3 | TESTING PROCEDURES

This section describes the testing procedures used in the paper, which are divided in three stages. First, we describe the computation of the factor loadings. Second, we describe the estimator for the parameters of interest in the construction of the test statistics. We then formalize the null hypothesis, the test statistics, and the corresponding limiting distributions of the tests.

3.1 | First stage: Realized covariance measures

Although nonstandard, the assumption that β_{it} is dynamic is consistent with recent literature in financial economics. This literature suggests that factor loadings may vary with conditioning variables; see Hansen and Richard (1987), Jagannathan and Wang (1996), and Wang (2003). The first step is the construction of the estimates of the factor loadings β_{it}. This has been a difficult task in the empirical asset pricing literature. Our paper overcomes this problem by using data at higher frequencies to obtain accurate measures of the factor loadings at lower frequencies. We follow Andersen et al. (2006) and explicitly allow for continuous evolution of β_{it} over time. The theoretical background for this assumption is the theory of quadratic variation and covariation. The following paragraphs provide an introduction to this theory and define the realized covariance measure that we will use to estimate the dynamics of β_{it}.

Let Δ denote the sampling frequency and $m = 1/\Delta$ be the number of sample observations per period t. We denote the intra-period continuously compounded returns from time $t + h\Delta$ to $t + (h + 1)\Delta$ by $R_{t+(h+1)\Delta} = p_{t+(h+1)\Delta} - p_{t+h\Delta}$ with $h = 0, \dots, m-1$ and $m\Delta = 1$. The corresponding inter-period return is defined as $R_{t+1} = \sum_{h=0}^{m-1} R_{t+(h+1)\Delta}$. For our purposes, it is helpful to consider this return process as a multidimensional process formed by the excess returns r^e_{t+1} on the N individual risky assets and K-factor returns f_t. We can do this under the assumption that the common pricing factors f_t are returns on traded assets. Let $R_{t+1} = \left(r^e_{1,t+1}, \dots, r^e_{N,t+1}, f_{1,t+1}, \dots, f_{K,t+1}\right)$; under these conditions, the $(N + K) \times (N + K)$ realized covariance matrix at time $t + 1$ is

$$\hat{\Omega}_{t+1} = \sum_{h=0}^{m-1} R_{t+(h+1)\Delta} R^{\top}_{t+(h+1)\Delta}. \tag{15}$$

This matrix is positive definite provided $N + K < m$. Moreover, this covariance measure can be defined over l periods as

$$\hat{\Omega}^{(l)}_{t+l} = \sum_{j=1}^{l} \hat{\Omega}_{t+j}.$$

The assumption of no-arbitrage in financial markets also entails a logarithmic $(N + K) \times 1$ price process, p_τ, with $\tau \in [0, T]$—that is, in the class of semimartingales. Then, it has the representation

$$p_\tau = p_0 + A_\tau + M_\tau, \tag{16}$$

where A_τ is a predictable drift component of finite variation, and M_τ is a local martingale, such that $A_0 = 0$ and $M_0 = 0$. A typical example of a process within this class is a multivariate continuous-time stochastic volatility diffusion process:

$$dp_\tau = \mu_\tau d\tau + \Omega_\tau dW_\tau, \tag{17}$$

where W_τ denotes a standard $N + K$-dimensional Brownian motion, and both the process for the $(N + K) \times (N + K)$ positive definite diffusion matrix, Ω_τ, and the $(N + K)$-dimensional instantaneous drift, μ_τ, are strictly stationary and jointly independent of the W_τ process. The cumulative return process at time $t + 1$ associated with Equation 17 is $R_{t+1} = p_{t+1} - p_t$. Then, for any partition Π_m of the interval $[t, t + 1]$ defined as $\Pi_m = \{t = \tau_0 < \tau_1 < \dots < \tau_m = t + 1\}$ the quadratic covariation (QC) of the return process from time t to $t + 1$ is defined as

