Tuning-free testing of factor regression against factor-augmented sparse alternatives

Jad Beyhum
Department of Economics, KU Leuven, Belgium
and
Jonas Striaukas
Department of Finance, Copenhagen Business School, Denmark

October 5, 2023

Abstract

This study introduces a bootstrap test of the validity of factor regression within a high-dimensional factor-augmented sparse regression model that integrates factor and sparse regression techniques. The test provides a means to assess the suitability of the classical dense factor regression model compared to a sparse plus dense alternative augmenting factor regression with idiosyncratic shocks. Our proposed test does not require tuning parameters, eliminates the need to estimate covariance matrices, and offers simplicity in implementation. The validity of the test is theoretically established under time-series dependence. Through simulation experiments, we demonstrate the favorable finite sample performance of our procedure. Moreover, using the `FRED-MD` dataset, we apply the test and reject the adequacy of the classical factor regression model when the dependent variable is inflation but not when it is industrial production. These findings offer insights into selecting appropriate models for high-dimensional datasets.

Keywords: sparse plus dense, high-dimensional inference, LASSO, factor models

1 Introduction

In this paper, we investigate a factor-augmented sparse regression model. Our analysis involves an observed sample of T real-valued outcomes $y_1, \ldots, y_T$, and high-dimensional regressors $x_1, \ldots, x_T \in \mathbb{R}^p$, which are interconnected as follows:

$$\begin{aligned} y_t &= f_t^\top \gamma^* + u_t^\top \beta^* + \varepsilon_t, \\ x_t &= B f_t + u_t, \quad t = 1 \ldots, T. \end{aligned} \tag{1}$$

Here, $\varepsilon_t \in \mathbb{R}$ represents a random error, u_t is a p-dimensional random vector of idiosyncratic shocks, f_t is a K-dimensional random vector of factors, and B is a $p \times K$ random matrix of loadings. The parameters of interest are $\gamma^* \in \mathbb{R}^K$ and $\beta^* \in \mathbb{R}^p$ and the right-hand side of (1) is unobserved. We consider the case where the number p of regressors is large with respect to the sample size T and sparsity conditions on the high-dimensional parameter vector β^* are imposed. The model formulation in equation (1) effectively merges two popular approaches in handling high-dimensional datasets: factor regression (Stock & Watson (2002), Bai & Ng (2006)) and sparse high-dimensional regression (Tibshirani (1996), Bickel et al. (2009)). Such a model allows the outcome to be related to the regressors through both common and idiosyncratic shocks and may better explain the data than factor regression or sparse regression alone (see Fan, Lou & Yu (2023), which introduces and studies model (1)). Note that, as in Stock & Watson (2002), Bai & Ng (2006), we could augment the model (1) with additional regressors w_t entering the first equation of (1) but not the second one. This case is discussed in the Appendix.

We develop a test for the hypothesis:

$$H_0 : \beta^* = 0 \quad \text{against} \quad H_1 : \beta^* \neq 0 \text{ is sparse,} \tag{2}$$

where our theory outlines the set of sparse alternatives against which our test has power. Our test has two main applications. First, it can be seen as a mean to assess the suitability of the classical factor regression model in comparison to factor-augmented sparse regression alternatives. It provides guidance on the choice between these two models in practical applications. In particular, it allows to infer if forecasting can be improved by using factor-augmented sparse regression instead of only factor regression.[1] Second, our test also sheds

[1] A traditional approach to testing equal forecasting accuracy in nested models is to use mean-squared

light on the data generating process by allowing us to determine if the underlying model is dense (as is the factor regression model) or sparse plus dense (as is the factor-augmented sparse regression model). This determination will then tell us if the relation between the regressors and the outcome is only driven by common shocks (factor regression) or if idiosyncratic shocks play a role as well (factor-augmented sparse regression). The question of the adequacy of sparse or dense representations has recently garnered significant attention (see, e.g., Abadie & Kasy (2019), Giannone et al. (2021)). However, existing studies mostly focus on the differences between sparse and dense models, and do not rely on formal frequentist tests. In contrast, we consider hypothesis testing with a sparse plus dense alternative.

Fan, Lou & Yu (2023) recently introduced the Factor-Adjusted deBiased Test (FabTest) for evaluating (2). However, the FabTest exhibits several limitations. The test relies on a desparsified LASSO estimator based on model (1). To achieve desparsification, Fan, Lou & Yu (2023) utilized the nodewise LASSO method proposed by Zhang & Zhang (2014) and van de Geer et al. (2014) for estimating the precision matrix of the idiosyncratic shocks. However, this approach introduces p additional tuning parameters, in addition to the one used in the original LASSO regression. Although the tuning parameters are selected through cross-validation in practice, Fan, Lou & Yu (2023) did not provide a theoretical justification for this selection procedure. Besides, inferential theory for LASSO-type regressions is not well understood when the tuning parameter is selected by cross-validation. Moreover, the test's performance may deteriorate due to errors associated with the nodewise LASSO estimates, and it incurs a heavy computational cost. Another limitation of the FabTest is its reliance on estimating the variance of ε_t, which can lead to imprecise results where variance estimation is challenging. Additionally, Fan, Lou & Yu (2023) only established the validity of the FabTest for i.i.d. sub-Gaussian data.[2]

In this paper, we propose a new bootstrap test for (2) that overcomes the limitations of the previously mentioned FabTest. Our proposed test does not require tuning parameters

error based bootstrap tests (Clark & McCracken (2001), Clark & West (2007)) Although such tests have been shown to work under certain conditions when the factors are estimated by principal components analysis (Gonçalves et al. (2017)) or common correlated effects (Stauskas & Westerlund (2022)), it is unclear if they are still valid when the LASSO estimator is also used.

[2]See Section 2 in Fan, Lou & Yu (2023).

or the estimation of variance or covariance matrices, making it easy to implement. We establish the validity of the test within a theoretical framework that accommodates scenarios where the number of variables, denoted by p, can significantly exceed T, the explanatory variables exhibit strong mixing and possess exponential tails. In simulations, our procedure shows improvement over the FabTest and demonstrates favorable performance. Furthermore, we apply our test to regression exercises using the `FRED-MD` dataset (McCracken & Ng (2016)). We reject the validity of the classical factor regression model to explain inflation but do not find evidence against the suitability of factor regression when the outcome is industrial production. In the Appendix, we explain how to adapt our test to the case where the model includes additional regressors w_t entering the first equation of (1).

