Ex-Post Risk Premia and Tests of Multi-Beta Models in Large Cross-Sections

VALENTINA RAPONI, CESARE ROBOTTI, and PAOLO ZAFFARONI[1]

April 2016

JEL # G12, C12, C52

[1]Raponi, Robotti, and Zaffaroni are from Imperial College London. We thank Adelina Barbalau, Fernando Chague, Soohun Kim, Lubos Pastor, Krishna Ramaswamy, Jay Shanken, Raman Uppal, Pietro Veronesi, seminar participants at Imperial College London, Luxembourg School of Finance, University of Georgia, University of Southampton, Tinbergen Institute, as well as conference participants at the 2016 Meetings of the Midwest Finance Association, the 2015 Meetings of the Brazilian Finance Society, and CFE 2015 for helpful discussions and comments. Corresponding author: Cesare Robotti, Finance Department, Imperial College Business School, South Kensington Campus, London, SW7 2AZ, UK; E-mail: c.robotti@imperial.ac.uk.

Ex-Post Risk Premia and Tests of Multi-Beta Models in Large Cross-Sections

Abstract

A limiting theory for estimating and testing linear beta-pricing models with a large number of assets and a fixed time-series sample size is presented. Since the ordinary least squares (OLS) estimator of the ex-ante risk premia is asymptotically biased and inconsistent in this context, the focus of the paper is on the modified OLS estimator of the ex-post risk premia proposed by Shanken (1992). We derive the asymptotic distribution of this estimator and show how its limiting variance can be consistently estimated. In addition, we characterize the asymptotic distribution of a cross-sectional test of the fundamental beta-pricing relation. Finally, we show how our results can be extended to deal with unbalanced panels. The practical relevance of our findings is demonstrated using Monte Carlo simulations and an empirical application to beta-pricing models with traded risk factors. Our analysis suggests that the market, size, and value factors are often priced in the cross-section of NYSE-AMEX-NASDAQ individual stock returns over short time spans. Overall, there is not much evidence of pricing for the profitability and investment factors of Fama and French (2015).

Keywords: Beta-pricing models; Ex-post risk premia; Two-pass cross-sectional regressions; Large N asymptotics; Specification test.

JEL classification numbers: C12; C13; G12.

The traditional empirical methodology for exploring asset pricing models entails estimation of asset betas (systematic risk measures) from time-series factor model regressions, followed by estimation of risk premia via cross-sectional regressions (CSR) of asset returns on the estimated betas. In the classic analysis of the capital asset pricing model (CAPM) by Fama and MacBeth (1973), a CSR is run each month, with inference ultimately based on the time-series mean and standard error of the monthly risk premium estimates. Also see the related paper by Black, Jensen, and Scholes (1972).

A formal econometric analysis of the two-pass methodology was first provided by Shanken (1992). He shows how the asymptotic standard errors of the second-pass ordinary least squares (OLS) and generalized least squares (GLS) risk premium estimators are influenced by estimation error in the first-pass betas, requiring an adjustment to the traditional Fama-MacBeth standard errors. A test of the validity of the pricing model's constraint on expected returns can also be derived from the CSR residuals (see, for example, Shanken (1985)).[1]

While most of the limiting results in Shanken (1992) require that the time-series sample size, T, is large and the number of securities, N, is fixed, Section 6 of his paper (see also Litzenberger and Ramaswamy (1979)) presents some preliminary results on the estimation of the second-pass risk premia when N is large and T is fixed. Shanken (1992, p. 20) argues that "This perspective is particularly relevant given the large number of securities for which returns data are available." Litzenberger and Ramaswamy (1979, p. 178) describe the pros and cons of using portfolios instead of individual securities in asset-pricing tests: "Because of the errors in variables problem, most previous empirical tests have grouped stocks into portfolios. Since errors in measurement in betas for different securities are less than perfectly correlated, grouping risky assets into portfolios would reduce the asymptotic bias in OLS estimators. However, grouping results in a reduction of efficiency caused by the loss of information." We don't take a stand on whether a researcher should use portfolios or individual securities in the estimation and testing of asset-pricing models. Motivated by the increasingly widespread use of large cross-sections of individual securities (with a limited time-series dimension) in empirical asset-pricing studies, we think that researchers would benefit

[1] Jagannathan and Wang (1998) relax the conditional homoskedasticity assumption in Shanken (1992) and derive expressions for the asymptotic variances of the OLS and GLS estimators that are valid under fairly general distributional assumptions. Hou and Kimmel (2006), Shanken and Zhou (2007), and Kan, Robotti, and Shanken (2013) provide a unifying treatment of the two-pass methodology in the presence of global (or fixed) model misspecification. We refer the readers to Jagannathan, Skoulakis, and Wang (2010), Kan and Robotti (2012), and Gospodinov and Robotti (2012) for a synthesis of the two-pass methodology.

from a rigorous large N and fixed T econometric framework in addition to the large T and fixed N methods that have already been proposed in the literature.

Since a consistent estimate of the "ex-ante" risk premia cannot be obtained with T fixed,[2] in the following analysis we follow Shanken (1992) and focus on "ex-post" risk premia. The ex-post risk premia equal the ex-ante risk premia plus the (unconditionally) unexpected factor outcomes. In this context, a risk premium estimator is $\sqrt{N}$-consistent if it converges in probability at speed $\sqrt{N}$ to the vector of ex-post risk premia as $N \to \infty$. We start by showing that Shanken's (1992) modified estimator of the ex-post risk premia is a member of a broader class of OLS bias-adjusted estimators. Under fairly standard assumptions, we derive the rate of convergence and show the $\sqrt{N}$-consistency of the modified Shanken's estimator. More importantly, we establish the asymptotic normality of the Shanken's bias-adjusted estimator and derive an explicit expression for its asymptotic covariance matrix. We also show how this covariance matrix can be consistently estimated and used to conduct inference on the risk premia estimates. Furthermore, we provide a new test for the validity of the beta-pricing restrictions and characterize its distribution when $N \to \infty$ and T is fixed. Finally, we show how our results need to be modified when the panel is unbalanced. It is worth emphasizing that our analysis is applicable to models with traded factors as well as to models that include also factors that are not traded.