$$\mathrm{QC}_{t+1} = \mathrm{plim}_{||\Pi_m||\to 0} \sum_{h=0}^{M-1} R_{\tau_{h+1}} R^{\top}_{\tau_{h+1}} \quad \text{as} \quad m \to \infty, \tag{18}$$

with plim denoting the limit in probability, and $||\Pi_m|| = \max_{h=0,\dots,m-1}(\tau_{h+1} - \tau_h)$. This process measures the realized sample path variation of the squared return process. For the process in Equation 17, it follows that

$$\mathrm{QC}_{t+1} = \int_t^{t+1} \Omega_\tau d\tau,$$

with $\int_t^{t+1} \Omega_\tau d\tau$ the integrated diffusion matrix at time $t+1$.

The main result derived and extended in Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2001, 2003), and Barndorff-Nielsen and Shephard (2004) is that the realized covariance matrix converges in probability to the quadratic variation measure as the sampling frequency m increases. Mathematically,

$$\widehat{\Omega}_{t+1} \xrightarrow{p} \mathrm{QC}_{t+1}, \quad \text{as} \quad m \to \infty. \tag{19}$$

It is important to note that the process QC_{t+1} is different from the conditional return covariance matrix $\mathrm{cov}_t[R_{t+1}, R_{t+1}]$. Nevertheless, Andersen, Bollerslev, Diebold, and Labys (2001) show that if the price process is square integrable and satisfies some further regularity conditions on the predictable drift component A_τ in Equation 16, it follows that $\mathrm{cov}_t[R_{t+1}, R_{t+1}] = \mathrm{E}_t\left[\mathrm{QC}_{t+1}\right]$. In our asset pricing setting outlined above, we are interested in the submatrix of $\mathrm{cov}_t[R_{t+1}, R_{t+1}]$ determined by the upper-right block matrix of dimension $N \times K$. The elements of this matrix correspond to the dynamic quantities $\mathrm{cov}_t\left[r^e_{i,t+1}, f_{j,t+1}\right]$ with $i = 1, \dots, N$ and $j = 1, \dots, K$, which are identified above as the dynamic factor loadings β_{it}. Hence, we have the identity $\beta_{it} = \mathrm{E}_t\left[\mathrm{QC}^{ur}_{i,t+1}\right]$, with $\mathrm{QC}^{ur}_{i,t+1}$ denoting the upper-right block matrix of QC_{t+1} with the quantities describing the quadratic variation between excess returns and common factors. These vectors can be consistently estimated by the vector $\widehat{\Omega}^{ur}_{i,t+1} = \left(\omega_{(i,N+1),t+1}, \dots, \omega_{(i,N+K),t+1}\right)$ obtained from realized estimators computed with data at higher frequencies. Most importantly, note that for a grid sufficiently dense ($m \to \infty$) Equation 19 shows that we can assume for each time t that

$$\beta_{it} = \mathrm{E}_t\left[\widehat{\Omega}^{ur}_{i,t+1}\right]. \tag{20}$$

This quantity is unobserved so we need to approximate it by a forecast of β_{it} obtained from a time series model. For simplicity, we propose the following autoregressive process for each element $\omega_{(i,j),t+1}$ of the vector $\widehat{\Omega}^{ur}_{i,t+1}$ with $i = 1, \dots, N$ and $j = N+1, \dots, N+K$:[1]

$$\omega_{(i,j),t+1} = \delta_{ij,0} + \delta_{ij,1}\omega_{(i,j),t} + \nu_{ij,t+1}, \tag{21}$$

with $\omega_{(i,j),t+1}$ converging to the quadratic covariation between $r^e_{i,t+1}$ and each of the pricing factors $f_{j,t+1}$ for $m \to \infty$. Furthermore, the autoregressive process implies that $\mathrm{E}_t\left[\omega_{(i,j),t+1}\right] = \delta_{ij,0} + \delta_{ij,1}\omega_{(i,j),t}$. By construction, see Equation 20, $\mathrm{E}_t\left[\omega_{(i,j),t+1}\right] = \beta_{ij,t}$ under the assumption of Equation 21, provides a correct specification of the dynamics of each element of the vector $\widehat{\Omega}^{ur}_{i,t+1}$. In this context, a consistent estimator of the time series $\beta_{ij,t}$ is