This work is related to a recent literature considering testing for high-dimensional parameters. There exists several approaches, see Fan et al. (2015), Zhu & Bradic (2018), Chernozhukov et al. (2019), Lederer & Vogt (2021), He et al. (2023) and references therein. These procedures differ in terms of the type of alternative hypotheses they consider: sparse, dense, or general. In the present paper, we are interested in sparse alternatives because we want to infer either superior predictive accuracy of the factor-augmented sparse regression or the presence of sparsity. Under dense factor-augmented regression models, the use of the RIDGE estimator is advisable (He (2023)). In a generalization of our model, Fan, Masini & Medeiros (2023) proposes a test for H_0 similar to that of Chernozhukov et al. (2019) against general hypotheses. Tests against general alternatives do not take advantage of sparsity assumptions and, under sparse alternatives, they should therefore exhibit lower power than tests relying on sparsity. This is a reason to rely on methods, such as ours, which use sparsity. Our strategy draws inspiration from Lederer & Vogt (2021), a recent paper that introduces a bootstrap procedure for selecting the penalty parameter of LASSO in standard sparse linear regression. They employ this procedure to test the null hypothesis that a specific high-dimensional parameter is equal to zero. We adapt their approach to the case with unobserved factors, which poses a challenge beyond the scope of the results in Lederer & Vogt (2021). In our case, the unobserved factors need to be estimated, indicating that they act as generated regressors.[3]

[3]Note that, again adapting Lederer & Vogt (2021), we could also devise a procedure to select the penalty parameter of LASSO-type estimators of model (1). We have experimented with such a procedure in Monte

Finally, we would like to note that this paper contributes to various other strands of literature. First, it complements papers that combine factor models and sparse regression (Hansen & Liao (2019), Fan, Lou & Yu (2023), Fan, Masini & Medeiros (2023), Vogt et al. (2022), Beyhum & Striaukas (2023) among others). The proposed strategy allows testing for the joint significance of the coefficients of the idiosyncratic shocks within this framework. Second, our work is related to the literature on inference on parameters of additional low-dimensional regressors in the factor regression model of Stock & Watson (2002), see Bai & Ng (2006), Gonçalves & Perron (2014, 2020). Third, our work connects with the literature on specification tests for models involving factors. Many papers test for the validity of the assumption that loadings are time-independent in the approximate factor model itself — the second equation in (1) — (Breitung & Eickmeier (2011), Chen et al. (2014), Han & Inoue (2015), Yamamoto & Tanaka (2015), Su & Wang (2017, 2020), Baltagi et al. (2021), Xu (2022), Fu et al. (2023)), while Corradi & Swanson (2014) tests for time-independence of all coefficients in the factor regression model of Stock & Watson (2002). Our approach complements this literature by proposing a specification test of the factor regression model under a different alternative, namely the factor-augmented sparse regression model.

Notation. For an integer $N \in \mathbb{N}$, let $[N] = \{1, \dots, N\}$. The transpose of a $n_1 \times n_2$ matrix A is written $A^\top$. Its k^{th} singular value is $\sigma_k(A)$. Let us also define the Euclidean norm $\|A\|_2^2 = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} A_{ij}^2$ and the sup-norm $\|A\|_\infty = \max_{i \in [n_1], j \in [n_2]} |A_{ij}|$. The quantity $n_1 \vee n_2$ is the maximum of n_1 and n_2, $n_1 \wedge n_2$ is the minimum of n_1 and n_2. For $N \in \mathbb{N}$, I_N is the identity matrix of size $N \times N$.

Carlo simulations and did not find that this procedure shows significant improvement over traditionally used cross-validation. For this reason, we decided to focus the present paper on the problem of testing (2), for which simulations yield excellent results.

2 The test

2.1 Testing procedure

In this subsection, we explain our testing procedure, which is then summarized in algorithmic form in subsection 2.2. To facilitate understanding, we rewrite the model in matrix form as follows:

$$Y = F\gamma^* + U\beta^* + \mathcal{E},$$
$$X = BF^\top + U,$$

where $Y = (y_1, \ldots, y_T)^\top$, $F = (f_1, \ldots, f_T)^\top$ is a $T \times K$ matrix, $U = (u_1, \ldots, u_T)^\top$ and $X = (x_1, \ldots, x_T)^\top$ are $T \times p$ matrices and $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_T)^\top$.

It is important to note that, under the null hypothesis H_0, we have $U^\top(Y - F\gamma^*) = U^\top\mathcal{E}$. This observation suggests a testing procedure that involves computing an estimate $2T^{-1}\left\|U^\top(Y - F\gamma^*)\right\|_\infty$ and comparing it with the (estimated) quantiles of $2T^{-1}\left\|U^\top\mathcal{E}\right\|_\infty$.[4]

We can estimate $U^\top(Y - F\gamma^*)$ by principal components analysis. As in Fan, Lou & Yu (2023), we let the columns of $\widehat{F}/\sqrt{T}$ be the eigenvectors corresponding to the leading K eigenvalues of $XX^\top$ and $\widehat{B} = (\widehat{F}^\top\widehat{F})^{-1}\widehat{F}^\top X = T^{-1}\widehat{F}^\top X$. When it is unknown, the number of factors K can be estimated by one of the many methods available in the literature (see for instance Bai & Ng (2002), Onatski (2010), Ahn & Horenstein (2013), Bai & Ng (2019), Fan et al. (2022)). Then, we project the data on the orthogonal of the vector space generated by the estimated factors. Let $\widehat{P} = T^{-1}\widehat{F}\left(\widehat{F}^\top\widehat{F}/T\right)^{-1}\widehat{F}^\top = T^{-1}\widehat{F}\widehat{F}^\top$ be the projector on the vector space generated by the columns of $\widehat{F}$. A natural estimate for U is $\widehat{U} = X - \widehat{F}\widehat{B}^\top = \left(I_T - \widehat{P}\right)X$. Similarly, we let $\widetilde{Y} = \left(I_T - \widehat{P}\right)Y$ be an estimate of $Y - F\gamma^*$. The final estimate of $2T^{-1}\left\|U^\top(Y - F\gamma^*)\right\|_\infty$ is our test statistic

$$2T^{-1}\left\|\widehat{U}^\top\widetilde{Y}\right\|_\infty. \tag{3}$$

Next, to estimate the quantiles of the distribution of $2T^{-1}\left\|U^\top\mathcal{E}\right\|_\infty$, we need an esti-

[4] We have a factor 2 in front of $T^{-1}\left\|U^\top(Y - F\gamma^*)\right\|_\infty$ and $T^{-1}\left\|U^\top\mathcal{E}\right\|_\infty$ because $2T^{-1}\left\|U^\top\mathcal{E}\right\|_\infty$ is the effective noise of the problem, a natural concept in the literature on the LASSO, see Lederer & Vogt (2021).

mate of $\mathcal{E}$. We obtain it through the following LASSO estimator:

$$\widehat{\beta}_\lambda = \arg\min_{\beta\in\mathbb{R}^p} \frac{1}{T}\left\|\widetilde{Y} - \widehat{U}\beta\right\|_2^2 + \lambda\|\beta\|_1, \tag{4}$$

where $\lambda > 0$ is a penalty parameter, the choice of which will be fully data driven in both theory and practice, making our test tuning-free. For $t \in [T]$, we denote by $\widetilde{y}_t$ the t^{th} element of $\widetilde{Y}$ and $\widehat{u}_t$ as the $T \times 1$ vector corresponding to the t^{th} row of $\widehat{U}$. For a given $\lambda > 0$, let $\widehat{\varepsilon}_{\lambda,t} = \widetilde{y}_t - \widehat{u}_t^\top \beta^*,\ t \in [T]$ be the estimate of ε_t. For a fixed $\alpha \in (0,1)$, we can then estimate q_α, the $(1-\alpha)$ quantile of the distribution of $2T^{-1}\|U^\top\mathcal{E}\|_\infty$, by the Gaussian multiplier bootstrap. Let $e = (e_1, \ldots, e_T)$ be a standard normal random vector independent of the data (X, Y) and define the criterion

$$\widehat{Q}(\lambda, e) = \left\|\frac{2}{T}\sum_{t=1}^{T} \widehat{u}_t\widehat{\varepsilon}_{\lambda,t}e_t,\right\|_\infty.$$

The estimate $\widehat{q}_\alpha(\lambda)$ of q_α is then the $(1-\alpha)$-quantile of the distribution of $\widehat{Q}(\lambda, e)$ given X and Y. Formally, $\widehat{q}_\alpha(\lambda) = \inf\left\{q : \mathbb{P}_e(\widehat{Q}(\lambda, e) \leq q) \geq 1-\alpha\right\}$, where $\mathbb{P}_e(\cdot) = \mathbb{P}(\cdot|X, Y)$.