We also explore the small N properties of our various test statistics via Monte Carlo simulations. We calibrate the model parameters using real data on individual stocks from the Center for Research in Security Prices (CRSP). Overall, our simulation results suggest that the tests are reliable for the sample sizes often encountered in empirical finance.

Empirically, our interest is in rigorously evaluating the performance of several prominent beta-pricing models using individual monthly stock returns from the CRSP database. In addition to the basic CAPM, we study the three-factor and five-factor models of Fama and French (1993, 2015). Although these models were primarily motivated by empirical observation, their size, value, profitability, and investment factors are sometimes viewed as proxies for more fundamental economic factors. While our specification test often rejects the hypothesis of a perfect fit for these three models, we find some convincing evidence of pricing for the market, size, and value factors over

[2]When T is fixed, increasing the number of assets affects the residual variation but does not eliminate the uncertainty about the unanticipated factor realizations (see Shanken (1992) for details).

short time spans. The evidence of pricing for the profitability and investment factors proposed by Fama and French (2015) is somewhat weaker.

To our knowledge, we are the first to provide a formal asymptotic analysis of Shanken's bias-adjusted estimator when N is large and T is fixed. Our limiting results should be viewed as useful complements to the simulation analysis in Chordia, Goyal, and Shanken (2015). There are only a few other papers concerned with pricing that deal with a large cross section and a small time series. Under the assumption that the covariance matrix of the disturbances is diagonal and known, Litzenberger and Ramaswamy (1979) employ a modified version of the weighted least squares second-pass estimator that is consistent without portfolio grouping. Kim and Skoulakis (2014) employ the regression calibration approach to obtain a $\sqrt{N}$-consistent estimator of the ex post risk premia in a two-pass CSR setting. However, the results in Kim and Skoulakis (2014) are only applicable when the factors are returns on zero net investment portfolios. When T, in addition to N, is allowed to go to infinity, Gagliardini, Ossola, and Scaillet (2015) and Bai and Zhou (2015) derive the limiting properties of some bias-adjusted estimators of the ex-post and ex-ante second-pass risk premia, respectively. In particular, the asymptotic theory in Gagliardini, Ossola, and Scaillet (2015) requires that N and T tend to infinity jointly such that N grows not faster than a power (less than 3) of T. Finally, Jegadeesh and Noh (2014) and Pukthuanthong, Roll, and Wang (2014) propose instrumental variable estimators of the second-pass risk premia as an alternative to the asymptotically biased OLS CSR estimator of Fama and MacBeth (1973).

When it comes to a test of the fundamental beta-pricing relation, we are the first to propose a two-pass CSR test of the model's pricing errors that is valid when N is large and T is fixed. Extending the classical test of Gibbons, Ross, and Shanken (1989), Pesaran and Yamagata (2012) propose a number of tests of the null hypothesis of zero alphas in the first-pass relation. Their setup accommodates only traded factors in the analysis and the feasible versions of their tests are justified only when both N and T tend to infinity jointly such that $N/T^3 \to 0$. Gagliardini, Ossola, and Scaillet (2015) extend the work of Pesaran and Yamagata (2012) to the case in which the factors are not necessarily traded portfolios.

The rest of the paper is organized as follows. Section I provides a brief review of the two-pass CSR methodology and introduces the main notation. Section II presents the asymptotic analysis of Shanken's (1992) bias-adjusted estimator of the ex-post risk premia when $N \to \infty$ and T is fixed.

Moreover, under the assumption of zero ex-ante pricing errors, we derive the limiting distribution of the specification test described above when N is large and T is fixed. Finally, we show how the main analysis can be extended to accommodate unbalanced panels. We explore the small N properties of the various tests in Section III. Section IV presents our main empirical findings, and Section V summarizes our conclusions. The proofs of the main results are collected in the Appendix.

1 Two-Pass Methodology

A beta-pricing model seeks to explain cross-sectional differences in expected asset returns in terms of asset betas computed relative to the model's systematic economic factors. Let $f_t = [f_{1t}, \ldots, f_{Kt}]'$ be a K-vector of observable factors at time t and $R_t = [R_{1t}, \ldots, R_{Nt}]'$ be an N-vector of test asset returns at time t ($t = 1, \ldots, T$).

Assume that asset returns are governed by the following multifactor model:

$$R_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \epsilon_{it} = \alpha_i + \beta_i' f_t + \epsilon_{it}, \qquad i = 1, \ldots, N, \qquad t = 1, \ldots, T, \quad (1)$$

where α_i is an asset specific intercept scalar, $\beta_i = [\beta_{i1}, \ldots, \beta_{iK}]'$ is a vector of multiple regression betas of asset i w.r.t. the K factors, and the ϵ_{it}'s are factor model residuals.

Assumption 1. (Loadings) As $N \to \infty$,

$$\frac{1}{N} \sum_{i=1}^{N} \beta_i \to \mu_\beta, \quad (2)$$

$$\frac{1}{N} \sum_{i=1}^{N} \beta_i \beta_i' \to \Sigma_\beta, \quad (3)$$

where Σ_β is a finite symmetric and positive-definite matrix.

The first part of Assumption 1 states that the limiting cross-sectional average of the betas exists, while the second part states that the limiting cross-sectional average of the squared betas exists and is a symmetric and positive-definite matrix. Without loss of generality, we assume the β_i to be non random.[3]

In vector and matrix notation, we can write the above model as

$$R_t = \alpha + B f_t + \epsilon_t, \qquad t = 1, \ldots, T, \quad (4)$$

[3] See Gagliardini, Ossola, and Scaillet (2015) for a treatment of beta-pricing models with random betas.

where $\alpha = [\alpha_1, \ldots, \alpha_N]'$, $B = [\beta_1, \ldots, \beta_N]'$, and $\epsilon_t = [\epsilon_{1t}, \ldots, \epsilon_{Nt}]'$.

