' \\
&\quad - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H} e_{j,t_h-k}(c_{i,t_h} - \tilde{c}_{i,t_h})\lambda_i\lambda_j' \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}(c_{i,t_h} - \tilde{c}_{i,t_h})(c_{j,t_h-k} - \tilde{c}_{j,t_h-k})\lambda_i\lambda_j' \\
&= I + II + III + IV,
\end{aligned}$$

where we obtain the second equality by using the decomposition such that $\tilde{e}_{i,t_h}\tilde{e}_{j,t_h-k} = e_{i,t_h}e_{j,t_h-k} - e_{i,t_h}(c_{j,t_h-k} - \tilde{c}_{j,t_h-k}) - e_{j,t_h-k}(c_{i,t_h} - \tilde{c}_{i,t_h}) + (c_{i,t_h} - \tilde{c}_{i,t_h})(c_{j,t_h-k} - \tilde{c}_{j,t_h-k})$, where $\tilde{c}_{i,t_h} = \tilde{\lambda}_i'\tilde{f}_{t_h}$ and $c_{i,t_h} = \lambda_i' f_{t_h}$. We can show that I is $O_p((T_H - k)^{-1/2})$ since it is zero mean process. By using $c_{j,t_h} - \tilde{c}_{j,t_h} = (H^{-1'}\lambda_j - \tilde{\lambda}_j)'\tilde{f}_{t_h} + \lambda_j' H^{-1}(Hf_{t_h} - \tilde{f}_{t_h})$ and following Bai and Ng (2006), we have $II \to 0$ if $\sqrt{n}/T_H \to 0$ and $n/\delta^2_{NT_H} \to 0$. Similarly, we have $III \to 0$ as $n/\delta^2_{NT_H} \to 0$. Finally, for IV, by Cauchy-Schwarz inequality, we have

$$\|IV\| \leq \left(\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(c_{i,t_h} - \tilde{c}_{i,t_h})\lambda_i\right\|^2\right)^{1/2}\left(\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(c_{j,t_h-k} - \tilde{c}_{j,t_h-k})\lambda_j\right\|^2\right)^{1/2}$$

Since $c_{i,t_h} - \tilde{c}_{i,t_h} = (H^{-1}\lambda_i - \tilde{\lambda}_i)'\tilde{f}_{t_h} + \lambda_i' H^{-1}(Hf_{t_h} - \tilde{f}_{t_h})$, by using c_r inequality,

$$\begin{aligned}
\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(c_{i,t_h} - \tilde{c}_{i,t_h})\lambda\right\|^2 &\leq 2\left(\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}\|f_{t_h}\|^2\right)\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda_i(H^{-1'}\lambda_i - \tilde{\lambda}_i)'\right\|^2 \\
&\quad + 2\left\|H^{-1}\right\|^2\left(\frac{1}{n}\sum_{i=1}^{n}\|\lambda_i\|^2\right)^2\frac{1}{T_H - k}\sum_{t_h=k+1}^{T_H}\left\|\tilde{f}_{t_h} - Hf_{t_h}\right\|^2.
\end{aligned}$$

The first term term and the second term converge to zero as $\sqrt{n}/T \to 0$ and $n/T_H \to 0$. The last remaining term is $\hat{\Gamma}_k - H^{-1'}\bar{\Gamma}_{n,k}H^{-1}$. We can rewrite this term as follows.

$$\begin{aligned}
\hat{\Gamma}_k - H^{-1'}\bar{\Gamma}_{n,k}H^{-1} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\sigma}_{ij,k}(\tilde{\lambda}_i\tilde{\lambda}_j' - H^{-1'}\lambda_i\lambda_j'H^{-1}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(\tilde{\sigma}_{ij,k} - \sigma_{ij,k})(\tilde{\lambda}_i\tilde{\lambda}_j' - H^{-1'}\lambda_i\lambda_j'H^{-1}) + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij,k}(\tilde{\lambda}_i\tilde{\lambda}_j' - H^{-1'}\lambda_i\lambda_j'H^{-1}) \\
&= I + II.
\end{aligned}$$

Then, $I \to 0$ using the fact that it is zero mean process. We decompose the second term II as follows.

$$II = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij,k}(\tilde{\lambda}_i - H^{-1}\lambda_i)\tilde{\lambda}_j' + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij,k}\lambda_i H^{-1}(\tilde{\lambda}_j - H^{-1'}\lambda_j)' = a + b.$$

Then, we can show that $a \to 0$ and $b \to 0$ since a and b are of order $O_p(T_H^{-1/2}) + O_p(\delta_{NT_H}^{-2})$. Since $H \xrightarrow{p} H_0$, we can complete the proof. ■

The proof of Proposition 2.1 is straightforward by applying Theorem 2.2.

C Bootstrap procedure

In Algorithm 1, we present a description of our bootstrap procedure using AR-sieve + CSD bootstrap. In step 1, we resample the residuals of the factor model by AR sieve + CSD bootstrap. This is identical to the bootstrap method we present in the main text (Algorithm 1 in the main text). In step 2, we resample the regression residuals and obtain the bootstrap sample for the MIDAS regression model. In step 3, using the two-step estimation procedure, we can obtain the bootstrap estimators.

Algorithm 1 Bootstrap for the factor-MIDAS regression model

1. **Generate bootstrap factor model.** For $t_h = 1, \ldots, T_H$, let

$$X^*_{i,t_h} = \tilde{\lambda}'_i \tilde{f}_{t_h} + e^*_{i,t_h} \quad \text{and} \quad X^*_{t_h} = \tilde{\Lambda} \tilde{f}_{t_h} + e^*_{t_h},$$

where e^*_{i,t_h} is obtained as follows. For each $i = 1, \ldots, N$, select an order $p_i = p_i(T_H)$ with $p_i \ll T_H$ (e.g., by AIC) and fit a p_i-th order autoregressive model to $\tilde{e}_{i,1}, \ldots, \tilde{e}_{i,T_H}$, where $\tilde{e}_{i,t_h} = X_{i,t_h} - \tilde{\lambda}'_i \tilde{f}_{t_h}$.

Denote $\tilde{\phi}_i(p_i) = (\tilde{\phi}_{i,j}(p_i), j = 1, \ldots, p_i)$ as the Yule-Walker autoregressive parameter estimators, such that $\tilde{\phi}_i(p_i) = \tilde{\Gamma}(p_i)^{-1}\tilde{\gamma}_{p_i}$, with $\tilde{\gamma}_{p_i} = (\tilde{\gamma}_e(1), \tilde{\gamma}_e(2), \ldots, \tilde{\gamma}_e(p_i))'$ and $\tilde{\Gamma}(p_i) = (\tilde{\gamma}_e(r-s))_{r,s=1,2,\ldots,p_i}$ such that

$$\tilde{\gamma}_e(\tau) = \frac{1}{T_H} \sum_{t_h=1}^{T_H - |\tau|} (\tilde{e}_{i,t_h} - \bar{e}_i)(\tilde{e}_{i,t_h+|\tau|} - \bar{e}_i),$$

for $\tau = 0, \ldots, p_i$ and $\bar{e}_i = T_H^{-1} \sum_{t_h=1}^{T_H} \tilde{e}_{i,t_h}$. With chosen lag length $p_i = p_i(T_H)$, generate

$$e^*_{i,t_h} = \sum_{j=1}^{p_i} \tilde{\phi}_{i,j}(p_i) e^*_{i,t_h-j} + u^*_{i,t_h}, \quad t_h = 1, \ldots, T_H,$$

where $u^*_{t_h} = (u^*_{1,t_h}, \ldots, u^*_{N,t_h})' = \tilde{\Sigma}_u^{1/2} \eta_{t_h}$ with $\eta_{t_h} \sim$ i.i.d.$(0, I_N)$. Set initial conditions $e^*_{i,0}, \ldots, e^*_{i,1-p_i} = 0$ for $i = 1, \ldots, N$.

Choose $\tilde{\Sigma}_u = (\hat{\sigma}_{u,ij})_{i,j=1,\ldots,N}$ by thresholding, with

$$\hat{\sigma}_{u,ij} = \begin{cases} \tilde{\sigma}_{u,ij} & i = j \\ \tilde{\sigma}_{u,ij} \mathbb{1}\left(|\tilde{\sigma}_{u,ij}| > \omega\right) & i \neq j, \end{cases} \quad \text{where} \quad \tilde{\sigma}_{u,ij} = \frac{1}{T_H} \sum_{t_h=1}^{T_H} \tilde{u}_{i,t_h} \tilde{u}_{j,t_h},$$

ω is a threshold, and $\tilde{u}_{i,t_h} = \tilde{e}_{i,t_h} - \sum_{j=1}^{p_i} \tilde{\phi}_{i,j}(p_i) \tilde{e}_{i,t_h-j}$ for $i = 1, \ldots, N$ and $t_h = 1 + p_i, \ldots, T_H$.

2. **Generate bootstrap factor-MIDAS regression model.** For $t = 1, \ldots, T$, construct

$$y^*_t = \tilde{\beta}_0 + \tilde{\beta}'_1 \tilde{F}_t(\tilde{\theta}) + \varepsilon^*_t,$$

where $\varepsilon^*_t = \nu_t \hat{\varepsilon}_t$, $\hat{\varepsilon}_t = y_t - \tilde{\beta}_0 - \tilde{\beta}'_1 \tilde{F}_t(\tilde{\theta})$, and $\nu_t \sim$ i.i.d.$(0, 1)$ across t, independent of η_{t_h}.

3. **Extract bootstrap factors and estimate bootstrap parameters.** Obtain the estimated factors $\tilde{f}^*_{t_h}$ by principal component analysis on the bootstrap panel $X^*_{t_h}$. After, regress y^*_t on 1 and temporally aggregated factors $(\tilde{f}^{*\prime}_{t-1/m}, \ldots, \tilde{f}^{*\prime}_{t-K/m})'$ and obtain the bootstrap estimates $\tilde{\beta}^*$ and $\tilde{\theta}^*$.

D Proof of results in Section 3

In this section, we first deliver the bootstrap high-level conditions under which our bootstrap data generating process yields a consistent bootstrap distribution. Our bootstrap data generating process (DGP) is similar to the one proposed by GP (2014). Let $\{e^*_{t_h} = (e^*_{1,t_h}, \dots, e^*_{N,t_h})'\}$ be a bootstrap sample from $\{\tilde{e}_{t_h} = (\tilde{e}_{1,t_h}, \dots, \tilde{e}_{N,t_h})'\}$, where $\tilde{e}_{t_h} = X_{t_h} - \tilde{\Lambda}\tilde{f}_{t_h}$ are the residuals from the original panel dataset. $\{\varepsilon^*_t\}$ are the resampled bootstrap residuals from $\{\tilde{\varepsilon}_t = y_t - g(\tilde{F}_t; \tilde{\alpha})\}$. Using these two bootstrap samples, $\{e^*_{t_h}\}$ and $\{\varepsilon^*_t\}$, the bootstrap data generating process (DGP) is as follows.

$$X^*_{t_h} = \tilde{\Lambda}\tilde{f}_{t_h} + e^*_{t_h}, \text{ for } t_h = 1, \dots, T_H,$$

$$y^*_t = \tilde{\beta}_0 + \tilde{\beta}'_1 \tilde{F}_t(\tilde{\theta}) + \varepsilon^*_t, \text{ for } t = 1, \dots, T.$$

We can obtain the bootstrap estimators by following a two-step process that is similar to the procedure used in the original sample: in the first step, we estimate the factors from a new bootstrap panel dataset $X^*_{t_h}$ and denote them by $\tilde{f}^*_{t_h}$, then in the second step, by regressing y^*_t on 1 and $\tilde{F}^*_t(\tilde{\theta})$, we can obtain the bootstrap estimators. We denote these estimators by $\tilde{\alpha}^*$, which are the analogues of NLS estimators from the original sample. Below conditions are our bootstrap high-level conditions. The conditions are similar to those of GP (2014).

Condition C.1* (*Weak Time Series and Cross Section Dependence in $e^*_{it_h}$*)

(a) $E^*(e^*_{i,t_h}) = 0$ *for all* (i, t_h).

(b) $\frac{1}{T_H}\sum_{t_h=1}^{T_H}\sum_{s_h=1}^{T_H} |\gamma^*_{s_h t_h}|^2 = O_p(1)$, *where* $\gamma^*_{s_h t_h} = E^*\left(\frac{1}{N}\sum_{i=1}^{N} e^*_{i,t_h} e^*_{i,s_h}\right)$.

(c) $\frac{1}{T_H^2}\sum_{t_h=1}^{T_H}\sum_{s_h=1}^{T_H} E^*\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(e^*_{i,t_h}e^*_{i,s_h} - E^*(e^*_{i,t_h}e^*_{i,s_h}))\right|^2 = O_p(1)$.

Condition C.2* (*Weak Dependence Among $\tilde{f}_{t_h}$, $\tilde{\lambda}_i$, and $\tilde{e}^*_{i,t_h}$*)

(a) $\frac{1}{T_H}\sum_{t_h=1}^{T_H}\sum_{s_h=1}^{T_H}\tilde{f}_{s_h}\tilde{f}'_{t_h}\gamma^*_{s_h t_h} = O_p(1)$.

(b) $\frac{1}{T_H}\sum_{t_h=1}^{T_H} E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{s_h=1}^{T_H}\sum_{i=1}^{N}\tilde{f}_{s_h}(e^*_{i,t_h}e^*_{i,s_h} - E^*(e^*_{i,t_h}e^*_{i,s_h}))\right\|^2 = O_p(1)$.

(c) $E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{t_h=1}^{T_H}\sum_{i=1}^{N}\tilde{f}_{t_h}\tilde{\lambda}'_i e^*_{i,t_h}\right\|^2 = O_p(1)$.