The only remaining element to make the test tuning-free is the procedure to select λ. We adapt the approach of Lederer & Vogt (2021) to our setting. Our choice of λ is

$$\widehat{\lambda}_\alpha = \inf\{\lambda > 0 \ : \ \widehat{q}_\alpha(\lambda') \leq \lambda' \text{ for all } \lambda' \geq \lambda\}. \tag{5}$$

We explain in Section 2.2 how to compute $\widehat{\lambda}_\alpha$ in practice. The infimum in (5) exists because for all $\lambda \geq \bar{\lambda} = 2T^{-1}\left\|\widehat{U}^\top\widetilde{Y}\right\|_\infty$, it holds that $\widehat{\beta}_\lambda = \widehat{\beta}_{\bar{\lambda}} = 0$. Moreover, since $\widehat{U}\widehat{\beta}_\lambda$ is a continuous function of λ, $\widehat{q}_\alpha(\lambda)$ is also continuous in λ and the infimum is attained at a point $\widehat{\lambda}_\alpha > 0$ such that $q_\alpha\left(\widehat{\lambda}_\alpha\right) = \widehat{\lambda}_\alpha$. Let us recall briefly the heuristics behind the choice of λ and refer the reader to Lederer & Vogt (2021) for more details. First, note that when λ is close to q_α, standard convergence bounds for the LASSO suggest that $\widehat{\beta}_\lambda$ is a precise estimate of β^*, so that $\widehat{\varepsilon}_{\lambda,t}$ is a good estimate of ε_t and, in turn, $\widehat{q}_\alpha(\lambda)$ is close to q_α. Second, when λ becomes (much) larger than q_α, the error $\widehat{\varepsilon}_{\lambda,t} - \varepsilon_t$ becomes large and dependent of $\widehat{u}_t$, which in turn increases $\widehat{q}_\alpha(\lambda)$ and leads it to be larger than q_α. We then let our estimator of q_α be $\widehat{\lambda}_\alpha = \widehat{q}_\alpha\left(\widehat{\lambda}_\alpha\right)$.

The test rejects H_0 at the level α when our test statistic given in (3) is larger than the estimate $\widehat{\lambda}_\alpha$ of q_α. Therefore, our testing procedure is free of tuning parameters stemming from the LASSO regression in equation (4).

2.2 Computation

Algorithm 1 below explains how to conduct the test in practice. Let us discuss Step 4 of Algorithm 1 in detail. It approximates $\widehat{\lambda}_\alpha$ as defined in (5). It is advisable to set the grid size M and the number of bootstrap samples L to be as large as possible. As mentioned in Lederer & Vogt (2021), one can speed up Step 4.2 by computing the LASSO with a warm start along the penalty parameter path. Furthermore, Step 4.3 can be accelerated through parallelization techniques. In our implementation, we use both suggestions which greatly speeds up the computations. We also note that to compute the p-value of the test, it suffices to conduct it on a grid of values of α and let the p-value be equal to the largest value of α in this grid such that the test of level α rejects H_0.

---

1. Estimate $\widehat{K}$ by one of the available estimators of the number of factors.

2. Let the columns of $\widehat{F}/\sqrt{T}$ be the eigenvectors corresponding to the leading $\widehat{K}$ eigenvalues of $XX^\top$.

3. Compute $\widehat{U} = \left(I_T - \widehat{P}\right) X$ and $\widetilde{Y} = \left(I_T - \widehat{P}\right) Y$, where $\widehat{P} = T^{-1}\widehat{F}\widehat{F}^\top$.

4. Calculate an approximation $\widehat{\lambda}_{\alpha,emp}$ of $\widehat{\lambda}_\alpha$ as follows:

4a. Specify a grid $0 < \lambda_1 < \cdots < \lambda_M < \bar{\lambda}$, with $\bar{\lambda} = 2T^{-1}\left\|\widehat{U}^\top\widetilde{Y}\right\|_\infty$.

4b. For $m \in [M]$ compute $\left\{\widehat{Q}\left(\lambda_m, e^{(\ell)}\right) :\ \ell \in [L]\right\}$ for L draws of $e \sim \mathcal{N}(0, I_T)$ and the corresponding empirical $(1-\alpha)$-quantile $\widehat{q}_{\alpha,emp}(\lambda_m)$ from them.

4c. Let $\widehat{\lambda}_{\alpha,emp} = \widehat{q}_{\alpha,emp}(\lambda_{\widehat{m}})$, with $\widehat{m} = \min\{m \in [M] :\ \widehat{q}_{\alpha,emp}(\lambda_{m'}) \leq \lambda_{m'}$ for all $m' \geq m\}$.

5. Reject H_0 when $2T^{-1}\left\|\widehat{U}^\top\widetilde{Y}\right\|_\infty > \widehat{\lambda}_{\alpha,emp}$.

Algorithm 1: Conducting a test of level $\alpha \in (0,1)$.

3 Asymptotic theory

In this section, we provide the asymptotic properties of the test in a theoretical framework allowing for time series dependence in the factors and the idiosyncratic shocks and exponential tails. We place ourselves in an asymptotic regime where T goes to infinity and p

goes to infinity as a function of T. The number of factors K is fixed with T. It would be possible to let it grow, see, for instance, Beyhum & Gautier (2023). For our theory, as is standard in the literature, we assume that K is known, so that $\widehat{K} = K$. However, our results would remain valid when one uses an estimator $\widehat{K}$ which is equal to K with probability going to 1 as $T \to \infty$. The distributions of the factors f_t and the error terms ε_t do not depend on T, while the distribution of the other variables are allowed to vary with T. All the constants we introduce are universal in the sense that they do not vary with the sample size. Our assumptions are significantly weaker than that of Fan, Lou & Yu (2023). We impose the usual identifiability condition for factor models (Bai (2003), Fan et al. (2013)):

$$\operatorname{cov}(f_t) = I_K \text{ and } B^\top B \text{ is diagonal.} \tag{6}$$

We introduce further notation. The loading b_{jk} corresponds to the j^{th} element of the k^{th} column of B. Let also $b_j = (b_{j1}, \dots, b_{jk})^\top$. We first state four assumptions similar to the usual ones found in the factor models literature (see e.g. Bai & Ng (2006), Fan et al. (2013)).