Assumption 2. (Shanken, 1992) Assume that the vector ϵ_t is independently and identically distributed (*i.i.d.*) over time with

$$E[\epsilon_t | F] = 0_N \tag{5}$$

and

$$\text{Var}[\epsilon_t | F] = \Sigma \qquad (\text{rank } N), \tag{6}$$

where Σ has generic element $\sigma_{ij} = E[\epsilon_{it}\epsilon_{jt}]$ (for $i, j = 1, \ldots, N$), $F = [f_1, \ldots, f_T]'$ is the $T \times K$ matrix of factors, and 0_N is an N-vector of zeros.

The *i.i.d.* assumption over time is common to many studies, including Shanken (1992). This assumption could be relaxed at the cost of a more cumbersome notation. The proposed K-factor beta-pricing model specifies that asset expected returns are linear in B, that is,

$$E[R_t] = X\Gamma, \tag{7}$$

where $X = [1_N,\ B]$ is assumed to be of full column rank, 1_N is an N-vector of ones, and $\Gamma = [\gamma_0,\ \gamma_1']'$ is a vector consisting of the zero-beta rate (γ_0) and ex-ante risk premia on the K factors (γ_1). When the model is misspecified, the N-vector of pricing errors, $e = E[R_t] - X\Gamma$, will be nonzero for all values of Γ.

Assumption 3. Assume that $E[f_t]$ does not vary over time and denote this expectation by $E[f]$.

To introduce the notion of ex-post risk premia, let $\bar{R}_i = \frac{1}{T}\sum_{t=1}^{T} R_{it}$, $\bar{R} = [\bar{R}_1, \ldots, \bar{R}_N]'$, and $\bar{\epsilon} = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t$. Averaging (4) over time, imposing (7), and noting that $E[R_t] = \alpha + BE[f]$ yield

$$\bar{R} = X\Gamma^P + \bar{\epsilon}, \tag{8}$$

where $\Gamma^P = [\gamma_0,\ \gamma_1^{P\prime}]'$ and

$$\gamma_1^P = \gamma_1 + \bar{f} - E[f]. \tag{9}$$

By (8), expected returns are still linear in the asset betas conditional on the factor outcomes. The random coefficient vector γ_1^P is referred to, accordingly, as the vector of ex-post risk premia. It equals the ex-ante risk premia plus the (unconditionally) unexpected factor outcomes. Assumption 3 rules out time variation in the ex-post risk premia γ_1^P. When $T \to \infty$, $\bar{f}$ will converge to $E[f]$ and the ex-post and ex-ante risk premia will coincide. However, in general, γ_1 and γ_1^P will differ when T is finite. Therefore, the notion of ex-post risk premia naturally emerges when estimating beta-pricing models in a large N and fixed T framework.

We now turn to estimation of the model. The popular two-pass method first obtains estimates $\hat{\beta}_i$, the betas of asset i, by running the following multivariate regression:

$$R_i = \alpha_i 1_T + F\beta_i + \epsilon_i, \quad i = 1, \ldots, N, \tag{10}$$

where $R_i = [R_{i1}, \ldots, R_{iT}]'$ is a time series of returns on asset i, 1_T is a T-vector of ones, and $\epsilon_i = [\epsilon_{i1}, \ldots, \epsilon_{iT}]'$. Defining $\tilde{F} = F - 1_T \bar{f}'$, it is easy to show that

$$\hat{\beta}_i = \beta_i + (\tilde{F}'\tilde{F})^{-1}\tilde{F}'\epsilon_i, \tag{11}$$

or, in matrix form,

$$\hat{B} = B + \epsilon'\mathcal{P}, \tag{12}$$

where $\epsilon = [\epsilon_1, \ldots, \epsilon_N]$ is a $T \times N$ matrix and $\mathcal{P} = \tilde{F}(\tilde{F}'\tilde{F})^{-1}$. We then run a single CSR of the sample mean vector $\bar{R}$ on $\hat{X} = [1_N,\ \hat{B}]$ to estimate Γ (Γ^P) in the second pass. Specifically, we have

$$\bar{R} = \hat{X}\Gamma + \eta, \tag{13}$$

where $\eta = \left(\bar{\epsilon} + B(\bar{f} - E[f]) - (\hat{X} - X)\Gamma\right)$ and

$$\bar{R} = \hat{X}\Gamma^P + \eta^P, \tag{14}$$

where $\eta^P = \left(\bar{\epsilon} - (\hat{X} - X)\Gamma^P\right)$. If we use the identity matrix as the weighting matrix in the second-pass CSR, we obtain the following OLS estimator for both the feasible representations in (13) and (14):

$$\hat{\Gamma} = \begin{bmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \end{bmatrix} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R}. \tag{15}$$

As Shanken (1992) points out, one cannot hope for a consistent estimate of Γ in (13) with T fixed. The reason is that $\bar{f}$ does not converge in probability to $E[f]$ unless $T \to \infty$. Although Bai and Zhou (2015) conjecture that the impact of the term $\bar{f} - E[f]$ is small in practice, we follow Shanken (1992) and conduct our analysis based on the representation in (14). Shanken (1992) and Bai and Zhou (2015) among others show that the OLS estimator of Γ^P in (15) is biased and inconsistent when T is fixed. Nevertheless, Shanken (1992) shows that the bias of the OLS estimator can be corrected. Denote the trace operator by $\text{tr}(\cdot)$ and a K-vector of zeros by 0_K. In addition, let

$\hat{\sigma}^2 = \frac{1}{N(T-K-1)}\text{tr}(\hat{\epsilon}'\hat{\epsilon})$, where $\hat{\epsilon}$ is the $T \times N$ matrix of residuals from the first-pass. Then, the bias-adjusted OLS estimator of Shanken (1992) is given by[4]

$$\hat{\Gamma}^* = \begin{bmatrix} \hat{\gamma}_0^* \\ \hat{\gamma}_1^* \end{bmatrix} = \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \frac{\hat{X}'\bar{R}}{N}, \tag{16}$$

where

$$\hat{\Sigma}_X = \frac{\hat{X}'\hat{X}}{N} \tag{17}$$

and

$$\hat{\Lambda} = \begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}. \tag{18}$$