(d) $\frac{1}{T_H}\sum_{t_h=1}^{T_H} E^*\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tilde{\lambda}_i e^*_{i,t_h}\right\|^2 = O_p(1)$.

(e) $\frac{1}{T_H}\sum_{t_h=1}^{T_H}\left(\frac{\tilde{\Lambda}' e^*_{t_h}}{\sqrt{N}}\right)\left(\frac{e^{*\prime}_{t_h}\tilde{\Lambda}}{\sqrt{N}}\right) - \tilde{\Gamma} = o_{p^*}(1)$, *in probability, where* $\tilde{\Gamma} \equiv \frac{1}{T_H}\sum_{t_h=1}^{T_H} Var^*\left(\frac{1}{\sqrt{N}}\tilde{\Lambda}' e^*_{t_h}\right) > 0$.

Condition C.3* (***Serial Dependence among*** $\tilde{f}_{t_h}$, $\tilde{\lambda}_i$, ***and*** $\tilde{e}^*_{i,t_h}$)

(a) $\frac{1}{T_H}\sum_{t_h=1}^{T_H} E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{s_h=1}^{T_H}\sum_{i=1}^{N}\tilde{f}_{s_h}(e^*_{i,t_h}e^*_{i,s_h-k} - E^*(e^*_{i,t_h}e^*_{i,s_h-k}))\right\|^2 = O_p(1)$ *for all* k.

(b) $E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{t_h=1}^{T_H}\tilde{f}_{t_h}e^{*\prime}_{t_h-k}\tilde{\Lambda}\right\|^2 = O_p(1)$ *for all* k.

(c) $\frac{1}{T_H}\sum_{t_h=1}^{T_H}\left(\frac{\tilde{\Lambda} e^*_{t_h}}{\sqrt{N}}\right)\left(\frac{e^{*\prime}_{t_h-k}\tilde{\Lambda}}{\sqrt{N}}\right) - \tilde{\Gamma}_k = o_{p^*}(1)$, *in probability, where* $\tilde{\Gamma}_k \equiv \frac{1}{T_H}\sum_{t_h=1}^{T_H} Cov^*\left(\frac{\tilde{\Lambda}' e^*_{t_h}}{\sqrt{N}}, \frac{\tilde{\Lambda}' e^*_{t_h-k}}{\sqrt{N}}\right) > 0$.

Condition C.4* (***Weak Dependence Between*** e^*_{i,t_h} ***and*** ε^*_t)

(a) $\frac{1}{T}\sum_{t=1}^{T} E^*\left|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\varepsilon^*_s(e^*_{i,t-j/m}e^*_{i,s-j/m} - E^*(e^*_{i,t-j/m}e^*_{i,s-j/m}))\right|^2 = O_p(1)$ *for* $j = 0, \ldots, m-1$.

(b) $E^*\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}\sum_{i=1}^{N}\tilde{\lambda}_i e^*_{i,t-j/m}\varepsilon^*_t\right\|^2 = O_p(1)$, *where* $E(e^*_{i,t-j/m}) = 0$ *for all* (i,t) *and* $j = 0, \ldots, m-1$.

Condition C.5* (***Bootstrap CLT***)

(a) $E^*(\varepsilon^*_t) = 0$ *and* $\frac{1}{T}\sum_{t=1}^{T} E^*|\varepsilon_t|^2 = O_p(1)$.

(b) $\tilde{\Omega}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{g}_{\alpha,t}\varepsilon_t^* \xrightarrow{d*} N(0, I_{r+p})$*, in probability, where* $E^*\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{g}_{\alpha,t}\varepsilon_t^*\right\|^2 = O_p(1)$ *and* $\tilde{g}_{\alpha,t} = \partial g(\tilde{F}_t, \alpha)/\partial\alpha$*, and* $\tilde{\Omega} \equiv Var^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{g}_{\alpha,t}\varepsilon_t^*\right) > 0$*.*

Condition C.6* (*Bootstrap Consistency*)

(a) $\operatorname{plim}\tilde{\Omega} = \Phi_0\Omega\Phi_0'$*, where* $\tilde{\Omega} = Var^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{g}_{\alpha,t}\varepsilon_t^*\right)$ *and* $\tilde{g}_{\alpha,t} \equiv \partial g(\tilde{F}_t, \alpha)/\partial\alpha$*.*

(b) $\operatorname{plim}\tilde{\Gamma} = H_0\Gamma H_0'$ *and* $\operatorname{plim}\tilde{\Gamma}_{j-l} = H_0\Gamma_{j-l}H_0'$*.*

Conditions C.1* through C.4* are the bootstrap analogues of Assumptions A.1 to A.6 in Appendix A. Conditions C.1*-C.2* are similar to the bootstrap high level conditions in GP (2014). The mean of bootstrap residuals are required to be zeros for all (i, t_h), which implies that we need to recenter the residuals when we resample them. Condition C.3* is a new set of high-level conditions required in our context. Unlike in GP (2014), since our bias contains the term which relies on serial dependence in the idiosyncratic error term in the factor model, we impose weak serial dependence among $\tilde{f}_{t_h}$, $\tilde{\lambda}_i$ and e_{i,t_h}^* in Condition C.3*. Note that since $\tilde{f}_{t_h}$ and $\tilde{\lambda}_i$ are fixed in the bootstrap world, serial dependence in the factors can be implied by restricting the serial dependence of e_{i,t_h}. Condition C.4* is similar to Condition C* in GP (2014), and we restrict the dependence between two bootstrap residuals. Condition C.5* implies that we can apply a central limit theorem on the score vector, $\tilde{g}_{\alpha,t}\varepsilon_t^*$. In Condition C.6*, we provide conditions for consistency of the bootstrap distribution. In Condition C.6*-(a), Ω denotes the bootstrap variance of the score vector in the bootstrap world and it is a bootstrap analogue of Ω. It implies that the bootstrap variance is rotated with a block diagonal matrix, Φ_0. This is because the score vector $\tilde{g}_{\alpha,t} = \left(F_t'(\theta)H', \beta'\frac{\partial F_t(\theta)}{\partial\theta'}\right)'$ is a rotated version of $g_{\alpha,t}$, where the rotation is given by Φ_0. Similarly, $\tilde{\Gamma}$ and $\tilde{\Gamma}_{j-l}$, defined in Condition C.2* and Condition C.3* are the bootstrap analogues of Γ and Γ_{j-l}, respectively.

Condition C.6*-(a) and (b) imply that it is crucial how we mimic the error terms of the MIDAS regression and the idiosyncratic factor error terms in the bootstrap world. Moreover, in our context, since the bias depends on both serial and cross-sectional dependence of e_{t_h}, the idiosyncratic error term in the bootstrap world should mimic the dependence in the time series and cross-sectional dimension.

Remark 1 *Note that $\tilde{\alpha}^*$ is obtained by regressing y_t^* on 1 and a temporally aggregated version of the lags of the bootstrap estimated factors, $\tilde{F}_t^*(\tilde{\theta})$. The bootstrap estimated factors, $\tilde{f}_{t_h}^*$, consistently estimate the rotated version of true "latent" bootstrap factors, $H^*\tilde{f}_{t_h}$, where $H^* = \tilde{V}^{*-1}\frac{\tilde{f}^{*\prime}\tilde{f}}{T_H}\frac{\tilde{\Lambda}'\tilde{\Lambda}}{N}$ and $\tilde{V}^*$ is the $r \times r$ diagonal matrix containing on the main diagonal the r largest eigenvalues of $X^*X^{*\prime}/NT_H$, in decreasing order. This matrix is the bootstrap analogue of the rotation matrix in the original sample, $H = \tilde{V}^{-1}\frac{\tilde{f}'f}{T_H}\frac{\Lambda'\Lambda}{N}$. As discussed in GP (2014), the indeterminacy of the rotation matrix is not a problem in the bootstrap world, as H^* does not depend on the population values. Moreover, H^* is asymptotically equal to $H_0^* = \mathrm{diag}(\pm 1)$, where the sign is determined by the sign of $\tilde{f}^{*\prime}\tilde{f}/T_H$. This implies that the bootstrap factors are identified up to a change of sign.*

Remark 2 *Similar to the discussion in GP (2014) regarding the rotation of the bootstrap estimators, our NLS estimators of bootstrap DGP rotate due to the rotation in the factors in the bootstrap world. Note that we can rewrite y_t^* as follows.*

$$y_t^* = \tilde{\beta}_0 + \tilde{\beta}_1' H^{*-1}\tilde{F}_t^*(\tilde{\theta}) + \tilde{\beta}_1' H^{*-1}(H^*\tilde{F}_t(\tilde{\theta}) - \tilde{F}_t^*(\tilde{\theta})) + \varepsilon_t^* = g(\tilde{F}_t^*, \tilde{\alpha}) + \xi_t^*,$$

where $g(\tilde{F}_t^*, \tilde{\alpha}) \equiv \tilde{\beta}_0 + \tilde{\beta}_1' H^{*-1}\tilde{F}_t^*(\tilde{\theta})$ and $\xi_t \equiv \tilde{\beta}_1' H^{*-1}(H^*\tilde{F}_t(\tilde{\theta}) - \tilde{F}_t^*(\tilde{\theta})) + \varepsilon_t^*$. Thus, $\tilde{\alpha}^*$ estimates $(\Phi^*)^{-1}\tilde{\alpha}$, where $\Phi^* = diag(1, H^*, I_p)$ is a block diagonal matrix. $(\Phi^*)^{-1}\tilde{\alpha}$ are the rotated version of NLS estimators in the original sample. As H^* is asymptotically equal *to* H_0^**,* $(\Phi^*)^{-1}\tilde{\alpha}$ *is equal to* $(\Phi_0^*)^{-1}\tilde{\alpha}$*, where* $\Phi_0^* = \text{diag}(1, H_0^*, I_p)$*, and* $(\Phi_0^*)^{-1}\tilde{\alpha}$ *is the sign-adjusted version of* $\tilde{\alpha}$*.*

Lemma D.1 *Let the Assumptions A.1-A.5 in Appendix A hold and consider any residual-based bootstrap scheme for which Conditions C.1*-C.5* are verified. Suppose* $\sqrt{T}/N \to c$*,* $0 \leq c < \infty$*. In addition, let the two following conditions hold: (1) Condition C.6*-(a) is verified and (2)* $c = 0$ *or Condition C.6*-(b) is verified; then as* $N, T \to \infty$*,*

$$\sqrt{T}(\tilde{\alpha}^* - (\Phi_0^*)^{-1}\tilde{\alpha}) \xrightarrow{d^*} N(-c(\Phi_0^*)^{-1}\Delta_\alpha, (\Phi_0^*)^{-1}\Sigma_\alpha(\Phi_0^*)^{-1}),$$

in probability and Δ_α *and* Σ_α *are defined in Theorem 2.1.*

Remark 3 *In Lemma D.1, we derive the bootstrap distribution of the estimators,* $\tilde{\alpha}^*$*. According to Lemma D.1, the distribution of* $\sqrt{T}(\tilde{\alpha}^* - (\Phi_0^*)^{-1}\tilde{\alpha})$ *follows a normal distribution with a non-zero mean vector,* $-c(\Phi_0^*)^{-1}\Delta_\alpha$*. The asymptotic bias is proportional to* $(H_0^*)^{-1}\tilde{\beta}$*. However, the weighting parameters* $\tilde{\theta}^*$ *are not affected by the rotation problem.*

Remark 4 *To match the bootstrap distribution with the limiting distribution of the estimators in the original sample to achieve bootstrap consistency since our rotation matrix* H^* *may not be an identity matrix. Therefore, we consider the rotated version of our bootstrap results, given by* $\sqrt{T}(\Phi^*\tilde{\alpha}^* - \tilde{\alpha})$*. For the consistency of the rotated bootstrap results, we rely on the Corollary 3.1. in GP (2014) such that* $\sup_{x\in\mathbb{R}^{r+p}} |P^*(\sqrt{T}(\Phi_0^*\tilde{\alpha}^* - \tilde{\alpha}) \leq x) - P(\sqrt{T}(\tilde{\alpha} - \alpha) \leq x)| \xrightarrow{p} 0$*. For detail, see GP (2014). This corollary justifies the use of a residual-based bootstrap method in the context of the factor-MIDAS regression models.*

Notation: P^* denotes the bootstrap probability measure, conditional on the original sample. The bootstrap measure P^* depends on the original sample size N, T and T_H, and sample realization ω, but for a simpler notation, we omit these and write P^* for $P^*_{NT,\omega}$. We write $T^*_{NT} = o_{p^*}(1)$, in probability, or $T^*_{NT} \xrightarrow{p*} 0$, in probability, for any bootstrap test statistics T^*_{NT}, if, when for any $\delta > 0$, $P^*(|T^*_{NT}| > \delta) = o_p(1)$. If for all $\delta > 0$, there exists $M_\delta < \infty$ such that $\lim_{N,T\to\infty} P[P^*(|T^*_{NT}| > M_\delta) > \delta] = 0$, we write as $T^*_{NT} = O_{p^*}(1)$, in probability. We write $T^*_{NT} \xrightarrow{d*} D$, in probability, if T^*_{NT} weakly converges to the distribution D under P^*, conditional on a sample with probability that converges to one, i.e. $E^*(f(T^*_{NT})) \xrightarrow{p} E(f(D))$ for all bounded and uniformly continuous function f.

Lemma D.2 $\frac{1}{T}\sum_{t=1}^{T} \varepsilon^*_t(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m}) = o_{p^*}(1)$.