Assumption 1 *All the eigenvalues of the $K \times K$ matrix $p^{-1}B^\top B$ are bounded away from 0 and ∞ as $p \to \infty$.*

Assumption 2 *The following holds:*

(i) $\{u_t, f_t, \varepsilon_t, \sum_{\ell=1}^p u_{t\ell} b_\ell\}_t$ *is strictly stationary. Moreover, it holds that*

$$\mathbb{E}[u_{tj}] = \mathbb{E}[f_{tk}] = \mathbb{E}[u_{tj} f_{tk}] = \mathbb{E}\left[f_{tk}\left(\sum_{\ell=1}^p u_{t\ell} b_{\ell h}\right)\right] = 0,$$

for all $t \in [T], j \in [p], k, h \in [K]$.

(ii) Let $\Sigma = \mathbb{E}[u_t u_t^\top]$. *There exist* $\kappa_1, \kappa_2 > 0$ *such that* $\sigma_p(\Sigma) > \kappa_1$, $\sigma_p(\mathbb{E}\left[\varepsilon_t^2 u_t u_t^\top\right]) > \kappa_1$, $\left|\mathbb{E}\left[\varepsilon_t^2 u_t u_t^\top\right]\right|_\infty < \kappa_2$, $\max_{j\in[p]} \sum_{\ell=1}^p |\Sigma_{j\ell}| < \kappa_2$ *and* $\min_{j,\ell\in[p]} \left(\mathbb{E}\left[(u_{tj}u_{t\ell})^2\right] - \mathbb{E}[u_{tj}u_{tk}]^2\right) > \kappa_1$.

(iii) There exist $K_1, \theta_1 > 0$ such that for any $z > 0$, $t \in [T]$, $j \in [p]$ and $k \in [K]$, we have

$$\mathbb{P}\left(|u_{tj}| > z\right) \leq \exp\left(-\left(\frac{z}{K_1}\right)^{\theta_1}\right);$$

$$\mathbb{P}\left(|f_{tk}| > z\right) \leq \exp\left(-\left(\frac{z}{K_1}\right)^{\theta_1}\right);$$

$$\mathbb{P}\left(\frac{1}{\sqrt{p}}\left|\sum_{j=1}^{p} b_{jk} u_{tj}\right| > z\right) \leq \exp\left(-\left(\frac{z}{K_1}\right)^{\theta_1}\right);$$

$$\mathbb{P}\left(|\varepsilon_t| > z\right) \leq \exp\left(-\left(\frac{z}{K_1}\right)^{\theta_1}\right).$$

(iv) $\{u_t \varepsilon_t\}_t$ is uncorrelated across t, and, for all $t \in [T], j \in [p], k \in [K]$,

$$\mathbb{E}[u_{tj}\varepsilon_t] = \mathbb{E}[f_{tk}\varepsilon_t] = \mathbb{E}\left[\varepsilon_t \sum_{\ell=1}^{p} u_{t\ell} b_{\ell k}\right] = 0.$$

Assumption 3 *Let α denote the strong mixing coefficients of $\{f_t, u_t, \varepsilon_t, \sum_{\ell=1}^{p} u_{t\ell} b_\ell\}_t$. There exists $\theta_2 > 0$ such that $2\theta_1^{-1} + \theta_2^{-1} > 1$ and $K_2 > 0$ such that for all $T \in \mathbb{Z}_+$, we have*

$$\alpha(T) \leq \exp\left(-K_2 T^{\theta_2}\right).$$

Assumption 4 *There exists $M > 0$ such that for all $s, t \in [T]$, we have*

(i) $\|B\|_\infty < M$ a.s.;

(ii) $\mathbb{E}\left[p^{-1/2}\left(u_s^\top u_t - \mathbb{E}\left[u_s^\top u_t\right]\right)\right]^4 < M$;

(iii) $\max_{j \in [p], k \in [K]} \left|T^{-1} \sum_{t=1}^{T} \mathbb{E}\left[u_{tj}\left(\sum_{\ell=1}^{p} u_{t\ell} b_{\ell k}\right)\right]\right| < M$.

Assumption 1 combined with the identifiability condition (6) constitutes a strong factor assumption (Bai (2003)). Assumption 2 restricts the moments and the tail behavior of the variables. Assumption 2 (i),(ii),(iv) contains conditions on the moments of the different variables similar to that of the literature (Bai & Ng (2006), Fan et al. (2013)). We assume that the variables in Assumption 2 (iii) have exponential tails with common parameter θ_1. It would be possible to have a different tail parameter for each variable or to allow for polynomial tails, but we avoid doing so to simplify our presentation. In the similar context of inference on factor regression models, Assumption E.2. in Bai & Ng (2006) and Assumption 7 in Gonçalves & Perron (2014) impose conditions analogous to the restriction that

$\{u_t\varepsilon_t\}_t$ is uncorrelated across t in Assumption 2 (iv). A sufficient condition for the latter restriction is that $\{\varepsilon_t\}_t$ is uncorrelated across t and independent of $\{u_t\}_t$. Note that this assumption could be avoided by using a block bootstrap method, but this would complicate the test and may not be justified in a setting where the factors might capture most of the serial correlation. Assumption 3 means that $\{f_t, u_t, \varepsilon_t, \sum_{\ell=1}^{p} u_{t\ell} b_{\ell h}\}_t$ are strongly mixing, which is a restriction on the time-series dependence of the variables. Next, Assumption 4 (i),(ii) is found in Fan et al. (2013) and contains a boundedness condition (i) and a moment condition (ii) on both the time-series and the cross-section dependence of the idiosyncratic shocks. Finally, Assumption 4 (iii) is a restriction on the cross-section correlation of the idiosyncratic shocks. This condition holds, for instance, if $\{u_t\}_t$ and $\{b_\ell\}_\ell$ are independent and $\max_{j\in[p]} \sum_{\ell=1}^{p} |\Sigma_{j\ell}| = O(1)$ (which is imposed in Assumption 2 (iii)).

Let us introduce $\theta^{-1} = 2\theta_1^{-1} + \theta_2^{-1}$, $\tau = 12 + 4\theta_2 + \frac{4}{\theta} + \frac{4}{\theta_2}$ and $\varphi^* = \gamma^* - B^\top \beta^*$. To interpret φ^*, note that the first equation of (1) can be rewritten $y_t = f_t^\top \varphi^* + x_t^\top \beta^* + \varepsilon_t$, which becomes a usual high-dimensional sparse regression model when $\varphi^* = 0$. Note φ^* is random and, therefore, we use probabilistic notation when stating bounds on its norm. The next assumption concerns the relative growth rate of T and p.

Assumption 5 *The following holds:*

(*i*) $\sqrt{\frac{\log(T\vee p)^\tau}{T}}(\|\beta^*\|_1 \vee 1) = o(1)$;

(*ii*) $\log(T \vee p)^{5/2} \frac{\sqrt{T}}{T\wedge p}(\|\varphi^*\|_2 \vee 1) = o_P(1)$.