Conditional on the factor realizations, in the next section we will provide a formal asymptotic analysis of $\hat{\Gamma}^*$ and we will investigate the limiting behavior of a model specification test based on the ex-post sample pricing errors

$$\hat{e}^P = \bar{R} - \hat{X}\hat{\Gamma}^*. \tag{19}$$

The errors-in-variables (EIV) correction in (16) entails subtracting the estimated covariance matrix of the beta estimation errors from $\hat{B}'\hat{B}$, in an attempt to better approximate the matrix $B'B$. However, it is possible that this EIV correction will overshoot and that the matrix $\left(\hat{\Sigma}_X - \hat{\Lambda}\right)$ will not be positive definite. This complicates the analysis of the finite-sample properties of $\hat{\Gamma}^*$. To deal with the possibility that the estimator will occasionally produce extreme results, in the simulation and empirical sections of the paper we multiply the matrix $\hat{\Lambda}$ by a scalar k ($0 \leq k \leq 1$), effectively implementing a shrinkage estimator. If k is zero, we get the OLS estimator back, whereas if k is one we obtain the modified Shanken's estimator. The choice of k, the parameter that determines the degree of shrinkage between the OLS and modified OLS estimators, is based on the eigenvalues of the matrix $\left(\hat{\Sigma}_X - k\hat{\Lambda}\right)$. Starting from $k = 1$, if the minimum eigenvalue of this matrix is negative and/or the condition number of this matrix is bigger than 20, then we lower k by an arbitrarily small amount. We iterate this procedure until the minimum eigenvalue is positive and the condition number becomes smaller than 20.[5] In our simulation experiments, we find that this shrinkage estimator is "virtually unbiased". This is mainly due to the fact that in our simulations we encounter very few instances in which $\left(\hat{\Sigma}_X - \hat{\Lambda}\right)$ is not positive definite.

[4] As in most past studies and to keep the notation somewhat manageable, we do not distinguish between general factors and factors that are portfolio returns. Therefore, we do not incorporate the additional pricing restriction that is implied when a given factor is a portfolio return.

[5] Following Greene (2008), Gagliardini, Ossola, and Scaillet (2015) rely on similar methods to implement their trimming conditions.

2 Asymptotic Analysis

The analysis in this section assumes that $N \to \infty$ and T is fixed. In addition, our results are derived under the assumption of weak cross-sectional dependence between the factor model disturbances (see Assumption 4 in Appendix A). This assumption is routinely made in most studies on large N asymptotics. We first present the limiting distribution of Shanken's bias-adjusted estimator and explain how its asymptotic covariance matrix can be consistently estimated. In addition, we show that the modified estimator of Shanken (1992) is a member of a broader family of OLS bias-adjusted estimators. Finally, we characterize the limiting behavior of a test of the beta-pricing restrictions and show how our analysis can be extended to deal with an unbalanced panel.

2.1 Asymptotic Distribution of the Bias-Adjusted OLS Estimator

In many empirical studies, interest lies in the point estimates of the second-pass risk premia. A statistically significant element of $\hat{\gamma}_1^*$ associated with a given factor is often interpreted as evidence that its risk premium is nonzero. In this subsection, we study the asymptotic distribution of $\hat{\Gamma}^*$ under the assumption that the model is correctly specified. Let $\Sigma_X = \begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta \end{bmatrix}$, $\sigma^2 = \lim \frac{1}{N}\sum_{i=1}^N \sigma_i^2$ with $\sigma_i^2 = E[\epsilon_{it}^2]$, $U_\epsilon = \lim \frac{1}{N}\sum_{i,j=1}^N E\left[\mathrm{vec}(\epsilon_i\epsilon_i' - \sigma_i^2 I_T)\mathrm{vec}(\epsilon_j\epsilon_j' - \sigma_j^2 I_T)'\right]$, $M = I_T - D(D'D)^{-1}D'$, where I_T is a $T \times T$ identity matrix and $D = [1_T,\ F]$, $Q = \frac{1_T}{T} - \mathcal{P}\gamma_1^P$, and $Z = (Q \otimes \mathcal{P}) + \frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}$. In the following theorem, we provide the rate of convergence and the limiting distribution of $\hat{\Gamma}^*$.

Theorem 1

(i) Under Assumptions 1–3 and 4 in Appendix A,

$$\hat{\Gamma}^* - \Gamma^P = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{20}$$

(ii) Under Assumptions 1–3 and 4–5 in Appendix A,

$$\sqrt{N}\left(\hat{\Gamma}^* - \Gamma^P\right) \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V + \Sigma_X^{-1} W \Sigma_X^{-1}\right), \tag{21}$$

where

$$V = \frac{\sigma^2}{T}\left[1 + \gamma_1^{P\prime}\left(\tilde{F}'\tilde{F}/T\right)^{-1}\gamma_1^P\right]\Sigma_X^{-1} \tag{22}$$

and

$$W = \begin{bmatrix} 0 & 0_K' \\ 0_K & Z'U_\epsilon Z \end{bmatrix}. \quad (23)$$

Proof: See Appendix B and Lemmas 1 to 5 in Appendix A.

Note that the variance expression in (21) is very simple. The first part of this asymptotic covariance matrix, V, accounts for the estimation error in the betas, and it is very similar to the large T variance expression of the second-pass OLS estimator proposed by Shanken (1992) in Theorem 1(ii). The term $c \equiv \gamma_1^{P\prime}\left(\tilde{F}'\tilde{F}/T\right)^{-1}\gamma_1^P$ is an asymptotic adjustment for EIV, and $c\frac{\sigma^2}{T}\Sigma_X^{-1}$ is the corresponding EIV component of variance. As Shanken (1992) points out, the EIV adjustment reflects the fact that the variance of the beta estimates is directly related to residual variance and inversely related to factor variability.