$$\widehat{\beta}_{ij,t} = \widehat{\mathrm{E}}_t[\omega_{(i,j),t+1}] = \widehat{\delta}_{ij,0} + \widehat{\delta}_{ij,1}\omega_{(i,j),t}, \tag{22}$$

with $(\widehat{\delta}_{ij,0}, \widehat{\delta}_{ij,1})$ the ordinary least squares parameter estimators of the time series regression Equation (21).

3.2 | Second stage: Estimation of factor risk premia

The observable estimated factor loadings obtained from Equation 21 allow us to obtain estimates of the factor risk premia from time series regressions between the excess returns on the risky assets and the estimated dynamic factor loadings. For ease of presentation, we reproduce the notation for the relevant quantities. Let $\eta_i = (\alpha_i, \lambda_i^\top)^\top$ denote the model parameters for $i = 1, \dots, N$, $\widehat{X}_{it} = \left(1, \widehat{\beta}_{it}^\top\right)$ with $\widehat{\beta}_{it}$ defined in Equation 22, and $X_{it} = \left(1, \beta_{it}^\top\right)$ with $\beta_{it} = \mathrm{E}_t\left[\mathrm{QC}^{ur}_{i,t+1}\right]$.

We propose the following time series regression equation that mimics Equation 10 for testing the above hypotheses:

$$r^e_{i,t+1} = \widehat{X}_{it}\eta_i + \nu_{i,t+1}, \tag{23}$$

[1]The autoregressive strategy in Equation 21 can be improved by constructing estimators of the conditional covariance matrix $\mathrm{cov}_t[R_{t+1}, R_{t+1}]$ that preserve the symmetric positive definiteness property of covariance matrices; see Noureldin, Shephard, and Sheppard (2012), Golosnoy, Gribisch, and Liesenfeld (2012), and Jin and Maheu (2013). To do this, one has to assume that $\widehat{\Omega}_{t+1}$ conditional on the information available up to time t follows a $N+K$-dimensional central Wishart distribution $W(\nu, S_t/\nu)$, where $\nu > N+K-1$ denotes the degrees of freedom and S_t/ν is a positive definite symmetric scale matrix of order $N \times K$. This assumption defines a conditional autoregressive Wishart model such that $\mathrm{E}_t[\widehat{\Omega}_{t+1}] = S_t$.

with

$$v_{i,t+1} = \beta_i^{*\top} \widetilde{g}_{t+1} + \varepsilon_{i,t+1}$$

being the pricing error of the asset pricing equation. The regression model (Equation 23) exhibits cross-correlation with the rest of the asset pricing equations indexed by $i = 1, \ldots, N$ due to the presence of observed and unobserved common factors in the pricing errors $v_{i,t+1}$. We assume R factors that will be treated as unobserved components and that will determine a factor model. The presence of these factors generates cross-sectional dependence in the test statistic. The error term $\varepsilon_{i,t+1}$ satisfies Assumption C in Ando and Bai (2015), namely $\mathrm{E}[\varepsilon_{it}] = 0$, $\mathrm{E}[|\varepsilon_{it}|^8] < C < \infty$ for all i and t, ε_{it} and ε_{js} are independent for $i \neq j$ and $t \neq s$. The error term is also independent of the regressors $\widehat{X}_{it}$, parameters η_i and factors $\widetilde{g}_t$.