Assumption 5 (i) contains sparsity restrictions on the alternative hypotheses. When $\|\beta\|_\infty = O(1)$, condition (i) corresponds, up to logarithmic factors, to the standard consistency condition for the LASSO with bounded regressors and error with sub-Gaussian tails that is $\sqrt{\log(p)/T}(s_0 \vee 1) = o(1)$, where s_0 is the number of nonzero coefficients of β^*. Our condition is slightly stronger because of the fact that the factors have to be estimated, the variables have exponential tails and are strongly mixing. Condition (ii) is a slightly more restrictive version of the standard condition that $\sqrt{T}/(T \wedge p) = o(1)$ for inference in the factor regression model.[5] Indeed, since φ^* is of size K, it is reasonable to assume that

[5]This condition is equivalently stated as $\sqrt{T}/p = o(1)$ in Bai & Ng (2006), Corradi & Swanson (2014) and many others.

$\|\varphi^*\|_2 = O_P(1)$. Under this condition, (ii) corresponds to $\sqrt{T}/(T \wedge p) = o(1)$ up to logarithmic factors. Additionally, it is worth noting that our proofs reveal that Assumption 5 is stronger than necessary, and the validity of the test could be established under more complex but weaker rate conditions. However, for the sake of clarity, we present Assumption 5 instead of a more intricate condition.

We have the following theorem.

Theorem 1 *Let Assumptions 1, 2, 3, 4 and 5 hold. For all $\alpha \in (0,1)$, we have*

(*i*) *If $\beta^* = 0$, then* $\mathbb{P}\left(T^{-1}\left\|\widehat{U}^\top \widetilde{Y}\right\|_\infty > \widehat{\lambda}_\alpha\right) \leq \alpha + o(1)$.

(*ii*) *If* $\sqrt{\frac{\log(T\vee p)}{T\wedge p}} = o_P\left(T^{-1}\left\|U^\top U\beta^*\right\|_\infty\right)$, *then* $\mathbb{P}\left(T^{-1}\left\|\widehat{U}^\top \widetilde{Y}\right\|_\infty > \widehat{\lambda}_\alpha\right) \to 1$.

The proof of Theorem 1 can be found in Online Appendix B. Statement (i) means that the empirical size of the test tends to the nominal size. Statement (ii) shows that the test has asymptotic power equal to 1 against sequences of alternatives such that $\sqrt{\frac{\log(T\vee p)}{T\wedge p}} = o_P\left(T^{-1}\left\|U^\top U\beta^*\right\|_\infty\right)$. As noted in Lederer & Vogt (2021), such a condition is inevitable because the presence of the error ε_t prevents us from distinguishing true $U\beta^*$ and ε_t when $U\beta^*$ is too small.

4 Simulations

In this section, we provide a Monte Carlo study which sheds light on the finite sample performance of our proposed testing procedure. We generate samples with $T = 100$ observations, $p = 100$ variables and $K = 2$ factors. The loadings are such that $b_{jk} \sim \mathcal{U}[-1,1], j \in [p], k \in [K]$. The factors are generated as $f_t = \rho_f f_{t-1} + \tilde{f}_t$ for $t = 2,\ldots,T$, where $\tilde{f}_t$ are i.i.d. $\mathcal{N}\left(0, I_K\left(1-\rho_f^2\right)\right)$. The idiosyncratic components $\{u_t\}$ are such that $u_t = \rho_u u_{t-1} + \tilde{u}_t$ for $t = 2,\ldots,T$, where $\tilde{u}_t$ are i.i.d. $\mathcal{N}\left(0, \Sigma\left(1-\rho_u^2\right)\right)$, with $\Sigma_{ij} = 0.6^{|i-j|}, i,j \in [p]$. We also let $\varepsilon_t = \rho_e \varepsilon_{t-1} + \tilde{\varepsilon}_t$ for $t = 2,\ldots,T$, where $\tilde{\varepsilon}_t$ are $\mathcal{N}\left(0, \left(1-\rho_e^2\right)\right)$.

The parameters ρ_f, ρ_u and ρ_e control the level of time series dependence. The stationary distributions of f_t, u_t, ε_t are, respectively, $\mathcal{N}(0, I_K)$, $\mathcal{N}(0,\Sigma)$ and $\mathcal{N}(0,1)$. We initialize f_0, u_0 and ε_0 as such. We consider three dependency designs:

Design 1. $\rho_f = \rho_u = \rho_e = 0$, so that the data are i.i.d. across t.

Design 2. $\rho_f = 0.6$, $\rho_u = 0.1$ and $\rho_e = 0$, which introduces time series dependence in the factors and the idiosyncratic shocks.

Design 3. $\rho_f = 0.6$ and $\rho_u = \rho_e = 0.1$, where there is time series dependence in the factors, the idiosyncratic shocks and the error terms.

The third design is not formally allowed in our theory but we want to show that our test performs well even under weak serial correlation of $\{\varepsilon_t\}_t$.

Finally, we set $\beta^* = (1, 0.5, \dots)^\top \times m$, where $m \in \{0, 0.1, 0.2, 0.3, 0.4\}$ controls signal strength and $\gamma^* = (0.5, 0.5)^\top$.

We compute the rejection probabilities of our test and the FabTest of Fan, Lou & Yu (2023) at the levels $\alpha \in \{0.1, 0.05, 0.01\}$ over 2000 replications. For our test, we set $M = 200$ and choose an equidistant grid of values of λ. We use $L = 200$ bootstrap replications. The results are insensitive to the choice of L and M as long as they are large enough. This is to be expected since their only role is in the approximation of theoretical quantities. In our experience, $L = M = 100$ yields already very precise results. The number of factors K is estimated through the eigenvalue ratio estimator of Ahn & Horenstein (2013). The test of Fan, Lou & Yu (2023) is implemented as in the simulations of Fan, Lou & Yu (2023).

The results are reported in Table 1. In the Online Appendix A, we present simulations under the same data generating processes, but with larger sample size ($T = 200$) and number of variables ($p = 200$). First, we see that both tests have an empirical size close to the nominal levels. For both testing procedures, we see that the empirical size is closer to the nominal levels for the dependent data case compared to the independent data case, but the differences are small. Notably, we see a large increase in the power of our test compared to the FabTest of Fan, Lou & Yu (2023). In both simulation designs, the power of our test increases much faster for larger values of m, suggesting that our procedure correctly rejects the null hypothesis even if the signal is relatively weak, while possessing similar control on the empirical size.

Finally, note that our test has a much lower computational time than the FabTest. For instance, on a Ryzen 9 processor, for Design 1 with $m = 0$ and $T = p = 100$, our test runs in around 2 seconds, while the FabTest takes 36 seconds (average over 100 replications).