The second part of the asymptotic covariance matrix, $\Sigma_X^{-1}W\Sigma_X^{-1}$, represents the additional fixed T and large N adjustment to the overall variance. This term is the inflation in variance due to the bias adjustment and vanishes only when $T \to \infty$. In addition, the W matrix accounts for the cross-sectional variation in the residual variances of asset returns.

To conduct statistical inference, we need a consistent estimator of the asymptotic covariance matrix in Theorem 1(ii). Let $M^{(2)} = M \odot M$, where $\odot$ denotes the Hadamard product operator. In addition, define

$$\hat{\sigma}_4 = \frac{\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N}\hat{\epsilon}_{it}^4}{3\mathrm{tr}\left(M^{(2)}\right)}. \quad (24)$$

and let

$$\hat{Z} = (\hat{Q} \otimes \mathcal{P}) + \frac{\mathrm{vec}(M)}{T-K-1}\hat{\gamma}_1^{*\prime}\mathcal{P}'\mathcal{P} \quad (25)$$

with

$$\hat{Q} = \frac{1_T}{T} - \mathcal{P}\hat{\gamma}_1^*. \quad (26)$$

The following theorem provides a consistent estimator of the asymptotic covariance matrix of the estimates.

Theorem 2 *Under Assumptions 1 to 3 and 4–5 in Appendix A, we have*

$$\hat{V} + \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\hat{W}\left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \xrightarrow{p} V + \Sigma_X^{-1}W\Sigma_X^{-1}, \quad (27)$$

where

$$\hat{V} = \frac{\hat{\sigma}^2}{T}\left[1+\hat{\gamma}_1^{*\prime}\left(\tilde{F}'\tilde{F}/T\right)^{-1}\hat{\gamma}_1^*\right](\hat{\Sigma}_X-\hat{\Lambda})^{-1}, \tag{28}$$

$$\hat{W} = \begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{Z}'\hat{U}_\epsilon\hat{Z} \end{bmatrix}, \tag{29}$$

and $\hat{U}_\epsilon$ is a consistent plug-in estimator of U_ϵ (see Appendix C).

Proof: See Appendix B and Lemmas 1 to 6 in Appendix A.

It is worth noting that the modified estimator of Shanken (1992) belongs to a larger family of OLS bias-adjusted estimators. Starting from the standard OLS estimator, it is easy to show that

$$\hat{\Gamma} = \Gamma^P + (\hat{X}'\hat{X})^{-1}\left(X'\bar{\epsilon} - X'(\hat{X}-X)\Gamma^P + (\hat{X}-X)'\bar{\epsilon} - (\hat{X}-X)'(\hat{X}-X)\Gamma^P\right). \tag{30}$$

By Lemmas 2(i), 3(i), 4(i), and 5(i) in Appendix A, we have

$$(\hat{X}'\hat{X})^{-1}\left(X'\bar{\epsilon} - X'(\hat{X}-X)\Gamma^P + (\hat{X}-X)'\bar{\epsilon}\right) = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{31}$$

Moreover, Lemmas 2(i) and 2(iii) in Appendix A imply that

$$(\hat{X}'\hat{X})^{-1}(\hat{X}-X)'(\hat{X}-X)\Gamma^P = O_p(1). \tag{32}$$

While the bias term in (31) vanishes when $N \to \infty$, the bias term in (32) renders the OLS estimator asymptotically biased and inconsistent. Nevertheless, the bias in (32) can be corrected. Using Lemma 1(i) in Appendix A, a $\sqrt{N}$-consistent (target) estimator of Γ^P is given by

$$\tilde{\Gamma} = \left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\bar{R} + \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}\breve{\Gamma}, \tag{33}$$

where $\breve{\Gamma}$ denotes any preliminary $\sqrt{N}$-consistent estimator of Γ^P. If we impose the restriction that $\tilde{\Gamma} = \breve{\Gamma}$, then

$$\left[I_{K+1} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}\right]\tilde{\Gamma} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R},$$

which implies that

$$\tilde{\Gamma} = \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\frac{\hat{X}'\bar{R}}{N} = \hat{\Gamma}^*, \tag{34}$$

that is, the modified estimator of Shanken (1992). Therefore, the bias-adjusted estimator of Shanken (1992) is obtained when the target and preliminary estimators are set equal to each other. In general, it is up to the econometrician to decide what $\breve{\Gamma}$ to use in the analysis and the choice of preliminary estimator clearly involves some degree of arbitrariness. In this sense, the modified estimator of Shanken (1992) has the attractive feature of allowing us to sidestep the issue of choosing a preliminary $\sqrt{N}$-consistent estimator of Γ^P.

2.2 Limiting Distribution of the Specification Test

In this section, we are interested in deriving an OLS-type test of the validity of the beta-pricing model. The null hypothesis underlying the beta-pricing restriction can be formulated as

$$H_0: \ e_i = 0 \qquad \forall i = 1, 2, \ldots, \tag{35}$$

where e_i is the pricing error associated with asset i. Let $X_i = [1, \ \beta_i']$, $\hat{X}_i = [1, \ \hat{\beta}_i']$, and denote by $\hat{e}_i^P$ the ex-post sample pricing error for asset i. Then, we have

$$\begin{aligned} \hat{e}_i^P &= \bar{R}_i - \hat{X}_i \hat{\Gamma}^* && (36) \\ &= e_i + Q'\epsilon_i - \hat{X}_i \left(\hat{\Gamma}^* - \Gamma^P \right). && (37) \end{aligned}$$

It follows that

$$\hat{e}_i^P \xrightarrow{p} e_i + Q'\epsilon_i \equiv e_i^P. \tag{38}$$

Equation (38) shows that even when the ex-ante pricing error, e_i, is zero, $\hat{e}_i^P$ will not converge in probability to zero. This is a consequence of the fact that when T is fixed, $Q'\epsilon_i = \bar{\epsilon}_i$ will not converge to zero even under the null of zero ex-ante pricing errors. This is the price that we have to pay when N is large and T is fixed. Nonetheless, a test of H_0 with good size and power properties can be developed. Since we estimate Γ^P via OLS cross-sectional regressions, we propose a test based on the sum of the squared ex-post sample pricing errors, that is,

$$\hat{\mathcal{Q}} = \frac{1}{N} \sum_{i=1}^{N} (\hat{e}_i^P)^2. \tag{39}$$

Consider the centered statistic

$$\mathcal{S} = \sqrt{N} \left(\hat{\mathcal{Q}} - \frac{\hat{\sigma}^2}{T} \left(1 + \hat{\gamma}_1^{*\prime} (\tilde{F}'\tilde{F}/T)^{-1} \hat{\gamma}_1^* \right) \right). \tag{40}$$

The following theorem provides the limiting distribution of $\mathcal{S}$ under $H_0 : e_i = 0$ for all i.