Lemma D.3 *If* $\sqrt{T}/N \to c$, *where* $0 \le c < \infty$,

(a) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})' = \frac{\sqrt{T}}{N}\tilde{V}^{*-1}H^*\Gamma^*H^*\tilde{V}^{*-1} + o_{p^*}(1)$,

(b) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})' = \frac{\sqrt{T}}{N}\tilde{V}^{*-1}H^*\Gamma^*_{j-l}H^*\tilde{V}^{*-1} + o_{p^*}(1)$,

(c) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^*\tilde{f}_{t-j/m}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})' = \frac{\sqrt{T}}{N}H^*\Gamma^*\left(\frac{1}{T_H}\sum_{s_h=1}^{T_H}\tilde{f}_{s_h}\tilde{f}^{*\prime}_{s_h}\right)\tilde{V}^{*-2} + o_{p^*}(1)$,

(d) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^*\tilde{f}_{t-l/m}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})' = \frac{\sqrt{T}}{N}H^*\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{f}'_{t-j/m}\right)\Gamma^*\left(\frac{1}{T_H}\sum_{s_h=1}^{T_H}\tilde{f}_{s_h}\tilde{f}^{*\prime}_{s_h}\right)\tilde{V}^{*-2} + o_{p^*}(1)$.

Lemma D.4 *If* $\sqrt{T}/N \to c$, *where* $0 \le c < \infty$,

(a) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{F}^*_t(\tilde{\theta}) - H^*\tilde{F}_t(\tilde{\theta}))(\tilde{F}^*_t(\tilde{\theta}) - H^*\tilde{F}_t(\tilde{\theta}))'$

$= cH^*_0\tilde{V}^{-1}\left(\sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^* w_j(\tilde{\theta}) + \sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^*_{j-l}w_l(\tilde{\theta})\right)\tilde{V}^{-1}H^*_0 + o_{p^*}(1)$,

(b) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^*\tilde{F}_t(\tilde{\theta})(\tilde{F}^*_t(\tilde{\theta}) - H^*\tilde{F}_t(\tilde{\theta}))'$

$= cH^*_0\left[\sum_{j=1}^{K} w_j^2(\tilde{\theta}) + \sum_{j=1}^{K}\sum_{l\neq j}^{K} w_j(\tilde{\theta})\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{f}'_{t-j/m}\right)w_l(\tilde{\theta})\right]\Gamma^*\tilde{V}^{-2}H^*_0 + o_{p^*}(1)$.

Proof of Lemma D.1. Since in the bootstrap world, $\tilde{\alpha}^*$ maximizes the following objective function:

$$\tilde{Q}_T^*(\tilde{\alpha}) = -\frac{1}{T}\sum_{t=1}^{T}[y_t - g(\tilde{F}_t^*, \tilde{\alpha})]^2.$$

where $g(\tilde{F}_t^*, \tilde{\alpha}) = \tilde{\beta}' H^{*-1}\tilde{F}_t^*(\tilde{\theta})$. Then, we have

$$\sqrt{T}(\tilde{\alpha}^* - (\Phi^*)^{-1}\tilde{\alpha}) = -\left[\frac{1}{T}\sum_{t=1}^{T} H(\tilde{F}_t^*, \tilde{\alpha}_T)\right]^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} s(\tilde{F}_t^*, \tilde{\alpha}),$$

where $s(\tilde{F}_t^*, \tilde{\alpha})$ is a score vector and $H(\tilde{F}_t^*, \tilde{\alpha})$ is a Hessian matrix in the bootstrap world. $\tilde{\alpha}_T$ is intermediate between $\tilde{\alpha}$ and $\tilde{\alpha}^*$. We analyse each term. We can write the score vector as follows.

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} s(\tilde{F}_t^*, \tilde{\alpha}) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\varepsilon_t^* + \tilde{\beta}' H^{*-1}(H^*\tilde{F}_t(\tilde{\theta}) - \tilde{F}_t^*(\tilde{\theta}))]\frac{\partial g(\tilde{F}_t^*, \tilde{\alpha})}{\partial \tilde{\alpha}},$$

where the partial derivative is

$$\frac{\partial g(\tilde{F}_t^*, \tilde{\alpha})}{\partial \tilde{\alpha}} = \Phi^*\frac{\partial g(\tilde{F}_t, \alpha)}{\partial \alpha} + P_t^*, \text{ where } P_t^* = \begin{bmatrix} 0 \\ \tilde{F}_t^*(\tilde{\theta}) - H^*\tilde{F}_t(\theta) \\ \left(\frac{\partial \tilde{F}_t^*(\tilde{\theta})'}{\partial \tilde{\theta}} H^{*-1'}\tilde{\beta} - \frac{\partial \tilde{F}_t(\theta)'}{\partial \theta} H^{-1'}\beta\right) \end{bmatrix},$$

and $\Phi^* = diag(1, H^*, I_p)$. Under this decomposition, we can analyse $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\frac{\partial g(\tilde{F}_t^*, \tilde{\alpha})}{\partial \tilde{\alpha}}$ into two non-zero blocks of P_t^*. The second block can be written as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*(\tilde{F}_t^*(\tilde{\theta})-H^*\tilde{F}_t(\theta)) \\
&\quad=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\left[\sum_{j=1}^{K}w_j(\tilde{\theta})(\tilde{f}_{t-j/m}^*-H^*\tilde{f}_{t-j/m})+\sum_{j=1}^{K}(w_j(\tilde{\theta})-w_j(\theta))H^*\tilde{f}_{t-j/m}\right] \\
&\quad=\sum_{j=1}^{K}w_j(\tilde{\theta})\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*(\tilde{f}_{t-j/m}^*-H^*\tilde{f}_{t-j/m})+\sum_{j=1}^{K}(w_j(\tilde{\theta})-w_j(\theta))H^*\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\tilde{f}_{t-j/m} \\
&\quad=o_{p^*}(1).
\end{aligned}
$$

Since $\tilde{\theta}\xrightarrow{p}\theta$ and weighting function is continuous function, we can use continuous mapping theorem and have the second part as $o_p(1)$. By Lemma D.2 we can show that the first part is $o_{p^*}(1)$. The third part can be argued similarly. Since it is easier to check for each row, we write k-th row of the third block in P_t^* as $\left(\frac{\partial\tilde{F}_{k,t}^*(\tilde{\theta}_k)}{\partial\tilde{\theta}_k}H_k^{*-1'}\tilde{\beta}_k-\frac{\partial\tilde{F}_{k,t}(\theta_k)}{\partial\theta_k}H_k^{-1'}\beta_k\right)$. Then, for this k-th row, we can write it as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\left(\frac{\partial\tilde{F}_{k,t}^*(\tilde{\theta}_k)}{\partial\tilde{\theta}_k}H_k^{*-1'}\tilde{\beta}_k-\frac{\partial\tilde{F}_{k,t}(\theta_k)}{\partial\theta_k}H_k^{-1'}\beta_k\right) \\
&=H_k^{*-1'}\tilde{\beta}_k\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\sum_{j=1}^{K}\frac{\partial w_{j,k}(\tilde{\theta}_k)}{\partial\tilde{\theta}_k}(\tilde{f}_{k,t-j/m}^*-H_k^*\tilde{f}_{k,t-j/m})\right. \\
&\qquad\left.+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\sum_{j=1}^{K}\left\{\frac{\partial w_{j,k}(\tilde{\theta}_k)}{\partial\tilde{\theta}_k}-\frac{\partial w_{j,k}(\theta_k)}{\partial\theta_k}\right\}\tilde{f}_{k,t-j/m}\right] \\
&\qquad+(\tilde{\beta}_k-H_k^{-1'}\beta_k)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\left[\sum_{j=1}^{K}\frac{\partial w_{j,k}(\theta_k)}{\partial\theta_k}\tilde{f}_{k,t-j/m}\right] \\
&=o_{p^*}(1),
\end{aligned}
$$

where H_k is the k-th diagonal element in the rotation matrix H and β_k is the k-th slope parameter in β. We can obtain the second equality because $\tilde{\beta}\xrightarrow{p}H^{-1'}\beta$ and Lemma D.2.

Finally, we have the following result.

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\frac{\partial g(\tilde{F}_t;\tilde{\alpha})}{\partial\tilde{\alpha}}\xrightarrow{d^*}N(0,\Phi_0^*\tilde{\Omega}\Phi_0^*), \tag{5}$$

where $\Phi_0^* = \operatorname{plim}\Phi^*$, $\tilde{\Omega}\equiv \operatorname{Var}^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t^*\tilde{g}_{\alpha,t}\right)$, and $\tilde{g}_{\alpha,t}=\partial g(\tilde{F}_t,\alpha)/\partial\alpha$.

Now, we analyse the second term in the score vector $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{\beta}'H^{*-1}(H^*\tilde{F}_t(\tilde{\theta})-\tilde{F}_t^*(\tilde{\theta}))\frac{\partial g(\tilde{F}_t^*,\alpha)}{\partial\tilde{\alpha}}$ with respect to $\tilde{\beta}$ and $\tilde{\theta}$, respectively. (Note that there is no bias with respect to $\tilde{\beta}_0$, therefore we focus on $\tilde{\beta}_1$ here.) By Lemma D.4, the score vector with respect to $\tilde{\beta}_1$ can be written as follows.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(H^*\tilde{F}_t(\tilde{\theta})-\tilde{F}_t^*(\tilde{\theta}))\tilde{F}_t^*(\tilde{\theta})'H^{*-1'}\tilde{\beta}_1\\
&=-\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{F}_t^*(\tilde{\theta})-H^*\tilde{F}_t(\tilde{\theta}))(\tilde{F}_t^*(\tilde{\theta})-H^*\tilde{F}_t(\tilde{\theta}))'+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}H^*\tilde{F}_t(\tilde{\theta})(\tilde{F}_t^*(\tilde{\theta})-H^*\tilde{F}_t(\tilde{\theta}))'\right]H^{*-1'}\tilde{\beta}_1\\
&=-cH_0^*\left[\tilde{V}^{-1}\left\{\sum_{j=1}^{K}w_j(\tilde{\theta})\tilde{\Gamma}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\tilde{\theta})\tilde{\Gamma}_{j-l}w_l(\tilde{\theta})\right\}\tilde{V}^{-1}\right.\\
&\qquad\left.+\left\{\sum_{j=1}^{K}w_j^2(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\tilde{\theta})\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_{t-j/m}\tilde{F}_{t-l/m}'\right)w_l(\tilde{\theta})\right\}\tilde{\Gamma}\tilde{V}^{-2}\right]\tilde{\beta}_1\\
&=-cH_0^*\tilde{B}_{\beta_1}+o_{p^*}(1)
\end{aligned}$$

in probability, where we define $\tilde{B}_{\beta_1}$ as follows.

$$\begin{aligned}
\tilde{B}_{\beta_1}\equiv&\left[\tilde{V}^{-1}\left\{\sum_{j=1}^{K}w_j(\tilde{\theta})\tilde{\Gamma}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\tilde{\theta})\tilde{\Gamma}_{j-l}w_l(\tilde{\theta})\right\}\tilde{V}^{-1}\right.\\
&\left.+\left\{\sum_{j=1}^{K}w_j^2(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}w_j(\tilde{\theta})\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_{t-j/m}\tilde{F}_{t-l/m}'\right)w_l(\tilde{\theta})\right\}\tilde{\Gamma}\tilde{V}^{-2}\right]\tilde{\beta}_1.
\end{aligned}$$

We can also rewrite the part with respect to $\tilde{\theta}$ by Lemma D.4 as follows.

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\partial \tilde{F}_t^*(\tilde{\theta})}{\partial\tilde{\theta}}H^{*-1'}\tilde{\beta}_1\tilde{\beta}_1'H^{*-1}[H^*\tilde{F}_t(\tilde{\theta})-\tilde{F}_t^*(\tilde{\theta})]\\
&\quad=-c\tilde{\beta}_1\circ\left[\tilde{V}^{-1}\left\{\sum_{j=1}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\tilde{\Gamma}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\tilde{\Gamma}_{j-l}w_l(\tilde{\theta})\right\}\tilde{V}^{-1}\right.\\
&\qquad\left.+\left\{\sum_{j=1}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_{t-j/m}\tilde{F}_{t-l/m}'\right)w_l(\tilde{\theta})\right\}\tilde{\Gamma}\tilde{V}^{-2}\right]\tilde{\beta}_1\\
&\quad=-c\tilde{B}_\theta+o_{p^*}(1),
\end{aligned}
$$

in probability, where we define $\tilde{B}_\theta$ as follows.

$$
\begin{aligned}
\tilde{B}_\theta\equiv\tilde{\beta}_1\circ&\left[\tilde{V}^{-1}\left\{\sum_{j=1}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\tilde{\Gamma}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\tilde{\Gamma}_{j-l}w_l(\tilde{\theta})\right\}\tilde{V}^{-1}\right.\\
&\left.+\left\{\sum_{j=1}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}w_j(\tilde{\theta})+\sum_{j=1}^{K}\sum_{l\neq j}^{K}\frac{\partial w_j(\tilde{\theta})}{\partial\tilde{\theta}}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{F}_{t-j/m}\tilde{F}_{t-l/m}'\right)w_l(\tilde{\theta})\right\}\tilde{\Gamma}\tilde{V}^{-2}\right]\tilde{\beta}_1.
\end{aligned}
$$

Next, we derive the hessian matrix. We first rewrite it as follows.

$$
\frac{1}{T}\sum_{t=1}^{T}H(\tilde{F}_t^*,\tilde{\alpha})=\frac{1}{T}\sum_{t=1}^{T}\xi_t\frac{\partial^2 g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}\partial\tilde{\alpha}'}+\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}}\frac{\partial g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}'}=H_1+H_2.
$$