Design 1: $\rho_f = \rho_u = \rho_e = 0$						
		Our test			FabTest	
m	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
0	0.0830	0.0390	0.0100	0.0800	0.0400	0.0085
0.1	0.1145	0.0515	0.0180	0.1020	0.0530	0.01350
0.2	0.3025	0.1945	0.0745	0.1515	0.0845	0.0225
0.3	0.6540	0.5375	0.3080	0.3192	0.2086	0.0800
0.4	0.9175	0.8555	0.6905	0.6740	0.5430	0.3245
Design 2: $\rho_f = 0.6$, $\rho_u = 0.1$ and $\rho_e = 0$						
		Our test			FabTest	
m	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
0	0.0875	0.0390	0.0065	0.0905	0.0410	0.0125
0.1	0.1090	0.0480	0.0140	0.1015	0.0460	0.0160
0.2	0.3075	0.2030	0.0750	0.1535	0.0805	0.0220
0.3	0.6570	0.5320	0.3145	0.3305	0.2220	0.0920
0.4	0.9195	0.8595	0.7005	0.6810	0.5580	0.3410
Design 3: $\rho_f = 0.6$ and $\rho_u = \rho_e = 0.1$						
		Our test			FabTest	
m	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
0	0.0935	0.0475	0.0120	0.0850	0.0425	0.0100
0.1	0.1295	0.0595	0.0160	0.1160	0.0530	0.0140
0.2	0.3200	0.2065	0.0800	0.1600	0.0875	0.0215
0.3	0.6645	0.5480	0.3215	0.3302	0.2151	0.0855
0.4	0.9190	0.8665	0.7050	0.6810	0.5555	0.3245

Table 1: Rejection probabilities with $T = 100$ and $p = 100$ for the three designs we consider.

5 Empirical application

We apply our test in two exercises where we use the `FRED-MD` monthly dataset of McCracken & Ng (2016). To avoid the (potential) structural breaks of the great recession and the COVID pandemic, we analyze the data between July 2009 (one month after the end of the NBER recession) and February 2020 (included). The variables are transformed and standardized as suggested in McCracken & Ng (2016). First, we consider inflation series where additionally to our test application, we investigate whether results found using the full sample generalize to out-of-sample forecasting. In the second example, motivated by a recent discussion on sparse versus dense models, see, e.g., Giannone et al. (2021), we test whether the dense factor regression model is an appropriate model for the industrial production variable versus the sparse plus dense alternative (factor-augmented sparse regression model). In this case, we additionally investigate all series that are available in the `FRED-MD` dataset and report the ratios of the number of times we reject the null for each subcategory of this dataset.

Inflation and forecasting. In the first exercise, we want to forecast inflation, denoted CPI_t, at date $t+1$ (the variable `CPIAUCSL` of `FRED-MD`). We focus on inflation as it is a canonical example in forecasting using factor regressions, see, e.g., Stock & Watson (1999). For this, we use all the variables x_t (including the lag of inflation) at date t from the `FRED-MD` dataset as regressors, and thus the regression therefore uses one lag of data. We study the following model

$$\begin{aligned} \mathrm{CPI}_{t+1} &= f_t^\top \gamma^* + u_t^\top \beta^* + \varepsilon_t, \\ x_t &= B f_t + u_t, \quad t = 1\dots, T. \end{aligned} \tag{7}$$

The final sample consists of $T = 127$ observations and $p = 127$ variables. We estimate that there are two factors with the eigenvalue ratio estimator of Ahn & Horenstein (2013). We apply our test, choosing an equidistant grid of $M = 2000$ values of λ and $L = 2000$ bootstrap draws (we use the same grid and values of M and L for our other tests implemented in this section). To compute the p-value, we perform the test for $\alpha \in \{0.001\ell, \ell \in \{0, \dots, 1000\}\}$ and let the p-value be equal to the largest value of α for which we reject H_0. For this exercise, we find a p-value of 0.022 and therefore reject the hypothesis H_0 of adequacy of the classical factor regression model at the 5% level. This suggests that using a factor-augmented sparse regression model could better explain future inflation compared to a factor regression model. Moreover, this indicates that the expected value of inflation given past `FRED-MD` variables may follow a sparse plus dense pattern rather than only a dense representation. We also implemented the FabTest on this data. Following the procedure in Fan, Lou & Yu (2023), i.e., using 2000 bootstrap replications, cross-validation to compute the parameters of the LASSOs regressions and refitted cross-validation based on iterated sure independent screening to estimate the variance of ε_t, the FabTest returns a p-value of 0.784. Hence, in contrast with our approach, the FabTest does not reject H_0 in this exercise.

Sometimes, practitioners include lags of the outcome variable in the regressors on top of the factors (Stock & Watson (1999, 2002), Bai & Ng (2006)). If such lags are significant, this could be the reason why we rejected H_0. To address this concern, we consider the alternative model

$$\begin{aligned} \mathrm{CPI}_{t+1} &= \mathrm{CPI}_t \delta^* + f_t^\top \gamma^* + u_t^\top \beta^* + \varepsilon_t, \\ x_t &= B f_t + u_t, \quad t = 1 \ldots, T, \end{aligned} \tag{8}$$

where this time x_t contains all variables at date t from the `FRED-MD` dataset except CPI_t. We apply the test for $H_0 : \beta^* = 0$, for the case with additional regressors as discussed in the Appendix. We find a p-value of 0.023, so that we can reject H_0 in this case as well.

Based on our full-sample testing results, we proceed to evaluate whether results generalize to out-of-sample predictions. In this case, we compare out-of-sample prediction accuracy of two models, namely, factor model versus the factor-augmented sparse alternative. Our objective in this exercise is to assess whether the inclusion of a high-dimensional sparse idiosyncratic component in the factor model results in improved or diminished forecasting accuracy. To conduct this assessment, we adopt the lag-augmented model specification outlined in equation (8). We use the first half of the sample, $t \in [1, \lfloor T/2 \rfloor - 1]$, to estimate our model and generate the first out-of-sample forecast. Employing an expanding window approach, we continuously add new observations into the estimation set, allowing us to produce updated forecasts. Once we cover the entire out-of-sample period, $t \in [\lfloor T/2 \rfloor, T]$, we calculate the out-of-sample forecast errors. We employ 5-fold cross-validation to select the tuning parameter in the LASSO regression.

We find that the out-of-sample mean squared error (MSE) ratio of the factor-augmented sparse model relative to the factor regression model is 0.812. This indicates that the inclusion of the high-dimensional idiosyncratic shocks leads to more accurate predictions out-of-sample when compared to the classical factor regression model. This seems to be in line with our full-sample testing results where using our test, we indeed find that the sparse component is significant at a 5% significance level.

Testing dense versus sparse plus dense alternative. Let us now turn to industrial production, denoted IP_t (the variable `INDPRO` of `FRED-MD`). We implement the same first regression exercise as for inflation, i.e., in the case where the lag of inflation is included in x_t just replacing inflation by industrial production (see equation (7)). We study the following model

$$\begin{aligned} \mathrm{IP}_{t+1} &= f_t^\top \gamma^* + u_t^\top \beta^* + \varepsilon_t, \\ x_t &= Bf_t + u_t, \quad t = 1 \ldots, T, \end{aligned} \tag{9}$$

and test the same hypothesis, i.e., $H_0 : \beta^* = 0$. The p-value of our test is equal to 0.121 and that of the FabTest is 0.880 (both are computed exactly as in the inflation exercise). Therefore, using both tests, we do not reject H_0. This indicates that the factor regression model is adequate to explain industrial production and there is no need to introduce a sparse component in the model. It also suggests that the data generating process is dense. Interestingly, this result confirms the findings of Giannone et al. (2021), who, using a Bayesian approach, also found that a dense representation was more suitable in a similar regression exercise of industrial production. Our strategy relies on a formal frequentist test and considers a different alternative. It is, therefore, complementary to that of Giannone et al. (2021).