The quantities η_i, β_i^* and $\widetilde{g}_{t+1}$, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$ are estimated from the following minimization problem:

$$l\left(\eta_i, \beta_i^*, \widetilde{g}_{t+1}\right) = \sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{i,t+1}^e - \widehat{X}_{it}\eta_i - \beta_i^{*\top}\widetilde{g}_{t+1}\right)^2, \tag{24}$$

subject to the normalization $\frac{G^\top G}{T} = I_R$ and $\frac{\beta^{*\top}\beta^*}{N}$ being diagonal, with $G = (\widetilde{g}_1, \ldots, \widetilde{g}_T)^\top$ a $T \times R$ matrix and $\beta^* = (\beta_1^*, \ldots, \beta_N^*)^\top$ a $N \times R$ matrix. For convenience of implementation, we adopt the iterative principal component approach initially proposed by Bai (2009) and extended by Song (2013). This approach decomposes the original estimation problem into two steps: estimation of the individual coefficients given common factors, and estimation of the common factors given individual coefficients. Following these authors, we maintain the assumption that the number of factors—that is, R—is known. The extension to an unknown number of factors under heterogeneous regression coefficients is cumbersome (see Song, 2013) and beyond the scope of this paper.

Bai (2009) and Song (2013) propose a tractable solution to the estimation problem by concentrating out the factor loadings from the objective function (Equation 24). More specifically, these authors assume that the factor loadings β_i^* satisfy a relationship of the form $\beta^* = \left(G^\top G\right)^{-1} G^\top \left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)$, with r_i^e the $T \times 1$ vector representation of the excess returns in Equation 23. Then, replacing this expression in Equation 24, the minimization problem is equivalent to

$$\max_G \ \mathrm{tr}\left[G^\top \left(\frac{1}{NT}\sum_{i=1}^{N}\left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)\left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)^\top\right) G\right], \tag{25}$$

subject to $\frac{G^\top G}{T} = I_R$, with tr denoting the trace of the matrix. Therefore, the estimator $(\{\widehat{\eta}_i\}_{i=1}^N, \hat{G})$ with $\widehat{\eta}_i = (\widehat{\alpha}_i, \widehat{\lambda}_i^\top)^\top$ should simultaneously solve a system of nonlinear equations:

$$\widehat{\eta}_i = \left(\widehat{X}_i^\top M_{\hat{G}} \widehat{X}_i\right)^{-1} \widehat{X}_i^\top M_{\hat{G}} r_i^e, \tag{26}$$

and

$$\left[\frac{1}{NT}\sum_{i=1}^{N}\left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)\left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)^\top\right]\hat{G} = \hat{G}\widehat{V}_{NT}, \tag{27}$$

where $\widehat{V}_{NT}$ is a diagonal matrix of R largest eigenvalues corresponding to $\hat{G}$, and $M_{\hat{G}} = I_R - \hat{G}(\hat{G}^\top\hat{G})^{-1}\hat{G}^\top$. The actual estimation procedure can be implemented by iterating each of the two steps in Equations 26 and 27 until convergence. The unknown factor loadings are obtained as $\widehat{\beta}^* = \left(\hat{G}^\top\hat{G}\right)^{-1}\hat{G}^\top\left(r_i^e - \widehat{X}_i\widehat{\eta}_i\right)$.

The presence of generated regressors $\widehat{X}_{it}$ in Equation 23 replacing X_{it} implies that the variance of the vector of parameter estimators $\widehat{\eta}_{\cdot} = (\widehat{\eta}_1, \ldots, \widehat{\eta}_N)^\top$ needs to be corrected. In what follows, we establish several asymptotic results when m, N, and T go to infinity sequentially. The sequential asymptotics are defined as m diverging to infinity first, and then N, T.[2] For a detailed discussion on sequential and simultaneous asymptotics for panel data, see Phillips and Moon (1999, 2000). In particular, we adopt the following notation: $(m, N, T)_{\mathrm{seq}} \to \infty$ means that first $m \to \infty$ and then $(N, T) \to \infty$.