Theorem 3 *Under Assumptions 1 to 3 and 4–5 in Appendix A and under $H_0: \ e_i = 0$ for all i, we have*

$$\mathcal{S} \xrightarrow{d} \mathcal{N}\left(0, \mathcal{V}\right), \tag{41}$$

where $\mathcal{V} = Z_Q' U_\epsilon Z_Q$ and $Z_Q = (Q \otimes Q) - \frac{\text{vec}(M)}{T-K-1} Q'Q$.

Proof: See Appendix B and Lemmas 1 to 5 in Appendix A.

The expression for the asymptotic variance of the test in (41) is rather simple. This variance can be consistently estimated by replacing Q with $\hat{Q}$ and U_ϵ with $\hat{U}_\epsilon$. Specifically, using Theorem 2 and Lemma 6 in Appendix A, we have

$$\hat{Z}_Q' \hat{U}_\epsilon \hat{Z}_Q \xrightarrow{p} Z_Q' U_\epsilon Z_Q, \tag{42}$$

where

$$\hat{Z}_Q = \left(\hat{Q} \otimes \hat{Q}\right) - \frac{\text{vec}(M)}{T-K-1} \hat{Q}'\hat{Q}. \tag{43}$$

Then, under H_0, it follows that

$$\frac{\mathcal{S}}{(\hat{Z}_Q' \hat{U}_\epsilon \hat{Z}_Q)^{\frac{1}{2}}} \xrightarrow{d} \mathcal{N}(0,1). \tag{44}$$

It can be shown that our test will have power when $e_i^2 > 0$ for a sufficiently large number of securities. In the next section, we will undertake a Monte Carlo simulation experiment calibrated to real data in order to determine whether our test possesses desirable size and power properties.

2.3 Unbalanced Panel

In this section, we extend out methodological results to the case of an unbalanced panel. Following Gagliardini, Ossola, and Scaillet (2015), we assume a missing at random design (see, for example, Rubin, 1976), that is independence between unobservability and return generating process. This allows us to keep the factor structure linear. In the following analysis, we explicitly account for the randomness of T_i, the time-series sample size for asset i. Define the following $T \times T$ matrix

$$J_i = \text{diag}(J_{i1} \ldots J_{it} \ldots J_{iT}) \qquad i = 1, \ldots, N, \tag{45}$$

where $J_{it} = 1$ if the return on asset i is observed by the econometrician at date t, and zero otherwise. We assume that J_{it} is *i.i.d.* across i and t. In addition, let $R_{i,u} = J_i R_i$, $F_{i,u} = J_i F$, and $\epsilon_{i,u} = J_i \epsilon_i$, and assume that asset returns are governed by the multifactor model

$$J_{it}R_{it} \quad = \quad J_{it}\alpha_i + J_{it}f_t'\beta_i + J_{it}\epsilon_{it}, \tag{46}$$

that is, the same data generating process of the previous section multiplied by J_{it}. Let $\bar{R}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it}R_{it}$, $\bar{f}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it}f_t$, and $\bar{\epsilon}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it}\epsilon_{it}$. Averaging (46) over time, imposing the beta-pricing restriction, and noting that $E[R_{it}] = \alpha_i + \beta_i' E[f]$ yields

$$\bar{R}_{i,u} = \gamma_0 + \hat{\beta}_{i,u}'\gamma_{1i,u}^P + \eta_{i,u}^P, \tag{47}$$

where $\gamma_{1i,u}^P = \gamma_1 + \bar{f}_{i,u} - E[f]$, $\eta_{i,u}^P = \bar{\epsilon}_{i,u} - (\hat{\beta}_{i,u} - \beta_i)'\gamma_{1i,u}^P$, $\hat{\beta}_{i,u} = \beta_i + \mathcal{P}_{i,u}'\epsilon_i$, $\mathcal{P}_{i,u} = \tilde{F}_{i,u}(\tilde{F}_{i,u}'\tilde{F}_{i,u})^{-1}$, and $\tilde{F}_{i,u} = F_{i,u} - J_i 1_T \bar{f}_{i,u}'$. Since the panel is unbalanced, there is now a sequence of ex-post risk premia, one for each asset i.

In matrix form, we have

$$\bar{R}_u = \gamma_0 1_N + \begin{bmatrix} \hat{\beta}_{1,u}' & & 0_{K(N-1)}' \\ \vdots & \ddots & \vdots \\ 0_{K(N-1)}' & & \hat{\beta}_{N,u}' \end{bmatrix} \begin{bmatrix} \gamma_{11,u}^P \\ \vdots \\ \gamma_{1N,u}^P \end{bmatrix} + \begin{bmatrix} \eta_{1,u}^P \\ \vdots \\ \eta_{N,u}^P \end{bmatrix}, \tag{48}$$

where $\bar{R}_u = (\bar{R}_{1,u}, \dots, \bar{R}_{N,u})'$. Define the $N \times K$ matrix $\hat{X}_u = \left[1_N,\ \hat{B}_u\right]$, where $\hat{B}_u = \left(\hat{\beta}_{1,u}, \dots, \hat{\beta}_{N,u}\right)'$. Denote by $\hat{\epsilon}_{i,u}$ the T-vector of residuals from the first-pass (unbalanced) OLS regressions in

$$R_{i,u} = \alpha_i J_i 1_T + F_{i,u}\beta_i + \epsilon_{i,u}, \quad i = 1, \dots, N. \tag{49}$$