Then, H_1 is $o_{p^*}(1)$ by Condition C.5*-(b) and the results in the proof for Lemma D.3. The second term H_2 converges in probability to $\Phi_0^*\tilde{\Sigma}\Phi_0^*$ as following:

$$
\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}}\frac{\partial g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}'}\xrightarrow{p^*}\Phi_0^*E\Big[\frac{\partial g(\tilde{F}_t,\alpha)}{\partial\alpha}\frac{\partial g(\tilde{F}_t,\alpha)}{\partial\alpha'}\Big]\Phi_0^*\equiv\Phi_0^*\tilde{\Sigma}\Phi_0^*, \tag{6}
$$

where $E\Big[\frac{\partial g(\tilde{F}_t,\alpha)}{\partial\alpha}\frac{\partial g(\tilde{F}_t,\alpha)}{\partial\alpha'}\Big]\equiv\tilde{\Sigma}$. We can obtain this by rewriting $\frac{\partial g(\tilde{F}_t^*,\tilde{\alpha})}{\partial\tilde{\alpha}}=\Phi^*\frac{\partial g(\tilde{F}_t,\alpha)}{\partial\alpha}+P_t^*$. Then, $\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(\tilde{F}_t,\tilde{\alpha})}{\partial\tilde{\alpha}}P_t^{*'}=o_{p^*}(1)$ and $\frac{1}{T}\sum_{t=1}^{T}P_t^*P_t^{*'}=o_{p^*}(1)$, in probability. By putting all together, we have

$$
\sqrt{T}(\tilde{\alpha}^*-(\Phi_0^*)^{-1}\tilde{\alpha})\xrightarrow{d^*}N(-c(\Phi_0^*\tilde{\Sigma}\Phi_0^*)^{-1}\Phi_0^*\tilde{B}_\alpha,\Phi_0^{*-1}\tilde{\Sigma}^{-1}\tilde{\Omega}\tilde{\Sigma}^{-1}\Phi_0^{*-1}), \tag{7}
$$

in probability, where $\tilde{B}_\alpha = (0, \tilde{B}_{\beta_1}, \tilde{B}_\theta)'$. Under Assumptions A.1-A.6, we have $\text{plim}\, \tilde{V} = V$, $\text{plim}\, \tilde{\alpha} = \Phi^{-1}\alpha$, $\text{plim}\, \Phi^* = \Phi_0^*$, and $\text{plim}\, \tilde{\Omega} = \Phi_0 \Omega \Phi_0$. This implies that $\sqrt{T}(\tilde{\alpha}^* - (\Phi_0^*)^{-1}\tilde{\alpha}) \xrightarrow{d^*} N(-c\Phi_0^{*-1}\Delta_\alpha, \Phi_0^{*-1}\Sigma_\alpha \Phi_0^{*-1})$, in probability. ■

The proof of Lemma D.3 is similar to the proof of Lemma B.2 in GP (2014) and Lemma D.3 - (a) and (c) are similar to the proof of Lemma B.3 - (a) and (b) in GP (2014), respectively. Thus, we focus here to prove Lemma D.3-(b) and (d), which are new.

Proof of Lemma D.3. Part(b): Using the identity in GP (2014), we can rewrite the part (b) as follows.

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})' \\
&= \tilde{V}^{*-1}\frac{1}{T}\sum_{t=1}^{T}(A^*_{1,t-j/m} + A^*_{2,t-j/m} + A^*_{3,t-j/m} + A^*_{4,t-j/m}) \\
&\quad \times (A^*_{1,t-l/m} + A^*_{2,t-l/m} + A^*_{3,t-l/m} + A^*_{4,t-l/m})'\tilde{V}^{*-1}.
\end{aligned}$$

Ignoring $\tilde{V}^{*-1} = O_{p^*}(1)$, we can show that the terms except $\frac{1}{T}\sum_{t=1}^{T} A^*_{3,t-j/m}A^{*\prime}_{3,t-l/m}$ are negligible. For example, we have $\frac{1}{T}\sum_{t=1}^{T} A^*_{1,t-j/m}A^{*\prime}_{1,t-l/m} = O_{p^*}(T^{-1})$, $\frac{1}{T}\sum_{t=1}^{T} A^*_{2,t-j/m}A^{*\prime}_{2,t-l/m} = O_{p^*}(N^{-1}\delta_{NT_H}^{-2})$, and $\frac{1}{T}\sum_{t=1}^{T} A^*_{4,t-j/m}A^{*\prime}_{4,t-l/m} = O_{p^*}(N^{-1}\delta_{NT_H}^{-2})$. The cross terms are: $\frac{1}{T}\sum_{t=1}^{T} A^*_{1,t-j/m}A^{*\prime}_{2,t-l/m} =$ $O_{p^*}(T^{-1/2}N^{-1/2}\delta_{NT_H}^{-1})$, $\frac{1}{T}\sum_{t=1}^{T} A^*_{1,t-j/m}A^{*\prime}_{3,t-l/m} = O_{p^*}(T^{-1/2}N^{-1/2})$, $\frac{1}{T}\sum_{t=1}^{T} A^*_{1,t-j/m}A^{*\prime}_{4,t-l/m} = O_{p^*}(T^{-1/2}N^{-1/2})$, $\frac{1}{T}\sum_{t=1}^{T} A^*_{2,t-j/m}A^{*\prime}_{3,t-l/m} = O_{p^*}(N^{-1}\delta_{NT_H}^{-2})$, $A^*_{2,t-j/m}A^{*\prime}_{4,t-l/m} = O_{p^*}(N^{-1}\delta_{NT_H}^{-1})$, and $\frac{1}{T}\sum_{t=1}^{T} A^*_{3,t-j/m}A^{*\prime}_{4,t-l/m} = O_{p^*}(N^{-1}\delta_{NT_H}^{-1})$. Since we can show that

$$\frac{1}{T}\sum_{t=1}^{T} A^*_{3,t-j/m}A^{*\prime}_{3,t-l/m} = \frac{1}{N}H^*\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\tilde{\Lambda}' e^*_{t-j/m}}{\sqrt{N}}\right)\left(\frac{e^{*\prime}_{t-l/m}\tilde{\Lambda}}{\sqrt{N}}\right)H^* + o_{p^*}(1),$$

we have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})' = \frac{\sqrt{T}}{N}\tilde{V}^{*-1}H^*\tilde{\Gamma}_{j-l}H^*\tilde{V}^{*-1} + o_{p^*}(1),$$

where we define $\Gamma^*_{j-l} \equiv \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\tilde{\Lambda}' e^*_{t-j/m}}{\sqrt{N}}\right)\left(\frac{e^{*\prime}_{t-l/m}\tilde{\Lambda}}{\sqrt{N}}\right)$. Part (d): Similar to the identity we used in part (b), we can rewrite part (d) as follows.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^* \tilde{f}_{t-j/m}(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})' \\
&\quad = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^* \tilde{f}_{t-j/m}(A^*_{1,t-l/m} + A^*_{2,t-l/m} + A^*_{3,t-l/m} + A^*_{4,t-l/m})'\tilde{V}^{*-1} \\
&\quad \equiv \sqrt{T} H^*(d^*_{f1} + d^*_{f2} + d^*_{f3} + d^*_{f4})'\tilde{V}^{*-1},
\end{aligned}$$

where $d^*_{fi} \equiv \frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-j/m}A^{*\prime}_{i,t-l/m}$ for $i = 1,2,3,4$. Then, we can obtain $d^*_{f1} = O_{p^*}(\delta^{-1}_{NT_H}T^{-1/2}) + O_{p^*}(T_H^{-1})$, $d^*_{f2} = O_{p^*}((TN)^{-1/2})$ by Condition C.3*-(a) and $d^*_{f3} = O_{p^*}((TN)^{-1/2})$ by Condition C.3*-(b). Finally, $d^*_{f4} = \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{F}'_{t-j/m}\right)\Gamma^*\left(\frac{1}{T_H}\sum_{t=1}^{T_H}\tilde{f}_t\tilde{f}^{*\prime}_t\right)\tilde{V}^{*-1} + o_{p^*}(1)$. Thus,

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^* \tilde{f}_{t-j/m}(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})' \\
&\quad = \frac{\sqrt{T}}{N} H^* \left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{f}'_{t-j/m}\right)\Gamma^*\left(\frac{1}{T_H}\sum_{s=1}^{T_H}\tilde{f}_s\tilde{f}^{*\prime}_s\right)\tilde{V}^{*-2} + o_{p^*}(1).
\end{aligned}$$

■

Proof of Lemma D.4. Part (a): We rewrite part (a) and apply Lemma D.3.

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\sum_{j=1}^{K} w_j(\tilde{\theta})(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})\right]\left[\sum_{j=1}^{K} w_j(\tilde{\theta})(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})\right]' \\
&= \sum_{j=1}^{K} w_j(\tilde{\theta})\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})'\right] w_j(\tilde{\theta}) \\
&\quad + \sum_{j=1}^{K}\sum_{l\neq j}^{K} w_j(\tilde{\theta})\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})(\tilde{f}^*_{t-l/m} - H^*\tilde{f}_{t-l/m})'\right] w_l(\tilde{\theta}) \\
&= c\tilde{V}^{*-1}H^*\left(\sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^* w_j(\tilde{\theta}) + \sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^*_{j-l} w_l(\tilde{\theta})\right) H^*\tilde{V}^{*-1} + o_{p^*}(1) \\
&= cH_0^*\tilde{V}^{-1}\left(\sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^* w_j(\tilde{\theta}) + \sum_{j=1}^{K} w_j(\tilde{\theta})\Gamma^*_{j-l} w_l(\tilde{\theta})\right)\tilde{V}^{-1}H_0^* + o_{p^*}(1).
\end{aligned}$$

We use Lemma B.1 in GP (2014) to obtain the final equality, $\tilde{V}^* = H^*\tilde{V}H^{*'} + O_{p^*}(\delta_{NT_H}^{-2}) = \tilde{V} + O_{p^*}(\delta_{NT_H}^{-2})$ and $H^* = H_0^* + O_{p^*}(\delta_{NT_H}^{-2})$ in probability.

Part (b):

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\sum_{j=1}^{K} w_j(\tilde{\theta})(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})\right]\left[\sum_{j=1}^{K} w_j(\tilde{\theta})H^*\tilde{f}_{t-j/m}\right]' \\
&= \sum_{j=1}^{K} w_j(\tilde{\theta})\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^*\tilde{f}_{t-j/m}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})' w_j(\tilde{\theta}) \\
&\quad + \sum_{j=1}^{K}\sum_{l\neq j}^{K} w_j(\tilde{\theta})\frac{1}{\sqrt{T}}\sum_{t=1}^{T} H^*\tilde{f}_{t-l/m}(\tilde{f}^*_{t-j/m} - H^*\tilde{f}_{t-j/m})' w_l(\tilde{\theta}) \\
&= cH^*\left[\sum_{j=1}^{K} w_j^2(\tilde{\theta}) + \sum_{j=1}^{K}\sum_{l\neq j}^{K} w_j(\tilde{\theta})\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{f}'_{t-j/m}\right) w_l(\tilde{\theta})\right]\Gamma^*\left(\frac{1}{T_H}\sum_{s=1}^{T_H}\tilde{f}_s\tilde{f}_s^{*'}\right)\tilde{V}^{*-2} + o_{p^*}(1) \\
&= cH_0^*\left[\sum_{j=1}^{K} w_j^2(\tilde{\theta}) + \sum_{j=1}^{K}\sum_{l\neq j}^{K} w_j(\tilde{\theta})\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{t-l/m}\tilde{f}'_{t-j/m}\right) w_l(\tilde{\theta})\right]\Gamma^*\tilde{V}^{-2}H_0^* + o_{p^*}(1),
\end{aligned}$$

in probability. The final equality is by applying Lemma B.1. in GP (2014) and by $\frac{\tilde{f}^{*'}\tilde{f}}{T_H}\tilde{V}^{*-1} = \tilde{V}^{-1}H^*$ and $H^*\tilde{V}^{*-1} = \tilde{V}^{-1}H^*$. ■

In the remaining part, we prove Theorem 3.1. Recall that

$$e^*_{i,t_h} = \sum_{j=1}^{p_i} \tilde{\phi}_{i,j}(p_i) e^*_{i,t_h-j} + u^*_{i,t_h} \quad \text{for } t_h = 1, \ldots, T_H, \tag{8}$$

where $\tilde{\phi}_i(p_i) = (\tilde{\phi}_{i,j}(p_i), j = 1, \ldots, p_i)$ is Yule-Walker autoregressive parameter estimators. By the fact that $\tilde{\phi}_i(p_i)$ is Yule-Walker estimator, we can represent (8) as moving average process of order ∞ as

$$e^*_{i,t_h} = \sum_{j=0}^{\infty} \tilde{\psi}_{i,j}(p_i) u^*_{i,t_h-j}, \tag{9}$$

with $\tilde{\psi}_{i,0}(p_i) = 1$. By stacking (8) and (9) over $i = 1, \ldots, N$, we can rewrite it as vector representation as follows.

$$e^*_{t_h} = \sum_{j=1}^{p_i} \tilde{\Phi}_j(p) e^*_{t_h-j} + u^*_t, \text{ and} \tag{10}$$

$$e^*_{t_h} = \sum_{j=0}^{\infty} \tilde{\Psi}_j(p) u^*_{t_h-j}, \tag{11}$$

with $\tilde{\Psi}_0(p) = I_N$ and $p = \max(p_1, \ldots, p_N)$. Note that $\tilde{\Phi}_j(p)$ is $N \times N$ high-dimensional matrix, but it is a diagonal matrix by the construction such that $\tilde{\Phi}_j(p) = \text{diag}(\tilde{\phi}_{1,j}(p_1), \ldots, \tilde{\phi}_{N,j}(p_N))$. To prove Theorem 3.1, we include an auxiliary Lemma below.

Lemma D.5

(a) $\sum_{j=0}^{\infty} \|\tilde{\Psi}_j(p) - \Psi_j\| = o_p(1)$, *where* Ψ_j *is MA coefficients for* e_t *such that* $e_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j}$.