We note that the generated regressors in this paper have two layers of potential estimation errors. First, as Equation 15 shows, we compute the realized covariance matrix with high-frequency data. Second, as shown in Equation 21, we employ

[2]In the definition of the simultaneous asymptotics, m, N, and T tend to infinity at the same time.

an autoregressive framework to model the correlation in the elements of that variance. Therefore, the sequential asymptotics help with the first issue, since by assuming that m diverges first, we obtain consistent estimates for the realized covariance, which are used in the autoregressive fitting, and only have to account for the estimation error of the autoregressive model. We also note that the sequential asymptotics assumption is plausible in our application since we use daily data to estimate quarterly observations.[3]

First, we derive the variance of the statistic $\sqrt{T}(\hat{\eta}_{\cdot} - \eta_{\cdot})$, with $\eta_{\cdot} = (\eta_1, \dots, \eta_N)^\top$ the true parameters of the model, as this quantity plays a fundamental role in deriving the asset pricing tests. Let S be an $N(K+1) \times N(K+1)$ block-diagonal matrix with elements $S_{ii} = \left[X_i^\top M_G X_i\right]/T$, and let L be a $N(K+1) \times N(K+1)$ matrix with elements $L_{ij} = a_{ij}\left(X_i^\top M_G X_j\right)/T$ with $a_{ij} = (\beta_i^*)^\top \left(G^\top G/N\right)^{-1} \beta_j^*$. Also, let H be a block-diagonal matrix with elements H_i defined as $T \times K$ matrices, with columns given by the vectors $Z_{ij}\left(\frac{Z_{ij}^\top Z_{ij}}{T}\right)^{-1} \frac{Z_{ij}^\top v_{ij}}{\sqrt{T}}$ for $i = 1, \dots, N$ and $j = 1, \dots, K$; $Z_{ij} = (1, \mathrm{QC}_{(i,j)})$ with $\mathrm{QC}_{(i,j)}$ the (i,j) element of the quadratic covariation matrix QC^{ur} defined in Equation 18 and containing the covariance between the excess returns r_t^e and the factors f_t.

Lemma 1. *Under Assumptions A–D and GR1–GR3 in the Supporting Information Appendix, the asymptotic variance of the quantity* $\sqrt{T}(\hat{\eta}_{\cdot} - \eta_{\cdot})$ *is*

$$\left[\left(S - \frac{1}{N}L^\top\right)\right]^{-1} W \left[\left(S - \frac{1}{N}L\right)\right]^{-1}, \tag{28}$$

as $(m, N, T)_{seq} \to \infty$, *with* W *a* $N(K+1) \times N(K+1)$ *block-diagonal matrix given by*

$$W = \lim_{T\to\infty} \mathrm{var}\left[\left(\lambda \odot \frac{M_G X}{T}\right)^\top H\right] + \lim_{T\to\infty}\left(\frac{X^\top M_G X}{T}\right)\mathrm{var}(\varepsilon). \tag{29}$$

The following lemma proposes a consistent estimator for the asymptotic variance of $\sqrt{T}(\hat{\eta}_{\cdot} - \eta_{\cdot})$. To do this, we need a set of assumptions to be satisfied by the vector of unobservable regressors β_{it} and the error term v of the regression equation 21. These high-level assumptions can be found before theorem 2 in Pagan (1984). We also need some set of assumptions for the error term v_{it} in Equation 23 limiting the amount of cross-sectional dependence on the pricing errors. These assumptions can be found in assumptions A–D in Song (2013).

Lemma 2. *Under assumptions in Lemma 1 and Assumptions E and F in the Supporting Information Appendix, a consistent estimator of the variance of* $\sqrt{T}(\hat{\eta}_{\cdot} - \eta_{\cdot})$ *is*

$$\left[\left(\hat{S} - \frac{1}{N}\hat{L}^\top\right)\right]^{-1} \widehat{W} \left[\left(\hat{S} - \frac{1}{N}\hat{L}\right)\right]^{-1}, \tag{30}$$