The proposed modified estimator of the ex-post risk premia in the unbalanced panel case is

$$\hat{\Gamma}_u^* = \begin{bmatrix} \hat{\gamma}_{0,u}^* \\ \hat{\gamma}_{1,u}^* \end{bmatrix} = \left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1} \frac{\hat{X}_u'\bar{R}_u}{N}, \tag{50}$$

where

$$\hat{\Sigma}_{X,u} = \frac{\hat{X}_u'\hat{X}_u}{N}, \qquad \hat{\Lambda}_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}_u^2 \hat{\mathcal{F}}_u \end{bmatrix}, \tag{51}$$

$$\hat{\sigma}_u^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T_i - K - 1}\mathrm{tr}\left(\hat{\epsilon}_{i,u}\hat{\epsilon}_{i,u}'\right)\right), \tag{52}$$

and

$$\hat{\mathcal{F}}_u = \frac{1}{N}\sum_{i=1}^{N}\left(\tilde{F}_{i,u}'\tilde{F}_{i,u}\right)^{-1}. \tag{53}$$

The estimator $\hat{\Gamma}_u^*$ in (50) generalizes the modified estimator of Shanken (1992) to the unbalanced panel case and coincides with the Shanken's estimator when the panel is balanced. Let $\tau = E\left[1/T_i\right]$, $\theta = \text{Var}\left(\frac{J_{it}}{T_i}\right)$, and assume that these moments exist. In addition, define $\Sigma_{X,i} = \begin{bmatrix} 1 & \beta_i' \\ \beta_i & \beta_i\beta_i' \end{bmatrix}$ and let $\Sigma_{F\beta} = \text{plim}\frac{1}{N}\sum_{i=1}^N \beta_i' F'F\,\beta_i\Sigma_{X,i}$, $\mathcal{F}_u = \text{plim}\frac{1}{N}\sum_{i=1}^N \mathcal{P}_{i,u}'\mathcal{P}_{i,u}$, and $Q_{i,u} = \frac{J_i 1_T}{T_i} - \mathcal{P}_{i,u}\gamma_1^P$. Finally, define $Z_{i,u} = \left[\left(Q_{i,u}\otimes\mathcal{P}_{i,u}\right) + \frac{\text{vec}\left(M_{i,u}\right)}{T_i-K-1}\gamma_1^{P\prime}\mathcal{P}_{i,u}'\mathcal{P}_{i,u}\right]$ and $M_{i,u} = \left[I_T - J_iD(D'J_iD)^{-1}D'J_i\right]J_i$. The consistency and asymptotic normality of the proposed estimator are provided in the following theorem.

Theorem 4

(i) Under Assumptions 1–3 and 4 in Appendix A,

$$\hat{\Gamma}_u^* - \Gamma^P = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{54}$$

(ii) Under Assumptions 1–3 (and 4–5 in Appendix A),

$$\sqrt{N}\left(\hat{\Gamma}_u^* - \Gamma^P\right) \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V_u + \Sigma_X^{-1}(W_u+\Theta)\Sigma_X^{-1}\right), \tag{55}$$

where

$$V_u = \sigma^2\left(\tau + \gamma_1^{P\prime}\mathcal{F}_u\gamma_1^P\right)\Sigma_X^{-1}, \tag{56}$$

$$W_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \text{plim}\frac{1}{N}\sum_{i=1}^N Z_{i,u}'U_\epsilon Z_{i,u} \end{bmatrix}, \tag{57}$$

$$\Theta = \theta\Sigma_{F\beta} - \sigma^2\Psi, \tag{58}$$

with

$$\Psi = \begin{bmatrix} 0 & \gamma_1^{P\prime}\mathcal{F}_\gamma \\ \mathcal{F}_\gamma'\gamma_1^P & \mathcal{F}_{\gamma\beta} \end{bmatrix}, \tag{59}$$

$$\mathcal{F}_\gamma = \text{plim}\frac{1}{N}\sum_{i=1}^N \mathcal{P}_{i,u}'\mathcal{P}_{i,u}(\bar{f}_{i,u}-\bar{f})'\beta_i, \tag{60}$$

and

$$\begin{aligned}\mathcal{F}_{\gamma\beta} = {} & \text{plim}\frac{1}{N}\sum_{i=1}^N(\beta_i\beta_i'(\bar{f}_{i,u}-\bar{f})\gamma_1^{P\prime}\mathcal{P}_{i,u}'\mathcal{P}_{i,u} + \mathcal{P}_{i,u}'\mathcal{P}_{i,u}\gamma_1^P(\bar{f}_{i,u}-\bar{f})'\beta_i\beta_i' \\ & -(\bar{f}_{i,u}-\bar{f})'\beta_i\beta_i'(\bar{f}_{i,u}-\bar{f})\mathcal{P}_{i,u}'\mathcal{P}_{i,u}).\end{aligned} \tag{61}$$

Proof: See Online Appendix.

It should be noted that the asymptotic covariance matrix in Theorem 4 is similar to the one for the balanced panel case provided in Theorem 1, and it is still has a sandwich form. The additional terms in part (ii) of Theorem 4 account for the randomness of the sample size T_i. When the panel is balanced, Theorem 4 reduces to Theorem 1 since $T_i = T$, $J_{it} = 1$, $\bar{f}_{i,u} = \bar{f}$, which implies that $\tau = 1/T$, $\theta = 0$, $\Psi = \Theta = 0_{(K+1)\times(K+1)}$, and all the relevant quantities do not depend on i anymore.