(b) $\sum_{j=0}^{\infty} |\tilde{\psi}_{i,j}|^8 = O_p(1)$ *for* $i = 1, \ldots, N$.

Proof of Lemma D.5. To prove Lemma D.5-(a), we use the arguments in Bi, Shang, Yang, and Zhu (2021), specifically, Lemma C.7 in their supplement appendix. The difference is that their bootstrap method is applied to the factors, whereas our bootstrap method is constructed using the idiosyncratic error terms. Using their arguments in the proof of their Lemma C.7 and the fact that $\tilde{e}_{i,t} - e_{i,t} = \tilde{c}_{i,t} - c_{i,t} = O_p(\delta_{NT_H}^{-1})$, we can obtain the same result as in Lemma D.5, which yields $\sum_{j=0}^{\infty} \|\tilde{\Psi}_j(p) - \Psi_j\| = o_p(1)$. For (b), we can use Lemma D.5 and Assumption 3 in the main text to conclude. ■

Proof of Theorem 3.1. Following Lemma D.1, Remark 3 and 4, it is sufficient to show that our bootstrap algorithm described in Section 3 satisfy the bootstrap high level conditions C.1*-C.6*. **Condition C.1***. Part (a): We can show that $E^*(e^*_{i,t_h}) = \sum_{j=0}^{\infty} \tilde{\psi}_{i,j}(p_i) E^*(u^*_{i,t_h-j}) = 0$ since $E^*(u^*_{i,t_h-j}) = 0$ by its construction such that $u^*_{t_h} = \tilde{\Sigma}_u^{1/2} \eta_{t_h}$ with $\eta_{t_h} \sim$ i.i.d.$(0, I_N)$. Part (b): We first write γ^*_{st} as follows.

$$
\begin{aligned}
\gamma^*_{st} &= E^*\left(\frac{1}{N} e^{*\prime}_t e^*_s\right) \\
&= E^*\left[\frac{1}{N}\left(\sum_{j_1=0}^{\infty} \tilde{\Psi}_{j_1}(p) u^*_{t-j_1}\right)' \left(\sum_{j_2=0}^{\infty} \tilde{\Psi}_{j_2}(p) u^*_{s-j_2}\right)\right] \\
&= E^*\left[\frac{1}{N} tr\left(\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} \tilde{\Psi}_{j_1}(p) u^*_{t-j_1} u^{*\prime}_{s-j_2} \tilde{\Psi}'_{j_2}\right)\right] \\
&= tr\left(\frac{1}{N}\sum_{j=0}^{\infty} \tilde{\Psi}_j(p) \tilde{\Sigma}_u \tilde{\Psi}_{s-t+j}(p)'\right),
\end{aligned} \tag{12}
$$

where we obtain the last equality since $E^*(u^*_{t-j_1} u^{*\prime}_{s-j_2}) = 0$ if $t - j_1 \neq s - j_2$. Using (12), we can write our condition as following:

$$
\begin{aligned}
\frac{1}{T_H}\sum_{s,t=1}^{T_H} |\gamma^*_{st}|^2 &= \frac{1}{T_H}\sum_{s,t=1}^{T_H} \left| tr\left(\frac{1}{N}\sum_{j=0}^{\infty} \tilde{\Psi}_j(p)\tilde{\Sigma}_u \tilde{\Psi}_{s-t+j}(p)'\right)\right|^2 \\
&\leq \left(\frac{\|\tilde{\Sigma}_u\|^2}{N}\right)\left(\frac{1}{N}\frac{1}{T_H}\sum_{s,t=1}^{T_H}\left\|\sum_{j=0}^{\infty} \tilde{\Psi}_{s-t+j}(p)'\tilde{\Psi}_j(p)\right\|^2\right) \\
&\leq \left(\frac{\|\tilde{\Sigma}_u\|^2}{N}\right)\frac{1}{N}\frac{1}{T_H}\sum_{s,t=1}^{T_H}\sum_{j=0}^{\infty}\left\|\tilde{\Psi}_{s-t+j}(p)\right\|^2\left\|\tilde{\Psi}_j(p)\right\|^2 = O_p(1).
\end{aligned}
$$

We can show that $\|\tilde{\Sigma}_u\|^2/N = O_p(1)$ since we can show the similar arguments in GP (2020) such that $\|\tilde{\Sigma}_u\| \leq \rho(\tilde{\Sigma}_u)\sqrt{\text{rank}(\tilde{\Sigma}_u)} \leq \rho(\tilde{\Sigma}_u)\sqrt{N}$ under Assumption 4-5 in the main text. We can also show that $\frac{1}{N}\frac{1}{T_H}\sum_{s,t=1}^{T_H}\sum_{j=0}^{\infty}\left\|\tilde{\Psi}_{s-t+j}(p)\right\|^2\left\|\tilde{\Psi}_j(p)\right\|^2 = O_p(1)$ under the summability condition. Part (c): First, note that we can write

$$
\begin{aligned}
&E^*\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(e^*_{i,t_h}e^*_{i,s_h} - E^*(e^*_{i,t_h}e^*_{i,s_h}))\right|^2 \\
&= \frac{1}{N}\sum_{i,j=1}^{N} Cov^*(e^*_{i,t_h}e^*_{i,s_h}, e^*_{j,t_h}e^*_{j,s_h}) \\
&= \frac{1}{N}\sum_{i,j=1}^{N}\sum_{k_1,k_2,k_3,k_4=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_3}\tilde{\psi}_{j,k_4}Cov^*(u^*_{i,t_h-k_1}u^*_{i,s_h-k_2}, u^*_{j,t_h-k_3}u^*_{j,s_h-k_4}).
\end{aligned}
$$

We can write $u_{i,t_h} = a_i'\eta_{t_h} = \sum_{l=1}^{N} a_{il}\eta_{l,t_h}$, where a_i' denotes the i-th row of $\tilde{\Sigma}_u^{1/2}$. For simpler notation, define $Cov^*(e^*_{i,t_h}e^*_{i,s_h}, e^*_{j,t_h}e^*_{j,s_h}) = \Delta_{ij,t_hs_h}$. We can rewrite Δ_{ij,t_hs_h} as follows.

$$
\begin{aligned}
\Delta_{ij,t_hs_h} = &\sum_{k_1,k_2,k_3,k_4=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_3}\tilde{\psi}_{j,k_4}\sum_{l_1,l_2,l_3,l_4=1}^{N} a_{i,l_1}a_{i,l_2}a_{j,l_3}a_{j,l_4} \\
&\times \underbrace{Cov^*(\eta_{l_1,t_h-k_1}\eta_{l_2,s_h-k_2}, \eta_{l_3,t_h-k_3}\eta_{l_4,s_h-k_4})}_{A*}.
\end{aligned}
$$

Since $\eta_{l,t_h} \sim$ i.i.d.$(0, I_N)$, we can consider $A*$ based on the choice of l_i for $i = 1,2,3,4$ and $t_h - k_1$, $s_h - k_2$, $t_h - k_3$ and $s_h - k_4$. We need $l_1 = l_2 = l_3 = l_4$, $l_1 = l_3 \neq l_2 = l_4$, or $l_1 = l_4 \neq l_2 = l_3$ for $A*$ to be non-zero. If $l_1 = l_2 = l_3 = l_4$, we need $t_h - k_1 = s_h - k_2 = t_h - k_3 = s_h - k_4$, $t_h - k_1 = t_h - k_3 \neq s_h - k_2 = s_h - k_4$, or $t_h - k_1 = s_h - k_4 \neq s_h - k_2 = t_h - k_3$. In this case, we have $A* = E^*(\eta^4_{l,t_h}) - 1$ or 1. If $l_1 = l_3 \neq l_2 = l_4$, we need $t_h - k_1 = t_h - k_3$ and $s_h - k_2 = s_h - k_4$, and we have $A* = 1$. Similarly, when $l_1 = l_4 \neq l_2 = l_3$, we need $t_h - k_1 = s_h - k_4$ and $s_h - k_2 = t_h - k_3$, and this yields $A* = 1$. Letting $\bar{\eta} \geq \max\{E^*(\eta^4_{l,t_h}) - 1, 1\}$, we can bound the condition as follows.

$$\begin{aligned}
&\frac{1}{T_H^2}\sum_{s_h,t_h=1}^{T_H}\frac{1}{N}\sum_{i,j=1}^{N}\Delta_{ij,t_hs_h}\\
&\quad\le\bar{\eta}\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,k_2}+\tilde{\psi}_{i,k_1}i,\tilde{k}_2\tilde{\psi}_{j,t-s+k_2}\tilde{\psi}_{j,s-t+k_1}\right)\left(\sum_{l=1}^{N}a_{i,l}a_{j,l}\right)^2\\
&\quad\le\bar{\eta}\left[\underbrace{\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}\right)^2\left(\sum_{l=1}^{N}a_{i,l}a_{j,l}\right)^2}_{=A_{ij}-(I)}+\underbrace{\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,s-t+k}\right)\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,t-s+k}\right)\left(\sum_{l=1}^{N}a_{i,l}a_{j,l}\right)^2}_{=A_{ij}-(II)}\right].
\end{aligned}$$

Thus, the condition is bounded by

$$\bar{\eta}\left(\frac{1}{T^2}\sum_{t,s=1}^{T_H}\frac{1}{N}\sum_{i,j=1}^{N}A_{ij}-(I)+\frac{1}{T^2}\sum_{t,s=1}^{T_H}\frac{1}{N}\sum_{i,j=1}^{N}A_{ij}-(II)\right).$$

We can show that $\sum_{i,j=1}^{N}A_{ij}-(I)=O_p(1)$ which is sufficient to show that the first term is $O_p(1)$. Note that we can bound it further by Cauchy-Schwarz inequality as follows.

$$\frac{1}{N}\sum_{i,j=1}^{N}\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}\right)^2\left(\sum_{l=1}^{N}a_{i,l}a_{j,l}\right)^2\le\left\{\frac{1}{N}\sum_{i,j=1}^{N}\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}\right)^4\right\}^{1/2}\left\{\frac{1}{N}\sum_{i,j=1}^{N}\left(\sum_{l=1}^{N}a_{i,l}a_{j,l}\right)^4\right\}^{1/2}.$$

We can show that for some positive constant M, by repetitive application of Hölder's inequality,

$$\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}\right)^4\le M\left(\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}|^4\right)\le M\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}|^4|\tilde{\psi}_{j,k}|^4.$$

By Cauchy-Schwarz inequality, we can show that

$$\frac{1}{N}\sum_{i,j=1}^{N}\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,k}\right)^4\le M\left(\frac{1}{N}\sum_{i=1}^{N}|\tilde{\psi}_{i,k}|^8\right)^{1/2}\left(\frac{1}{N}\sum_{j=1}^{N}|\tilde{\psi}_{j,k}|^8\right)^{1/2}.$$

We can show that this is $O_p(1)$ by Assumption 3 in the main text. We can also show that

$\left\{\frac{1}{N}\sum_{i,j=1}^N (a_i'a_j)^4\right\}^{1/2} = O_p(1)$, because we have

$$\left(\frac{1}{N}\sum_{i,j=1}^N (a_i'a_j)^2\right)^{1/2} \le \sqrt{tr(\tilde{\Sigma}_u^4)/N} \le \sqrt{\{tr(\tilde{\Sigma}_u^2)\}^2/N} = \|\tilde{\Sigma}_u\|/\sqrt{N} = O_p(1).$$

We can obtain the final equality by Assumption 5 and by applying the arguments in GP (2020) to $\tilde{\Sigma}_u$ such that $\|\tilde{\Sigma}_u\| \le \rho(\tilde{\Sigma}_u)\sqrt{\text{rank}(\tilde{\Sigma}_u)} \le \rho(\tilde{\Sigma}_u)\sqrt{N}$ (in their proof of Theorem 3.1). For the second term involved with $A_{ij} - (II)$, by applying Cauchy-Schwarz inequality, we have

$$\frac{1}{T_H^2}\sum_{s,t=1}^{T_H}\frac{1}{N}\sum_{i,j=1}^N A_{ij} - (II) \le \left\{\frac{1}{N}\sum_{i,j=1}^N\left(\sum_{l=1}^N a_{i,l}a_{j,l}\right)^2\right\}^{1/2}\left\{\frac{1}{N}\sum_{i,j=1}^N\left(\frac{1}{T_H^2}\sum_{s,t=1}^{T_H}\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,s-t+k}\right)^2\right)^2\right\}^{1/2}.$$

We can show that $\left\{\frac{1}{N}\sum_{i,j=1}^N\left(\sum_{l=1}^N a_{i,l}a_{j,l}\right)^2\right\}^{1/2} = O_p(1)$ by using the similar arguments above. For the remaining term, we use Cauchy-Schwarz inequality as follows.

$$\frac{1}{N}\sum_{i,j=1}^N\left(\frac{1}{T_H^2}\sum_{s,t=1}^{T_H}\left(\sum_{k=0}^{\infty}\tilde{\psi}_{i,k}\tilde{\psi}_{j,s-t+k}\right)^2\right)^2 \le \frac{1}{N}\sum_{i,j=1}^N\frac{1}{T_H^2}\left(\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}|^2\frac{1}{T_H}\sum_{s,t=1}^{T_H}|\tilde{\psi}_{j,s-t+k}|^2\right)^2$$

Since we can show that $\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}|^2\frac{1}{T_H}\sum_{s,t=1}^{T_H}|\tilde{\psi}_{j,s-t+k}|^2 = O_p(1)$, the order of the above term is $O_p(N/T_H^2)$.