Motivated by contrasting results for inflation and industrial production variables, we analyze each series in the `FRED-MD` dataset. That is, for each series we test β using two model specifications we considered in earlier exercises, namely using only factors (equations (7), (9)) and adding a lag of the outcome variable to the factors (equation (8)). We report the rejection ratios at a 5% significance level for each subcategory of this dataset (see McCracken & Ng (2016)). Results appear in Table 2. First, results show that, in most categories, we reject the null hypothesis in less than 50% of cases, whether we rely solely on factors or include a lag of the outcome variable. The exception is the *Prices* category, where we reject the null hypothesis for 65% of the series. Notably, the addition of a lag tends to reduce rejection ratios across all categories, except for *Housing*. This outcome is expected since the significance of the lag may contribute to our rejection of the null hypothesis. Interestingly, even after accounting for the lag, the *Prices* category continues to exhibit a relatively high percentage of series for which we reject the null hypothesis — specifically, 35%.

	Factors	Factors+Lag
Consumption, Orders, and Inventories	0.400	0.100
Housing	0.000	0.100
Interest and Exchange Rates	0.455	0.091
Labor Market	0.452	0.032
Money and Credit	0.308	0.077
Output and Income	0.250	0.000
Prices	0.650	0.350
Stock Market	0.200	0.000

Table 2: Rejection ratios for each subcategory using factors (column *Factor*) and factors with a lag of the outcome variable (column *Factors+Lag*).

6 Conclusion

This paper proposes a new tuning-free test for the adequacy of the factor regression model against factor-augmented sparse alternatives. We establish the asymptotic validity of our test under time series dependence. In a Monte Carlo study, we show that our procedure has excellent finite sample properties. An empirical application illustrates the usefulness of our method by testing the adequacy of factor regression against factor-augmented sparse alternatives using the well-established `FRED-MD` dataset.

In the first case, our test rejects the null hypothesis of no sparse idiosyncratic shocks component in a regression model to forecast inflation. This result remains robust even when we include the lag of inflation as a regressor. These findings are in line with out-of-sample prediction results, where factor-augmented sparse regression reduces the MSE by 19% compared to the traditional factor regression model.

In the second case, we examine whether the combination of sparse and dense models is suitable for industrial production using our testing approach. In this instance, we do not find evidence to reject the null hypothesis. Given the contrasting results observed for inflation and industrial production, we extend our analysis to encompass all series in the dataset. We calculate rejection ratios for each subcategory of variables. We reject the null the for many series. However, our results suggest that, for the majority of series, the dense model is the appropriate choice, except for the *Prices* category. In this particular case, we reject the null hypothesis 35% of the time, even when incorporating the lag of the dependent variable as a regressor.

One possible limitation of this paper is that we modeled the dense component by a factor model. We leave other approaches to model dense components to future research.

Appendix: testing with additional regressors

A Alternative model

As in Stock & Watson (2002), Bai & Ng (2006), we augment the model with additional low-dimensional regressors $w_1, \ldots, w_t \in \mathbb{R}^\ell$ (where ℓ is fixed with T). We consider the alternative model.

$$\begin{aligned} y_t &= f_t^\top \gamma^* + w_t^\top \delta^* + u_t^\top \beta^* + \varepsilon_t, \\ x_t &= B f_t + u_t, \quad t = 1 \ldots, T, \end{aligned} \tag{10}$$

Here, again, $\varepsilon_t \in \mathbb{R}$ represents a random error, u_t is a p-dimensional random vector of idiosyncratic shocks, f_t is a K-dimensional random vector of factors, and B is a $p \times K$ random matrix of loadings. The parameters are $\gamma^* \in \mathbb{R}^K$, $\delta^* \in \mathbb{R}^\ell$, $\beta^* \in \mathbb{R}^p$. Note that here w_t plays the role of an observed factor (with loading equal to 0). This will be key to understanding the alternative testing procedure of Section B.

We focus on testing

$$H_0 : \beta^* = 0 \quad \text{against} \quad H_1 : \beta^* \neq 0. \tag{11}$$

To facilitate understanding, we again rewrite the model in matrix form as follows:

$$\begin{aligned} Y &= F^\top \gamma^* + W \delta^* + U^\top \beta^* + \mathcal{E}, \\ X &= BF + U, \end{aligned}$$

where $Y = (y_1, \ldots, y_T)^\top$, $F = (f_1, \ldots, f_T)^\top$ is a $T \times K$ matrix, $U = (u_1, \ldots, u_T)^\top$, $W = (w_1, \ldots, w_T)^\top$ and $X = (x_1, \ldots, x_T)^\top$ are $T \times p$ matrices and $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_T)^\top$.

B Alternative testing procedure

Algorithm 2 present the test in this alternative model. It is similar to Algorithm 1. The only difference is that $\widehat{P}$ is now the projector on the columns of the $T \times (\widehat{K} + \ell)$ matrix $(\widehat{F}\ W)$ in Step 3. Essentially, w_t is treated as an observed factor.

---

1. Estimate $\widehat{K}$ by one of the available estimators of the number of factors.

2. Let the columns of $\widehat{F}/\sqrt{T}$ be the eigenvectors corresponding to the leading $\widehat{K}$ eigenvalues of $XX^\top$.

3. Compute $\widehat{U} = \left(I_T - \widehat{P}\right) X$ and $\widetilde{Y} = \left(I_T - \widehat{P}\right) Y$, where $\widehat{P}$ is the projector on the columns of the $T \times \left(\widehat{K} + \ell\right)$ matrix $\left(\widehat{F}\ W\right)$.

4. Calculate an approximation $\widehat{\lambda}_{\alpha,emp}$ of $\widehat{\lambda}_\alpha$ as follows:

4a. Specify a grid $0 < \lambda_1 < \cdots < \lambda_M < \bar{\lambda}$, with $\bar{\lambda} = 2T^{-1} \left\|\widehat{U}^\top \widetilde{Y}\right\|_\infty$.

4b. For $m \in [M]$ compute $\left\{\widehat{Q}\left(\lambda_m, e^{(\ell)}\right) : \ell \in [L]\right\}$ for L draws of $e \sim \mathcal{N}(0, I_T)$ and the corresponding empirical $(1-\alpha)$-quantile $\widehat{q}_{\alpha,emp}(\lambda_m)$ from them.

4c. Let $\widehat{\lambda}_{\alpha,emp} = \widehat{q}_{\alpha,emp}(\lambda_{\widehat{m}})$, with $\widehat{m} = \min\{m \in [M] : \widehat{q}_{\alpha,emp}(\lambda_{m'}) \leq \lambda_{m'} \text{ for all } m' \geq m\}$.

5. Reject H_0 when $2T^{-1} \left\|\widehat{U}^\top \widetilde{Y}\right\|_\infty > \widehat{\lambda}_{\alpha,emp}$.

Algorithm 2: Conducting a test of level $\alpha \in (0, 1)$ with additional regressors.

Supplementary material

Online Appendix: Additional simulation results and the proof of Theorem 1 (.pdf file).