where $\hat{S}$ *be a* $N(K+1) \times N(K+1)$ *block-diagonal matrix with elements* $\hat{S}_{ii} = \left(\widehat{X}_i^\top M_{\hat{G}} \widehat{X}_i\right)/T$, $\hat{L}$ *be a* $N(K+1) \times N(K+1)$ *matrix with elements* $\hat{L}_{ij} = \hat{a}_{ij}\left(\widehat{X}_i^\top M_{\hat{G}} \widehat{X}_j\right)/T$ *and* $\hat{a}_{ij} = (\hat{\beta}_i^*)^\top \left(\hat{G}^\top \hat{G}/N\right)^{-1} \hat{\beta}_j^*$, *and* $\widehat{W}$ *a block-diagonal matrix with elements*

$$\widehat{W}_i = \left(\hat{\lambda}_i \odot \frac{M_{\hat{G}} \widehat{X}_i}{T}\right)^\top T^{-1} diag\left(\hat{H}_i^\top \hat{H}_i\right)\left(\hat{\lambda}_i \odot \frac{M_{\hat{G}} \widehat{X}_i}{T}\right) + \left(\frac{\widehat{X}_i^\top M_{\hat{G}} \widehat{X}_i}{T}\right)\hat{\sigma}^2, \tag{31}$$

with $\hat{H}_i$ *defined as* $T \times K$ *matrices with columns given by the vectors* $\widehat{Z}_{ij}\left(\frac{\widehat{Z}_{ij}^\top \widehat{Z}_{ij}}{T}\right)^{-1} \frac{\widehat{Z}_{ij}^\top \hat{v}_{ij}}{\sqrt{T}}$ *for* $i = 1, \dots, N$ *and* $j = 1, \dots, K$, *where* $\widehat{Z}_{ij} = (1, \omega_{(i,j)})$ *with* $\omega_{(i,j)}$ *the elements of the realized covariance matrix* $\widehat{\Omega}^{ur}$ *defined in Equation 15 and* $\hat{v}_{ij}$ *are the residuals of the regression model (Equation 21). The quantity* $\hat{\sigma}^2$ *is defined as* $\hat{\sigma}^2 = \frac{1}{NT - N(K+1) - (N+T)R}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(r_{i,t+1}^e - \widehat{X}_{it}\hat{\eta}_i - \hat{\beta}_i^{*\top}\hat{g}_{t+1}\right)^2$.

[3]As noted by an anonymous referee, we usually face a bias when using a plugin based on infill asymptotics; see, for instance, Li and Xiu (2016). In our paper, by employing the sequential asymptotic framework we are able to use consistent estimates of the realized covariances in the autoregressive model, and hence reduce estimation error from the first step to only the error produced by using forecasts of the autoregressive model (Equation 21) without having to rely on a bias-corrected estimator for the second step.

In order to complete the estimation of the model parameters, we introduce a panel data estimator for the factor risk premia λ. In contrast to the above estimators of λ that are idiosyncratic to each risky asset, our panel data estimator is common across the N risky assets. More formally:

$$\hat{\lambda} = \frac{1}{N}\sum_{i=1}^{N}\hat{\lambda}_i. \tag{32}$$

This estimator is an average of the estimators of the factor risk premia obtained from each individual asset pricing regression equation. Our strategy in the next section to test the null hypotheses $H_0^{\alpha,\lambda}$, H_0^{α}, and H_0^{λ} is to compare the estimates $\hat{\eta}_i$ obtained from each individual asset pricing regression equation with the vector $\hat{\eta} = (0, \hat{\lambda}^{\top})^{\top}$.

3.3 | Third stage: Homogeneity tests

Now we consider asset pricing tests with cross-sectional dependence. There are several testing procedures for slope homogeneity available in the literature; see, for example, Pesaran, Smith, and Im (1996), Phillips and Sul (2003), Pesaran and Yamagata (2008), Blomquist and Westerlund (2013), and Su and Chen (2013). Ando and Bai (2015) extend these tests by accommodating the presence of cross-sectional correlation between the error terms of the different regression models indexed by $i = 1, \ldots, N$. However, to the best of our knowledge there is no available test of slope homogeneity in panel data models that accounts for generated regressors. In the following, we adapt existing tests, namely those of Pesaran and Yamagata (2008) and Ando and Bai (2015), to our asset pricing context characterized by the presence of estimated factor loadings. We also extend these tests to be suitable for testing the null hypothesis $H_0^{\alpha,\lambda}$, which includes testing for the statistical significance of the intercept of the asset pricing models. We should note that the test in Pesaran and Yamagata (2008) corresponds to exact asset pricing factor models and Ando and Bai (2015) to approximate asset pricing factor models. We discuss first Ando and Bai's (2015) test as this is more relevant empirically due to the presence of cross-sectional correlation between the pricing errors.