Condition C.2*. Part (a): By Cauchy-Schwarz inequality, we can bound the condition as follows.

$$\left\|\frac{1}{T_H}\sum_{t=1}^{T_H}\tilde{f}_s\tilde{f}_t'\gamma_{st}^*\right\| \le \left(\frac{1}{T_H}\sum_{s,t=1}^{T_H}\|\tilde{f}_s\tilde{f}_t'\|^2\right)^{1/2}\left(\frac{1}{T_H}\sum_{s,t=1}^{T_H}|\gamma_{st}^*|^2\right)^{1/2} = O_p(1).$$

We can show the term in the first parenthesis $O_p(1)$ since we can show that $\frac{1}{T_H}\sum_{t=1}^{T_H}\|\tilde{f}_t\|^4 = O_p(1)$ by using Lemma C.1-(i) in GP (2014) and use Cauchy-Schwarz inequality. The term in the second parenthesis is $O_p(1)$ by Condition C.1*-(b). Part (b): For simpler notation, in the remaining proof, we let $\tilde{\psi}_{i,j} = \tilde{\psi}_{i,j}(p_i)$ and $\tilde{\Psi}_j = \tilde{\Psi}_j(p)$. Note that we can rewrite the condition as follows.

$$\frac{1}{T_H}\sum_{t=1}^{T_H}\frac{1}{T_H}\sum_{s,l=1}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^2\frac{1}{N}\sum_{i,j=1}^{N}Cov^*(e_{i,t}^*e_{i,s}^*,e_{j,t}^*e_{j,l}^*).$$

By considering the combination of i,j and t,s and l, the covariance term $Cov^*(e_{i,t}^*e_{i,s}^*,e_{j,t}^*e_{j,l}^*)$ can be further bounded as follows.

$$\begin{aligned}Cov^*(e_{i,t}^*e_{i,s}^*,e_{j,t}^*e_{j,l}^*)\leq\bar{\eta}&\left\{\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}\right)\left(\sum_{m=1}^{N}a_{i,m}a_{j,m}\right)^2\right.\\&\left.+\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,t-s+k_2}\tilde{\psi}_{j,l-t+k_1}\right)\left(\sum_{m=1}^{N}a_{i,m}a_{j,m}\right)^2\right\}\\&=\bar{\eta}(B_{ij}-(I)+B_{ij}-(II)),\end{aligned}$$

where we denote $B_{ij}-(I)=\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}\right)\left(\sum_{m=1}^{N}a_{i,m}a_{j,m}\right)^2$ and $B_{ij}-(II)=\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,t-s+k_2}\tilde{\psi}_{j,l-t+k_1}\right)\left(\sum_{m=1}^{N}a_{i,m}a_{j,m}\right)^2$. Then, using this bound on the covariance term, the condition is bounded by the following equation.

$$\bar{\eta}\left[\frac{1}{T_H^2}\sum_{t,s,l=1}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^2\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)+\frac{1}{T_H^2}\sum_{t,s,l=1}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^2\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(II)\right]$$

The first term in the square bracket can be bounded by Cauchy-Schwarz inequality as follows.

$$\frac{1}{T_H^2}\sum_{t,s,l=1}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^2\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)\leq\frac{1}{T_H}\sum_{t=1}^{T_H}\left(\frac{1}{T_H}\sum_{s,l=1}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^4\right)^{1/2}\left(\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left|\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)\right|^2\right)^{1/2}$$

We can show that $\frac{1}{T_H}\sum_{s,l=}^{T_H}\|\tilde{f}_s'\tilde{f}_l\|^4=O_p(1)$ by applying Lemma C.1 in GP (2014) with $p=8$ (this can be verified under our Assumption 1 in the main text). To show that $\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left|\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)\right|^2=O_p(1)$, we first bound it by Cauchy-Schwarz inequality as follows.

$$\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left|\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)\right|^2 \le \left(\frac{1}{N}\sum_{i,j=1}^{N}(a_i'a_j)^2\right)\left(\frac{1}{N}\sum_{i,j=1}^{N}\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}\right)^2\right).$$

As in the proof of Condition C.1*-(c), we can show that $\frac{1}{N}\sum_{i,j=1}^{N}(a_i'a_j)^2 = O_p(1)$. First, note that by using Hölder's inequality, we can show that $\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}\right)^2 \le M$ $\sum_{k_1,k_2=0}^{\infty}|\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}|^2$, for some positive constant M. Then, we apply Cauchy-Schwarz inequality and Hölder's inequality to obtain the following inequality. For some positive constant M,

$$\begin{aligned}
&\frac{1}{N}\sum_{i,j=1}^{N}\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,k_1}\tilde{\psi}_{j,l-s+k_2}\right)^2\\
&\quad\le M\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{k_1=0}^{\infty}|\tilde{\psi}_{i,k_1}|^4\sum_{k_2=0}^{\infty}|\tilde{\psi}_{i,k_2}|^4\right)^{1/2}\left(\frac{1}{N}\sum_{j=1}^{N}\sum_{k_1=0}^{\infty}|\tilde{\psi}_{j,k_1}|^4\left(\sum_{k_2=0}^{\infty}\frac{1}{T_H}\sum_{s,l=1}^{T_H}|\tilde{\psi}_{j,l-s+k_2}|^2\right)^2\right)^{1/2}.
\end{aligned}$$

Note that $\frac{1}{T_H}\sum_{s,l=1}^{T_H}|\tilde{\psi}_{j,l-s+k_2}|^2 = \sum_{\tau=0}^{T_H-1}\left(1-\frac{\tau}{T_H}\right)|\tilde{\psi}_{j,\tau+k_2}|^2 \le \sum_{\tau=0}^{\infty}|\tilde{\psi}_{j,\tau+k_2}|^2$. Then, since $\sum_{k_2=0}^{\infty}\sum_{\tau=0}^{\infty}|\tilde{\psi}_{j,\tau+k_2}|^2 = \sum_{k_3=0}^{\infty}(k_3+1)|\tilde{\psi}_{j,k_3}|^2$, we can show that $\left(\sum_{k_2=0}^{\infty}\frac{1}{T_H}\sum_{s,l=1}^{T_H}|\tilde{\psi}_{j,l-s+k_2}|^2\right)^2 \le$ $M_1\sum_{k_3=0}^{\infty}(k_3+1)^2|\tilde{\psi}_{j,k_3}|^4$ for some positive constant M_1. Therefore, we can show that the second term is $O_p(1)$ by Assumption 3 with $r=2$. By Assumption 3, we can show that $\frac{1}{N}\sum_{i=1}^{N}\sum_{k_1=0}^{\infty}|\tilde{\psi}_{i,k_1}|^4\sum_{k_2=0}^{\infty}|\tilde{\psi}_{i,k_2}|^4 = O_p(1)$ and we can also show that the remaining term in the above inequality is $O_p(1)$. Next, we show that $\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left|\frac{1}{N}\sum_{i,j=1}^{N}B_{ij}-(I)\right|^2 = O_p(1)$. By applying Cauchy-Schwarz inequality repetitively, it is sufficient to show that

$$\begin{aligned}
&\frac{1}{N}\sum_{i,j=1}^{N}\frac{1}{T_H}\sum_{s,l=1}^{T_H}\left|\frac{1}{T_H}\sum_{t=1}^{T_H}\left(\sum_{k_1,k_2=0}^{\infty}\tilde{\psi}_{i,k_1}\tilde{\psi}_{i,k_2}\tilde{\psi}_{j,t-s+k_2}\tilde{\psi}_{j,l-t+k_1}\right)\right|^2\\
&\quad\le M\frac{1}{T_H}\left(\frac{1}{N}\sum_{i,j=1}^{N}\left|\frac{1}{T_H}\sum_{t,l=1}^{T_H}\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}\tilde{\psi}_{j,l-t+k}|^2\right|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i,j=1}^{N}\left|\frac{1}{T_H}\sum_{t,s=1}^{T_H}\sum_{k_2=0}^{\infty}|\tilde{\psi}_{i,k}\tilde{\psi}_{j,t-s+k}|^2\right|^2\right)^{1/2}\\
&= O_p(1).
\end{aligned}$$

In fact, we can use Cauchy-Schwarz inequality and Assumption 3 to show that the term is $O_p(1)$. For example, we can show that

$$\begin{aligned}\frac{1}{N}\sum_{i,j=1}^{N}\left|\frac{1}{T_H}\sum_{t,l=1}^{T_H}\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}\tilde{\psi}_{j,l-t+k}|^2\right|^2 &= \frac{1}{N}\sum_{i,j=1}^{N}\left|\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}|^2\frac{1}{T_H}\sum_{t,l=1}^{T_H}|\tilde{\psi}_{j,l-t+k}|^2\right|^2\\ &\leq \left(M_1\frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{\infty}|\tilde{\psi}_{i,k}|^8\right)^{1/2}\left(M_2\frac{1}{N}\sum_{j=1}^{N}\sum_{k=0}^{\infty}|(1+k)^4|\tilde{\psi}_{j,k}|^8\right)^{1/2}\\ &= O_p(1),\end{aligned}$$

for some positive constants M_1 and M_2. We obtain the final equality by Assumption 3 with $r = 4$. Part (c): First, note that we can write the condition as follows.

$$E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{t=1}^{T_H}\sum_{i=1}^{N}\tilde{f}_t\tilde{\lambda}_i' e_{i,t}^*\right\|^2 = \frac{1}{T_H}\sum_{t,s=1}^{T_H} tr(\tilde{f}_t\tilde{f}_s')E^*\left(\frac{e_s^{*\prime}\tilde{\Lambda}}{\sqrt{N}}\frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}}\right)$$

Since $E^*\left(\frac{e_s^{*\prime}\tilde{\Lambda}}{\sqrt{N}}\frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}}\right) = E^*\left[tr\left(\frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}}\frac{e_s^{*\prime}\tilde{\Lambda}}{\sqrt{N}}\right)\right]$, we focus on $E^*(e_t^* e_s^{*\prime})$. Under vector MA(∞) representation of e_t^*, we can write it as follows.

$$E^*(e_t^* e_s^{*\prime}) = \sum_{k_1,k_2=0}^{\infty}\tilde{\Psi}_{k_1}E^*(u_{t-k_1}^* u_{s-k_2}^{*\prime})\tilde{\Psi}_{k_2} = \sum_{k=0}^{\infty}\tilde{\Psi}_k\tilde{\Sigma}_u\tilde{\Psi}_{s-t+k}'$$

By plugging this back into the condition and using Cauchy-Schwarz inequality,

$$\begin{aligned}E^*\left\|\frac{1}{\sqrt{T_H N}}\sum_{t=1}^{T_H}\sum_{i=1}^{N}\tilde{f}_t\tilde{\lambda}_i' e_{i,t}^*\right\|^2 &= \frac{1}{T_H}\sum_{t,s=1}^{T_H} tr(\tilde{f}_t\tilde{f}_s')tr\left(\frac{\tilde{\Lambda}'\sum_{k=0}^{\infty}\tilde{\Psi}_k\tilde{\Sigma}_u\tilde{\Psi}_{s-t+k}'\tilde{\Lambda}}{N}\right)\\ &\leq \left(\frac{1}{T_H}\sum_{t,s=1}^{T_H}|tr(\tilde{f}_t\tilde{f}_s')|^2\right)^{1/2}\left(\frac{1}{T_H}\sum_{t,s=1}^{T_H}\left|tr(\tilde{\Gamma}_{s-t})\right|^2\right)^{1/2},\end{aligned}$$

where we denote $\tilde{\Gamma}_{s-t} = \frac{1}{N}\tilde{\Lambda}'\left(\sum_{k=0}^{\infty}\tilde{\Psi}_k\tilde{\Sigma}_u\tilde{\Psi}_{s-t+k}'\right)\tilde{\Lambda}$. We can show that the first term is $O_p(1)$ by Assumption 1 and using the results in Lemma C.1 in GP (2014). For the second term, it is sufficient to show that $tr(\tilde{\Gamma}_t) = O_p(1)$. This is implied by Condition C.6*-(b), which will be verified. Part (d): We can rewrite the condition as follows.

$$\frac{1}{T_H}\sum_{t=1}^{T_H} E^*\left\|\frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}}\right\|^2 = \frac{1}{T_H}\sum_{t=1}^{T_H} E^*\left[tr\left(\frac{\tilde{\Lambda}' e_t^* e_t^{*\prime}\tilde{\Lambda}}{N}\right)\right] = \frac{1}{T_H}\sum_{t=1}^{T_H} tr\left(\frac{\tilde{\Lambda}' E^*(e_t^* e_t^{*\prime})\tilde{\Lambda}}{N}\right)$$