R package: an R package implementing our test is available at `https://github.com/jstriaukas/bootml`.

Acknowledgments

The authors thank Michael Vogt for sharing his code for the procedure of Lederer & Vogt (2021). Jad Beyhum undertook most of this work while employed by CREST, ENSAI

(Rennes). Jad Beyhum gratefully acknowledges financial support from the Research Fund KU Leuven through the grant STG/23/014. Jonas Striaukas gratefully acknowledges the financial support from the European Commission, MSCA-2022-PF Individual Fellowship, Project 101103508. Project Acronym: MACROML.

References

Abadie, A. & Kasy, M. (2019), 'Choosing among regularized estimators in empirical economics: The risk of machine learning', *Review of Economics and Statistics* **101**(5), 743–762.

Ahn, S. C. & Horenstein, A. R. (2013), 'Eigenvalue ratio test for the number of factors', *Econometrica* **81**(3), 1203–1227.

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. (2019), 'Rank regularized estimation of approximate factor models', *Journal of Econometrics* **212**(1), 78–96.

Baltagi, B. H., Kao, C. & Wang, F. (2021), 'Estimating and testing high dimensional factor models with multiple structural changes', *Journal of Econometrics* **220**(2), 349–365.

Beyhum, J. & Gautier, E. (2023), 'Factor and factor loading augmented estimators for panel regression with possibly nonstrong factors', *Journal of Business & Economic Statistics* **41**(1), 270–281.

Beyhum, J. & Striaukas, J. (2023), 'Factor-augmented sparse MIDAS regression for nowcasting', *arXiv preprint arXiv:2306.13362* .

Bickel, P. J., Ritov, Y. & Tsybakov, A. B. (2009), 'Simultaneous analysis of lasso and dantzig selector', *Annals of Statistics* **37**(4), 1705 – 1732.

Breitung, J. & Eickmeier, S. (2011), 'Testing for structural breaks in dynamic factor models', *Journal of Econometrics* **163**(1), 71–84.

Chen, L., Dolado, J. J. & Gonzalo, J. (2014), 'Detecting big structural breaks in large factor models', *Journal of Econometrics* **180**(1), 30–48.

Chernozhukov, V., Chetverikov, D. & Kato, K. (2019), 'Inference on causal and structural parameters using many moment inequalities', *Review of Economic Studies* **86**(5), 1867–1900.

Clark, T. E. & McCracken, M. W. (2001), 'Tests of equal forecast accuracy and encompassing for nested models', *Journal of econometrics* **105**(1), 85–110.

Clark, T. E. & West, K. D. (2007), 'Approximately normal tests for equal predictive accuracy in nested models', *Journal of econometrics* **138**(1), 291–311.

Corradi, V. & Swanson, N. R. (2014), 'Testing for structural stability of factor augmented forecasting models', *Journal of Econometrics* **182**(1), 100–118.

Fan, J., Guo, J. & Zheng, S. (2022), 'Estimating number of factors by adjusted eigenvalues thresholding', *Journal of the American Statistical Association* **117**(538), 852–861.

Fan, J., Liao, Y. & Mincheva, M. (2013), 'Large covariance estimation by thresholding principal orthogonal complements', *Journal of the Royal Statistical Society. Series B, Statistical methodology* **75**(4).

Fan, J., Liao, Y. & Yao, J. (2015), 'Power enhancement in high-dimensional cross-sectional tests', *Econometrica* **83**(4), 1497–1541.

Fan, J., Lou, Z. & Yu, M. (2023), 'Are latent factor regression and sparse regression adequate?', *Journal of the American Statistical Association* (just-accepted), 1–77.

Fan, J., Masini, R. & Medeiros, M. C. (2023), Bridging factor and sparse models. *Annals of Statistics* (forthcoming).

Fu, Z., Hong, Y. & Wang, X. (2023), 'Testing for structural changes in large dimensional factor models via discrete fourier transform', *Journal of Econometrics* **233**(1), 302–331.

Giannone, D., Lenza, M. & Primiceri, G. E. (2021), 'Economic predictions with big data: The illusion of sparsity', *Econometrica* **89**(5), 2409–2437.

Gonçalves, S., McCracken, M. W. & Perron, B. (2017), 'Tests of equal accuracy for nested models with estimated factors', *Journal of Econometrics* **198**(2), 231–252.

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.

Han, X. & Inoue, A. (2015), 'Tests for parameter instability in dynamic factor models', *Econometric Theory* **31**(5), 1117–1152.

Hansen, C. & Liao, Y. (2019), 'The factor-lasso and k-step bootstrap approach for inference in high-dimensional economic applications', *Econometric Theory* **35**(3), 465–509.

He, Y. (2023), 'Ridge regression under dense factor augmented models', *Journal of the American Statistical Association* pp. 1–13.

He, Y., Jaidee, S. & Gao, J. (2023), 'Most powerful test against a sequence of high dimensional local alternatives', *Journal of Econometrics* **234**(1), 151–177.

Lederer, J. & Vogt, M. (2021), 'Estimating the Lasso's effective noise.', *Journal of Machine Learning Research* **22**, 276–1.

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

Onatski, A. (2010), 'Determining the number of factors from empirical distribution of eigenvalues', *Review of Economics and Statistics* **92**(4), 1004–1016.

Stauskas, O. & Westerlund, J. (2022), 'Tests of equal forecasting accuracy for nested models with estimated cce factors', *Journal of Business & Economic Statistics* **40**(4), 1745–1758.

Stock, J. H. & Watson, M. W. (1999), 'Forecasting inflation', *Journal of Monetary Economics* **44**(2), 293–335.

Stock, J. H. & Watson, M. W. (2002), 'Forecasting using principal components from a large number of predictors', *Journal of the American Statistical Association* **97**(460), 1167–1179.

Su, L. & Wang, X. (2017), 'On time-varying factor models: Estimation and testing', *Journal of Econometrics* **198**(1), 84–101.

Su, L. & Wang, X. (2020), 'Testing for structural changes in factor models via a nonparametric regression', *Econometric Theory* **36**(6), 1127–1158.

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**(1), 267–288.

van de Geer, S., Bühlmann, P., Ritov, Y. & Dezeure, R. (2014), 'On asymptotically optimal confidence regions and tests for high-dimensional models', *Annals of Statistics* **42**(3), 1166 – 1202.

Vogt, M., Walsh, C. & Linton, O. (2022), 'CCE estimation of high-dimensional panel data models with interactive fixed effects', *arXiv preprint arXiv:2206.12152* .

Xu, W. (2022), 'Testing for time-varying factor loadings in high-dimensional factor models', *Econometric Reviews* **41**(8), 918–965.

Yamamoto, Y. & Tanaka, S. (2015), 'Testing for factor loading structural change under common breaks', *Journal of Econometrics* **189**(1), 187–206.

Zhang, C.-H. & Zhang, S. S. (2014), 'Confidence intervals for low dimensional parameters in high dimensional linear models', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **76**(1), 217–242.

Zhu, Y. & Bradic, J. (2018), 'Linear hypothesis testing in dense high-dimensional linear models', *Journal of the American Statistical Association* **113**(524), 1583–1600.