We consider the following test statistic:

$$\hat{\Gamma}_{\alpha,\lambda} = \frac{T\left(\hat{\eta}_{\cdot} - \hat{\eta}\iota_N\right)^{\top}\left(\hat{S} - \frac{1}{N}\hat{L}^{\top}\right)\hat{W}^{-1}\left(\hat{S} - \frac{1}{N}\hat{L}\right)(\hat{\eta}_{\cdot} - \hat{\eta}\iota_N) - [(N-1)K + N]}{\sqrt{2[(N-1)K+N]}}, \tag{33}$$

with ι_N denoting a vector of ones of dimension N and $(N-1)K + N$ denoting the number of restrictions under the null hypothesis $H_0^{\alpha,\lambda}$.

Proposition 1. *Under the null hypothesis $H_0^{\alpha,\lambda}$, and assumptions in Lemmas 1 and 2:*

$$\hat{\Gamma}_{\alpha,\lambda} \xrightarrow{d} N(0,1),$$

as $(m, N, T)_{seq} \to \infty$, with $\frac{\sqrt{T}}{N} \to 0$.

The above method can be also used to test the joint null hypothesis $H_0^{\alpha}: \alpha_i = 0$ for $i = 1, \ldots, N$. In particular, we can apply the results in Lemma 2 to the intercept parameter estimator only. This test statistic extends Gibbons et al.'s (1989) Wald-type test by considering the effect of using estimated factor loadings on the variance–covariance matrices W_α and accommodating the presence of cross-sectional dependence in the pricing errors. The corresponding test statistic is

$$\hat{\Gamma}_{\alpha} = \frac{T\hat{\alpha}^{\top}\hat{B}_{\alpha}\hat{\alpha} - N}{\sqrt{2N}}, \tag{34}$$

with $\hat{B}_\alpha$ the submatrix of $\left(\hat{S} - \frac{1}{N}\hat{L}^{\top}\right)\hat{W}^{-1}\left(\hat{S} - \frac{1}{N}\hat{L}\right)$ only containing the intercept parameter elements. Note that the joint null hypothesis H_0^{α} entertains N restrictions.[4] This test statistic will be used in the empirical application as a valid Wald-type test to assess the null hypothesis of no intercept under the presence of generated regressors.

[4] Gagliardini et al. (2016) also develop a test that extends Gibbons et al. (1989) Wald-type test. Their test accommodates estimated factor loadings within an approximate linear factor structure for large N and T. Our approach differs from that of Gagliardini et al. (2016). We work on an unconditional linear factor model, use high-frequency data to obtain consistent estimates of the factor loadings (assuming that $m \to \infty$), and our tests do not only consider the hypothesis of null intercept but also consider the homogeneity of the slope parameters that characterize the price of risk for the cross-section of risky assets.

Proposition 2. *Under the null hypothesis H_0^{α}, and assumptions in Lemmas 1 and 2:*

$$\hat{\Gamma}_{\alpha} \xrightarrow{d} N(0,1),$$

as $(m,N,T)_{seq} \to \infty$, *with* $\frac{\sqrt{T}}{N} \to 0$.