As we have shown previously in the proof of Condition C.2*-(c), we can write $E^*(e_t^* e_t^{*\prime}) = \sum_{k=0}^{\infty}\tilde{\Psi}_k\tilde{\Sigma}_u\tilde{\Psi}_k'$. Therefore, the condition is $\frac{1}{T_H}\sum_{t=1}^{T_H} tr\left(\frac{\tilde{\Lambda}'\sum_{k=0}^{\infty}\tilde{\Psi}_k\tilde{\Sigma}_u\tilde{\Psi}_k'\tilde{\Lambda}}{N}\right)$, and this is $O_p(1)$ given that $tr(\tilde{\Gamma}_0) = O_p(1)$. Part (e): To verify this condition, we use $r = 1$ (recall that r is the number of factors) for a simpler notation. Therefore, it suffices to show that $Var^*(A^*) = o_p(1)$, where $A^* = \frac{1}{T_H}\sum_{t=1}^{T_H}\left(\frac{\tilde{\Lambda}' e_t^*}{\sqrt{N}}\right)\left(\frac{e_t^{*\prime}\tilde{\Lambda}}{\sqrt{N}}\right)$. Note that

$$\begin{aligned}
Var^*(A^*) &= \frac{1}{T_H^2}\sum_{t,s=1}^{T_H}\frac{1}{N^2}\sum_{i,j,k,l}^{N}\tilde{\lambda}_i\tilde{\lambda}_j\tilde{\lambda}_k\tilde{\lambda}_l Cov^*(e_{i,t}^* e_{j,t}^*, e_{l,s}^* e_{k,s}^*)\\
&\leq 2\bar{\eta}\frac{1}{T_H^2}\sum_{t,s=1}^{T_H}\frac{1}{N^2}\sum_{i,j,k,l}^{N}\tilde{\lambda}_i\tilde{\lambda}_j\tilde{\lambda}_k\tilde{\lambda}_l\left(\sum_{p_1,p_2=0}^{\infty}\tilde{\psi}_{i,p_1}\tilde{\psi}_{j,p_2}\tilde{\psi}_{l,s-t+p_1}\tilde{\psi}_{k,s-t+p_2}\right)\\
&\quad\times\left(\sum_{m_1,m_2=1}^{N} a_{i,m_1}a_{j,m_2}a_{l,m_1}a_{k,m_2}\right)\\
&= 2\bar{\eta}\frac{1}{T_H^2}\sum_{t,s=1}^{T_H}\left\{\left(\frac{1}{N}\sum_{i,l=1}^{N}\tilde{\lambda}_i\tilde{\lambda}_l\right)\left(\sum_{p_1=0}^{\infty}\tilde{\psi}_{i,p_1}\tilde{\psi}_{l,s-t+p_1}\right)\left(\sum_{m_1=1}^{N} a_{i,m_1}a_{l,m_1}\right)\right\}^2\\
&= 2\bar{\eta}\frac{1}{T_H^2}\sum_{t,s=1}^{T_H}\left(\frac{\tilde{\Lambda}'\sum_{p=0}^{\infty}\tilde{\Psi}_p\tilde{\Sigma}_u\tilde{\Psi}_{s-t+p}\tilde{\Lambda}}{N}\right)^2,
\end{aligned}$$

where we obtain the second inequality by taking account of the covariance term given the combination of i, j, k, and l and t and s, similar to the proof of Condition C.1*-(c). Note that given that $tr(\tilde{\Gamma}_{s-t}) = O_p(1)$, we can show that $\frac{1}{T_H}\sum_{t,s=1}^{T_H}\tilde{\Gamma}_{s-t}^2 = O_p(1)$. Therefore, $Var^*(A^*) = O_p(1/T_H) = o_p(1)$. The proof to verify **Condition C.3*** is very similar to the proof of Condition C.2*. For example, Condition C.3*-(b) and (c) can be verified given that $tr(\tilde{\Gamma}_\tau) = O_p(1)$ with $\tau \neq 0$.

Condition C.4* Part (a): Given that ε_t^* and $e_{t-j/m}^*$ are independent in Assumption 2, it is sufficient to show that

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{NT}\sum_{s,l=1}^{T}\sum_{i=1}^{N}Cov^*(e^*_{i,t-j/m}e^*_{i,s-j/m},e^*_{i,t-j/m}e^*_{i,l-j/m}) = O_p(1).$$

We show a similar term is $O_p(1)$ in Condition C.2*-(b). Part (b): Similarly, given the independence of ε_t^* and $e^*_{i,t-j/m}$, it suffices to show that $E^*\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{\tilde{\Lambda}'e^*_{t-j/m}}{\sqrt{N}}\right\|^2 = O_p(1)$, which is verified in Condition C.2*-(c). **Condition C.5*** and **Condition C.6***-(a) can be verified using the arguments in GP (2014), since ε_t^* is constructed in the same way. **Condition C.6***- Part (b): Note that $\tilde{\Gamma}_k$ can be rewritten as follows.

$$\tilde{\Gamma}_k = \frac{1}{T_H}\sum_{t=1}^{T_H}\frac{1}{N}\tilde{\Lambda}'\sum_{p=0}^{\infty}\tilde{\Psi}_p\tilde{\Sigma}_u\tilde{\Psi}'_{p-k}\tilde{\Lambda} = \frac{\tilde{\Lambda}'\tilde{\Sigma}_{e,k}\tilde{\Lambda}}{N},$$

where we let $\tilde{\Sigma}_{e,k} \equiv \sum_{p=0}^{\infty}\tilde{\Psi}_p\tilde{\Sigma}_u\tilde{\Psi}'_{p-k}$. Let $\bar{\Gamma}_k = \frac{\Lambda'\tilde{\Sigma}_{e,k}\Lambda}{N}$. Then, by adding and subtracting appropriately, we have the following:

$$\begin{aligned}\tilde{\Gamma}_k - H_0\Gamma_k H_0' &= \tilde{\Gamma}_k - H_0\bar{\Gamma}_k H_0' + H_0\bar{\Gamma}_k H_0' - H_0\Gamma_k H_0' \\ &= \underbrace{(\tilde{\Gamma}_k - H_0\bar{\Gamma}_k H_0')}_{\equiv D_1} + \underbrace{H_0(\bar{\Gamma}_k - \Gamma_k)H_0'}_{\equiv D_2}.\end{aligned}$$

We can show that D_1 and D_2 are $o_p(1)$. In order to show that $D_2 = o_p(1)$, it is sufficient to show that $\tilde{\Sigma}_{e,k} - \Sigma_{e,k} \to 0$, where $\Sigma_{e,k} \equiv \sum_{p=0}^{\infty}\Psi_p\Sigma_u\Psi'_{p-k}$ with $\Sigma_u = E(u_t u_t')$. Note that we can expand $\tilde{\Sigma}_{e,k} - \Sigma_{e,k}$ as follows.

$$\tilde{\Sigma}_{e,k} - \Sigma_{e,k} = \underbrace{\sum_{p=0}^{\infty}(\tilde{\Psi}_p - \Psi_p)\tilde{\Sigma}_u\tilde{\Psi}_{p-k}}_{D_{21}} + \underbrace{\sum_{p=0}^{\infty}\Psi_p(\tilde{\Sigma}_u - \Sigma_u)\tilde{\Psi}_{p-k}}_{D_{22}} + \underbrace{\sum_{p=0}^{\infty}\Psi_p\Sigma_u(\tilde{\Psi}_{p-k} - \Psi_{p-k})}_{D_{23}}.$$

We can show that $D_{22} = o_p(1)$ since $\rho(\tilde{\Sigma}_u - \Sigma_u) \xrightarrow{p} 0$ under Assumptions 4-5 using the arguments in GP (2020). We can show that D_{21} and D_{23} are of order $o_p(1)$ by Lemma D.5.

Next, we show that D_1 is $o_p(1)$. We can decompose D_1 further as follows.

$$D_1 = \underbrace{\frac{1}{N}(\tilde{\Lambda} - \Lambda H^{-1})'\tilde{\Sigma}_{e,k}(\tilde{\Lambda} - \Lambda H^{-1})}_{D_{11}} + \underbrace{\frac{1}{N}H^{-1\prime}\Lambda'\tilde{\Sigma}_{e,k}(\tilde{\Lambda} - \Lambda H^{-1})}_{D_{12}} + \underbrace{\frac{1}{N}(\tilde{\Lambda} - \Lambda H^{-1})'\tilde{\Sigma}_{e,k}\Lambda H^{-1}}_{D_{12}'}.$$

$D_{11} = o_p(1)$ by applying Cauchy-Schwarz inequality as follows.

$$\|D_{11}\| \leq \underbrace{\left\|\frac{1}{\sqrt{N}}(\tilde{\Lambda} - \Lambda H^{-1})\right\|^2}_{=o_p(1)} \underbrace{\left\|\sum_{p=0}^{\infty} \tilde{\Psi}_p \tilde{\Sigma}_u \tilde{\Psi}_{p-k}'\right\|}_{=O_p(1)} = o_p(1),$$

where we use the fact that

$$\left\|\sum_{p=0}^{\infty} \tilde{\Psi}_p \tilde{\Sigma}_u \tilde{\Psi}_{p-k}'\right\| \leq \left\|\sum_{p=0}^{\infty} \tilde{\Psi}_p \tilde{\Psi}_{p-k}'\right\| \left\|\tilde{\Sigma}_u\right\| \leq \sum_{p=0}^{\infty} \left\|\tilde{\Psi}_p \tilde{\Psi}_{p-k}'\right\| \rho(\tilde{\Sigma}_u) = O_p(1),$$

and use the arguments in GP (2020). Since we have

$$\|D_{12}\| \leq \|H^{-1}\| \left\|\Lambda/\sqrt{N}\right\| \left\|\tilde{\Sigma}_{e,k}\right\| \left(\left\|\frac{1}{\sqrt{N}}(\tilde{\Lambda} - \Lambda H^{-1})\right\|^2\right)^{1/2},$$

we can show that this is $o_p(1)$ using similar arguments as we did for D_{11}. ■

E Additional simulation results

E.1 Simulation: results of DGP 1 and 2 of the factor-MIDAS regression model

Table 1 presents the results of DGP 1 and 2 in each panel. The results indicate that there is no bias when using the true factor, however, a bias does exist when using the estimated factor as a regressor. Increasing the sample size in both cross-sectional and time series dimensions results in a decrease in bias. If the cross-sectional dimension is small (50 and 100), the plug-in bias tends to overestimate the bias size. Both bootstrap methods perform similarly and replicate bias size well. When no method is used to correct the bias, size distortion occurs in terms of coverage rates. The plug-in bias somewhat recovers the size distortion, but bootstrap methods outperform the plug-in bias method. The results of DGP 1 and DGP 2 are similar, and both bootstrap methods are valid for these scenarios since the idiosyncratic error terms are randomly generated from a standard normal distribution.

E.2 Simulation experiment: increase in autoregressive coefficient

Table 2 shows the bias and 95% coverage rate of β when the idiosyncratic error term follows simple AR (1) process as:

$$e_{i,t_h} = \rho_i e_{i,t_h-1} + v_{i,t_h} \text{ for } t_h = 1, \ldots, T_H$$

where v_{i,t_h} is i.i.d. randomly generated from $N(0,1)$. We let ρ_i indicate the auto-regressive coefficient, which implies the persistence of auto-regressive process. For simplicity, we impose that each variable shares same autoregressive coefficient, $\rho_i = \rho$. In Table 2, we compare the results by varying persistence. We increase the coefficient from 0 to 0.7. When the persistence in the idiosyncratic error term is $\rho = 0.5$, the bias is around twice bigger than the bias where there is no serial-dependence. Moreover, the size of bias increase as the persistence increases.

E.3 Simulation experiment: unrestricted MIDAS regression model

Table 3-5 show the performance of bootstrap methods (wild bootstrap and AR-sieve + CSD bootstrap method) as well as plug-in bias estimation method under the framework of unrestricted MIDAS regression model. We consider the unrestricted MIDAS regression with

Table 1: DGP 1 & DGP 2 - Bias and coverage rate of 95% CIs for β

		$N=50$			$N=100$			$N=200$		
		$T=50$	100	200	50	100	200	50	100	200
		$T_H=150$	300	600	150	300	600	150	300	600
	bias									
DGP 1: homo & homo	True Factor	-0.01	-0.01	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00
	Estimated Factor	-0.32	-0.31	-0.29	-0.20	-0.17	-0.16	-0.12	-0.10	-0.08
	Plug-in	-0.38	-0.34	-0.32	-0.21	-0.19	-0.18	-0.10	-0.10	-0.09
	WB	-0.25	-0.24	-0.23	-0.16	-0.15	-0.14	-0.11	-0.09	-0.08
	AR-sieve+CSD	-0.24	-0.24	-0.23	-0.16	-0.15	-0.14	-0.10	-0.09	-0.08
	95% coverage rate									
	Estimated Factor	84.8	82.0	73.9	89.6	90.5	88.3	91.7	92.7	93.4
	Plug-in	87.6	89.1	89.3	90.4	92.1	92.4	91.2	92.7	93.6
	WB	94.1	94.7	93.3	95.0	95.6	94.5	92.7	95.4	94.9
	AR-sieve+CSD	95.8	94.9	92.4	95.8	96.1	95.0	96.0	96.3	95.3
	bias									
DGP 2: hetero & homo	True Factor	-0.01	0.00	0.00	0.00	0.01	-0.01	0.01	-0.01	0.00
	Estimated Factor	-0.34	-0.30	-0.29	-0.19	-0.16	-0.16	-0.10	-0.10	-0.09
	Plug-in	-0.37	-0.34	-0.32	-0.20	-0.19	-0.18	-0.10	-0.10	-0.09
	WB	-0.24	-0.24	-0.23	-0.16	-0.15	-0.14	-0.10	-0.09	-0.08
	AR-sieve+CSD	-0.24	-0.24	-0.23	-0.16	-0.15	-0.14	-0.10	-0.09	-0.08
	95% coverage rate									
	Estimated Factor	78.1	76.2	68.4	85.9	88.1	86.2	88.7	91.5	91.6
	Plug-in	82.7	86.8	88.3	86.6	89.8	92.5	88.9	92.3	92.5
	WB	91.7	93.0	93.1	92.6	93.3	94.2	91.0	94.4	94.0
	AR-sieve+CSD	92.5	92.9	92.2	94.0	95.2	93.8	93.5	94.8	94.8

In DGP 1, both error terms are homoskedastic. In DGP 2, MIDAS regression error terms are heteroskedastic and idiosyncratic error terms are homoskedastic. The results of coverage rates, when we use the estimated factors and plug-in bias, are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

a single factor as follows.

$$y_t = \beta + \sum_{j=1}^{K} \alpha_k f_{t-j/3} + \varepsilon_t$$

$$X_{t-k/3} = \Lambda f_{t-k/3} + e_{t-k/3},$$

Table 2: Bias and 95% coverage rate of β

N	T_H	$\rho=0$		$\rho=0.5$		$\rho=0.6$		$\rho=0.7$	
		bias	95%	bias	95%	bias	95%	bias	95%
50	150	-0.3380	84.7	-0.5887	68.02	-0.6808	60.42	-0.7993	49.18
	300	-0.3100	81.76	-0.5362	57.94	-0.6197	48.16	-0.7278	35.18
	600	-0.2890	74	-0.4970	40.96	-0.5746	29.32	-0.6761	17.2
100	150	-0.2022	89.82	-0.3763	83.18	-0.4450	79.34	-0.5372	72.62
	300	-0.1709	90.72	-0.3157	81.1	-0.3729	75.68	-0.4502	67.1
	600	-0.1565	88.7	-0.2849	75.36	-0.3358	67.44	-0.4047	56.16
200	150	-0.1343	91.48	-0.2639	87.6	-0.3163	85.38	-0.3890	81.8
	300	-0.1027	92.5	-0.1996	89.18	-0.2393	87.28	-0.2943	83.54
	600	-0.0865	92.44	-0.1647	88.02	-0.1968	85.48	-0.2411	80.7

for $k = 0, 1, 2$ and $t = 1, \ldots, T$. The simulation design is identical to that in Section 4 in the main text: $f_{t-k/m} \sim$ i.i.d. $N(0, 1)$ and $\lambda_i \sim$ i.i.d. $U[0, 1]$. We consider six data generating processes as detailed in Table 1 in the main text. In this setup, y_t is predicted using six lags of the factor ($K = 6$). We set $\beta = 0$ and $\alpha_k = \alpha^k$ with $\alpha = 0.8$. The estimation procedure is similar to restricted MIDAS, which proceeds in two steps: we first estimate the factors from $X_{t-k/m}$ and then in the second step, we regress y_t on the temporally aggregated estimated factors up to six lags. We report the bias in α_1 associated with the true factor, estimated factor, plug-in estimation method, as well as two bootstrap methods: the wild bootstrap method and the AR-sieve + CSD bootstrap method. In addition, we provide the 95% coverage rates associated with the estimated factor, plug-in estimation method, and both bootstrap methods. Note that the wild bootstrap is not valid in DGPs 4 to 6.