As discussed earlier, we also consider tests for the null hypothesis H_0^{λ}—that is, slope parameters only. The corresponding test statistic is

$$\hat{\Gamma}_{\lambda} = \frac{T\left(\hat{\lambda}_{\cdot} - \hat{\lambda}\iota_N\right)^{\top} \hat{B}_{\lambda} \left(\hat{\lambda}_{\cdot} - \hat{\lambda}\iota_N\right) - (N-1)K}{\sqrt{2(N-1)K}}, \tag{35}$$

with $\hat{B}_{\lambda}$ the submatrix of $\left(\hat{S} - \frac{1}{N}\hat{L}^{\top}\right)\hat{W}^{-1}\left(\hat{S} - \frac{1}{N}\hat{L}\right)$ only containing the slope parameter elements. Note that the joint null hypothesis H_0^{λ} has $(N-1)K$ restrictions. This test statistic will be used in the empirical application as a valid Wald-type test to assess the null hypothesis of slope homogeneity under the presence of generated regressors.

Proposition 3. *Under the null hypothesis H_0^{λ}, and assumptions in Lemmas 1 and 2:*

$$\hat{\Gamma}_{\lambda} \to N(0,1),$$

as $(m,N,T)_{seq} \to \infty$, *with* $\frac{\sqrt{T}}{N} \to 0$.

Propositions 1, 2, and 3 formalize the limiting distributions of the asset pricing tests with cross-sectional dependence. In Supporting Information Appendix B we also consider the simpler case of asset pricing tests without cross-sectional dependence. In this scenario we assume no cross-sectional dependence among the pricing errors ν_{it} in Equation 23. Although this is not possible in our model setup due to the presence of correlation in the pricing errors induced by the observed common factors $\tilde{f}_{t+1}$ in Equation 10 or the set of observed and unobserved common factors $\tilde{g}_{t+1}$ in Equation 14, we present this for completeness as an extension of the Swamy-type tests derived by Pesaran and Yamagata (2008) to entertaining generated regressors. In turn, these tests, together with the standard Swamy-type tests, serve to illustrate the contribution of our tests for achieving correct size and power.

We study the finite-sample performance of these tests with Monte Carlo simulations. The results are provided in the Supporting Information Appendix C. The test results show that: (i) the tests are able to separate heterogeneity arising from the intercepts (i.e., α_i) and the slopes (i.e., λ_i); (ii) in the presence of unobserved factors, Ando and Bai (2015) type tests are the only ones that have correct empirical size; (iii) the presence of generated regressors requires a variance correction that is achieved with our proposed test. We also note in this detailed simulation exercise that there are no significant differences in the finite-sample properties of the tests for the null intercept hypothesis between Ando and Bai and our correction for generated regressors. Nevertheless, there are significant differences from the Gibbons et al. (1989) test, which is oversized when the factor loadings are estimated. Overall, our proposed tests, $\hat{\Gamma}_{\alpha}$, $\hat{\Gamma}_{\lambda}$, and $\hat{\Gamma}_{\alpha,\lambda}$, have the best performance in terms of correct empirical size and power for detection of departures from the different null hypotheses.

4 | EMPIRICAL APPLICATION

In this section we apply the asset pricing tests developed above to explain the excess returns on 47 US industry portfolios maintained in Kenneth French's data library spanning the period July 1963 to December 2014.[5] The return on the risk-free asset is proxied by daily returns on the US 3-month Treasury bill, also available from this website. Our aim is to test the suitability of the beta asset pricing model for different specifications of the pricing factors. The empirical strategy for the practical implementation of the above tests is as follows.

First, for each individual industry portfolio we compute the realized covariance measures $\hat{\beta}_{it}$ proxying the dynamic quantities β_{it}. These observed measures are unbiased estimators of the actual unobserved conditional covariances between the excess returns on the risky portfolios and the pricing factors. In a second stage, we regress the excess industry portfolio

[5]The dataset provided on Kenneth French's website comprises the returns on 49 industry portfolios at daily frequencies; however, reliable information at daily frequency is not available for the full sample period for Healthcare (available from July 1969) and Computer Software (available from July 1965). Hence these industries are removed from the empirical study. In a similar vein, Fama and French (2008) suggest excluding from the empirical datasets the firms (and thus the industry portfolios) with Standard Industrial Classification codes between 6000 and 6999 (banking, insurance, real estate, and trading sectors).