DGPs 1 to 3 yield comparable outcomes: the plug-in estimation method and the two bootstrap methods are perform similarly, and effectively capture the size of the bias. Regarding the coverage rate, the bootstrap methods outperform the plug-in estimation method.

In DGP 4, where the idiosyncratic error terms of the factor model are serially dependent, the AR-sieve + CSD bootstrap method outperforms the plug-in estimation method in terms of replicating the bias and correcting the distortion induced by the bias. In DGP 5, the plug-in estimation method performs the best in estimating the bias size. In terms of coverage rate, the plug-in estimation method outperforms the wild bootstrap method when N is small, while the AR-sieve + CSD bootstrap method outperforms other two methods across all sample sizes. Finally, in DGP 6, both the plug-in and the AR-sieve + CSD bootstrap methods replicate the bias size well, with the AR-sieve + CSD bootstrap method performing the best at recovering the distortion in the coverage rate.

Table 3: DGP 1 & DGP 2 - Bias and coverage rate of 95% CIs for β

		$N = 50$			$N = 100$			$N = 200$		
		$T = 50$	100	200	50	100	200	50	100	200
		$T_H = 150$	300	600	150	300	600	150	300	600
	bias									
DGP 1: homo & homo	True Factor	0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00
	Estimated Factor	-0.10	-0.10	-0.09	-0.07	-0.05	-0.05	-0.04	-0.03	-0.03
	Plug-in	-0.09	-0.08	-0.08	-0.05	-0.05	-0.04	-0.03	-0.02	-0.02
	WB	-0.08	-0.08	-0.07	-0.05	-0.05	-0.04	-0.04	-0.03	-0.03
	AR-sieve+CSD	-0.08	-0.08	-0.07	-0.05	-0.05	-0.04	-0.04	-0.03	-0.03
	95% coverage rate									
	Estimated Factor	86.8	81.5	71.0	91.6	90.6	88.1	93.3	94.1	93.6
	Plug-in	89.7	89.7	90.0	92.1	92.6	92.5	93.2	94.0	94.4
	WB	94.3	93.5	92.6	95.3	94.5	93.8	95.7	95.3	95.1
	AR-sieve+CSD	94.4	93.1	92.6	95.6	94.4	94.0	95.6	95.3	95.1
	bias									
DGP 2: hetero & homo	True Factor	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	Estimated Factor	-0.11	-0.10	-0.09	-0.07	-0.05	-0.05	-0.04	-0.03	-0.03
	Plug-in	-0.09	-0.08	-0.08	-0.05	-0.05	-0.04	-0.03	-0.02	-0.02
	WB	-0.08	-0.08	-0.07	-0.05	-0.05	-0.04	-0.03	-0.03	-0.03
	AR-sieve+CSD	-0.08	-0.08	-0.07	-0.05	-0.05	-0.04	-0.03	-0.03	-0.03
	95% coverage rate									
	Estimated Factor	78.0	74.5	65.4	86.6	87.3	86.2	89.2	91.6	92.0
	Plug-in	84.3	86.8	89.3	88.2	90.2	91.5	89.7	92.1	93.7
	WB	90.7	91.2	91.7	92.5	92.7	92.8	92.6	93.9	94.6
	AR-sieve+CSD	90.9	91.3	91.5	92.8	92.7	93.1	92.6	93.7	94.5

In DGP 1, both error terms are homoskedastic. In DGP 2, MIDAS regression error terms are heteroskedastic and idiosyncratic error terms are homoskedastic. The results of coverage rates, when we use the estimated factors and plug-in bias, are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

F Other empirical result

In Table 6, we present the results after excluding the COVID pandemic period. The results are similar to Table 5 in the main text. When using the bootstrap method, the confidence

Table 4: DGP 3 & DGP 4 - Bias and coverage rate of 95% CIs for β

		$N = 50$			$N = 100$			$N = 200$		
		$T = 50$	100	200	50	100	200	50	100	200
		$T_H = 150$	300	600	150	300	600	150	300	600
	bias									
DGP 3: hetero & hetero	True Factor	-0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	0.00
	Estimated Factor	-0.11	-0.11	-0.10	-0.07	-0.05	-0.05	-0.04	-0.03	-0.03
	Plug-in	-0.10	-0.09	-0.09	-0.05	-0.05	-0.05	-0.03	-0.03	-0.03
	WB	-0.09	-0.09	-0.08	-0.06	-0.05	-0.05	-0.04	-0.03	-0.03
	AR-sieve+CSD	-0.09	-0.08	-0.08	-0.06	-0.05	-0.05	-0.04	-0.03	-0.03
	95% coverage rate									
	Estimated Factor	75.9	72.7	61.4	85.3	87.6	84.4	89.2	91.6	91.3
	Plug-in	84.6	87.6	88.5	87.7	90.5	91.8	89.3	92.2	93.4
	WB	91.1	92.0	91.7	91.8	92.7	93.3	92.9	93.6	94.2
	AR-sieve+CSD	91.1	91.7	90.6	91.9	92.6	93.1	92.7	93.9	94.1
	bias									
DGP 4: hetero & AR	True Factor	-0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	0.00
	Estimated Factor	-0.15	-0.14	-0.13	-0.10	-0.07	-0.07	-0.06	-0.05	-0.04
	Plug-in	-0.08	-0.08	-0.08	-0.05	-0.05	-0.05	-0.03	-0.03	-0.02
	WB	-0.08	-0.08	-0.08	-0.05	-0.05	-0.05	-0.04	-0.03	-0.03
	AR-sieve+CSD	-0.10	-0.10	-0.09	-0.07	-0.07	-0.06	-0.05	-0.04	-0.04
	95% coverage rate									
	Estimated Factor	69.6	63.1	48.7	81.2	83.3	78.9	87.1	89.7	89.0
	Plug-in	80.1	83.1	81.4	85.0	89.3	89.6	88.1	91.2	92.4
	WB	87.7	88.1	85.0	90.6	92.2	91.3	92.3	93.3	93.6
	AR-sieve+CSD	89.7	90.5	88.3	92.0	93.0	92.8	92.6	93.9	94.2

In DGP 3, both error terms are heteroskedastic. In DGP 4, the idiosyncratic error term is generated as the autoregressive process of lag 1 for each variable and with heteroskedastic. For coverage rates, the results for estimated factors and plug-ins are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

intervals associated with the factors shift. However, the bias does not have a significant impact on the estimates for the lags of the dependent variable. Additionally, it is worth noting that as we exclude the COVID period, the sign of the estimates associated with the

Table 5: DGP 5 & DGP 6 - Bias and coverage rate of 95% CIs for β

		$N=50$			$N=100$			$N=200$		
		$T=50$	100	200	50	100	200	50	100	200
		$T_H=150$	300	600	150	300	600	150	300	600
	bias									
DGP 5: hetero & CSD	True Factor	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	Estimated Factor	-0.09	-0.09	-0.09	-0.06	-0.05	-0.05	-0.03	-0.03	-0.02
	Plug-in	-0.07	-0.06	-0.06	-0.04	-0.04	-0.04	-0.02	-0.02	-0.02
	WB	-0.03	-0.03	-0.03	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	AR-sieve+CSD	-0.05	-0.05	-0.05	-0.03	-0.03	-0.03	-0.02	-0.02	-0.02
	95% coverage rate									
	Estimated Factor	80.9	77.1	67.6	87.8	88.1	86.8	90.1	91.9	92.7
	Plug-in	84.6	86.4	86.3	88.5	90.2	91.3	89.8	92.5	93.6
	WB	89.3	87.5	82.7	92.1	91.7	91.0	92.9	93.7	94.2
	AR-sieve+CSD	90.7	90.3	88.8	92.6	92.5	92.7	92.8	93.9	94.7
	bias									
DGP 6: hetero & CSD+AR	True Factor	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	Estimated Factor	-0.12	-0.12	-0.12	-0.07	-0.06	-0.06	-0.04	-0.04	-0.03
	Plug-in	-0.06	-0.06	-0.06	-0.04	-0.04	-0.03	-0.02	-0.02	-0.02
	WB	-0.03	-0.03	-0.03	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	AR-sieve+CSD	-0.06	-0.06	-0.06	-0.04	-0.04	-0.04	-0.03	-0.02	-0.02
	95% coverage rate									
	Estimated Factor	76.5	70.6	57.1	85.8	85.8	83.0	88.7	90.7	90.9
	Plug-in	82.3	82.3	79.2	86.9	89.4	89.4	89.2	91.8	93.4
	WB	86.3	82.1	73.2	90.7	89.9	88.0	92.3	92.9	93.4
	AR-sieve+CSD	89.6	87.9	84.8	92.1	92.2	91.4	92.8	93.7	94.5

In DGP 5 and 6, both error terms are heteroskedastic. In DGP 5, the idiosyncratic error term contains the cross-sectional dependence. In DGP 6, we impose the dependence in both dimensions for the idiosyncratic error terms. For coverage rates, the results for estimated factors and plug-in are based on asymptotic theory. The bootstrap coverage rates use the bootstrap equal-tailed percentile t method.

two factors is reversed. Previously, the slope coefficient for the aggregated factors was positive, whereas it becomes negative without the COVID period. This suggests that monthly information during the COVID period has a considerable influence on nowcasting the GDP growth rate.

Table 6: Estimation result of long period (1984 Q1 - 2019 Q4)

		$h=2$		$h=1$		$h=0$	
constant		0.87		0.92		0.88	
	Asymptotic	0.70	1.03	0.79	1.06	0.75	1.02
	WB	0.76	1.03	0.84	1.09	0.77	1.02
	AR sieve+CSD	0.79	1.05	0.86	1.11	0.79	1.04
first factor		-1.10		-1.34		-1.27	
	Asymptotic	-1.48	-0.73	-1.67	-1.01	-1.53	-1.00
	WB	-1.52	-0.92	-1.78	-1.20	-1.61	-1.12
	AR sieve+CSD	-1.56	-0.98	-1.83	-1.27	-1.66	-1.16
second factor		0.09		-0.14		-0.01	
	Asymptotic	-0.67	0.84	-0.35	0.07	-0.58	0.56
	WB	-0.13	0.26	-0.40	0.03	-0.23	0.14
	AR sieve+CSD	-0.17	0.24	-0.48	0.02	-0.28	0.13
y_{t-1}		-0.11		-0.19		-0.17	
	Asymptotic	-0.24	0.03	-0.31	-0.06	-0.30	-0.04
	WB	-0.26	0.00	-0.33	-0.10	-0.31	-0.06
	AR sieve+CSD	-0.26	-0.01	-0.35	-0.11	-0.31	-0.06
y_{t-2}		-0.06		-0.09		-0.04	
	Asymptotic	-0.24	0.12	-0.24	0.05	-0.17	0.09
	WB	-0.24	0.08	-0.27	0.03	-0.17	0.08
	AR sieve+CSD	-0.24	0.08	-0.27	0.02	-0.18	0.07
ρ_3		-0.16		-0.14		-0.15	
	Asymptotic	-0.29	-0.02	-0.26	-0.03	-0.26	-0.03
	WB	-0.28	-0.04	-0.26	-0.04	-0.26	-0.04
	AR sieve+CSD	-0.29	-0.04	-0.27	-0.05	-0.26	-0.04

The interval based on the asymptotic theory is obtained by adding and subtracting 1.645 times the heteroskedasticity robust standard errors. The confidence intervals based on bootstrap methods are obtained with equal-tailed bootstrap intervals with a bootstrap number 4999. WB indicates that we use wild bootstrap and AR sieve + CSD indicates that we use the bootstrap algorithm described in Section 3 in the main text.

References

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

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

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

Bi, D., Shang, H. L., Yang, Y., & Zhu, H. (2021). *Ar-sieve bootstrap for high-dimensional time series.* Retrieved from https://arxiv.org/abs/2112.00414

Gonçalves, S., & Perron, B. (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics*, *182*(1), 156–173.