In conducting statistical tests, we need a consistent estimator of the asymptotic covariance matrix in Theorem 4(ii). Let $\hat{\tau} = \frac{1}{N}\sum_{i=1}^N \frac{1}{T_i}$, $\hat{\Sigma}_{X,i}^a = \begin{bmatrix} 1 & \hat{\beta}_{i,u}' \\ \hat{\beta}_{i,u} & \hat{\Sigma}_{\hat{\beta}_{i,u}}^a \end{bmatrix}$, where $\hat{\Sigma}_{\hat{\beta}_{i,u}}^a = \hat{\beta}_{i,u}\hat{\beta}_{i,u}' - \hat{\sigma}_u^2 \mathcal{P}_{i,u}'\mathcal{P}_{i,u}$, $\hat{b}_i = \operatorname{tr}(F'F\hat{\Sigma}_{\hat{\beta}_{i,u}}^a)$, and $A_i = \mathcal{P}_{i,u}'\mathcal{P}_{i,u}F'F$. Also, let $\hat{\mathcal{U}}_i = \sum_{t=1}^T (\mathcal{P}_{i,u}' \otimes f_t'\mathcal{P}_{i,u}')\hat{U}_\epsilon(\mathcal{P}_{i,u} \otimes \mathcal{P}_{i,u} f_t)$, where $\hat{U}_\epsilon$ (as in the balanced panel case) is a plug-in estimator of U_ϵ that depends only on $\hat{\sigma}_{4,u} = \frac{\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \hat{\epsilon}_{it,u}^4}{3\,\frac{1}{N}\sum_{i=1}^N \operatorname{tr}\left(M_{i,u}^{(2)}\right)}$, with $\hat{\epsilon}_{it,u}$ being the t-th element of $\hat{\epsilon}_{i,u}$ and $M_{i,u}^{(2)} = M_{i,u} \odot M_{i,u}$. Finally, let $\hat{\Sigma}_{F\beta} = \frac{1}{N}\sum_{i=1}^N \hat{b}_i\hat{\Sigma}_{X,i}^a - \hat{\Upsilon}$, where $\hat{\Upsilon} = \frac{1}{N}\sum_{i=1}^N \begin{bmatrix} 0 & 2\hat{\sigma}_u^2\hat{\beta}_{i,u}'A_i' \\ 2\hat{\sigma}_u^2 A_i\hat{\beta}_{i,u} & 2\hat{\sigma}_u^2(A_i\hat{\Sigma}_{\hat{\beta}_{i,u}}^a + \hat{\Sigma}_{\hat{\beta}_{i,u}}^a A_i') + \hat{\mathcal{U}}_i \end{bmatrix}$, $\hat{\theta} = \frac{1}{NT}\sum_{t=1}^T\sum_{i=1}^N \frac{J_{it}}{T_i^2} - \frac{1}{T^2}$, and $\hat{Z}_{i,u} = \left[\left(\hat{Q}_{i,u} \otimes \mathcal{P}_{i,u}\right) + \frac{\operatorname{vec}(M_{i,u})}{T_i - K - 1}\hat{\gamma}_{1,u}^{*\prime}\mathcal{P}_{i,u}'\mathcal{P}_{i,u}\right]$, where $\hat{Q}_{i,u} = \frac{J_i 1_T}{T_i} - \mathcal{P}_{i,u}\hat{\gamma}_{1,u}^*$.

The following theorem provides a consistent estimator of the asymptotic covariance matrix of the estimates.

Theorem 5 *Under Assumptions 1 to 3 and 4-5 in Appendix A, we have*

$$\hat{V}_u + \left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1}(\hat{W}_u + \hat{\Theta})\left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1} \xrightarrow{p} V_u + \Sigma_X^{-1}(W_u + \Theta)\Sigma_X^{-1}, \tag{62}$$

where

$$\hat{V}_u = \left[\hat{\sigma}_u^2\left(\hat{\tau} + \hat{\gamma}_{1,u}^{*\prime}\hat{\mathcal{F}}_u\hat{\gamma}_{1,u}^*\right)\right]\left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1}, \tag{63}$$

$$\hat{W}_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \frac{1}{N}\sum_{i=1}^N \hat{Z}_{i,u}'\hat{U}_\epsilon\hat{Z}_{i,u} \end{bmatrix}, \tag{64}$$

$$\hat{\Theta} = \hat{\theta}\hat{\Sigma}_{F\beta} - \hat{\sigma}_u^2\hat{\Psi}, \tag{65}$$

with

$$\hat{\Psi} = \begin{bmatrix} 0 & \hat{\gamma}_{1,u}^{*\prime}\hat{\mathcal{F}}_\gamma \\ \hat{\mathcal{F}}_\gamma'\hat{\gamma}_{1,u}^* & \hat{\mathcal{F}}_{\gamma\beta} \end{bmatrix}, \tag{66}$$

$$\hat{\mathcal{F}}_{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{P}_{i,u}'\mathcal{P}_{i,u}(\bar{f}_{i,u}-\bar{f})'\hat{\beta}_{i,u}, \tag{67}$$

$$\begin{aligned}\hat{\mathcal{F}}_{\gamma\beta} \;=\; & \frac{1}{N}\sum_{i=1}^{N}\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\hat{\gamma}^{*\prime}_{1,u}\mathcal{P}_{i,u}'\mathcal{P}_{i,u} \\ & +\frac{1}{N}\sum_{i=1}^{N}\mathcal{P}_{i,u}'\mathcal{P}_{i,u}\hat{\gamma}^{*}_{1,u}(\bar{f}_{i,u}-\bar{f})'\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}} \\ & -\frac{1}{N}\sum_{i=1}^{N}(\bar{f}_{i,u}-\bar{f})'\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\mathcal{P}_{i,u}'\mathcal{P}_{i,u}. \end{aligned} \tag{68}$$

Proof: See Online Appendix.

Turning to the specification test analysis, let

$$\hat{e}^{P}_{u} = \bar{R}_{u} - \hat{X}_{u}\hat{\Gamma}^{*}_{u} \tag{69}$$

be the N-vector of ex-post sample pricing errors. Define $\hat{\mathcal{Q}}_u = \frac{\hat{e}^{P\prime}_{u}\hat{e}^{P}_{u}}{N}$ as the sum of squared ex-post sample pricing errors and denote by $\hat{\Sigma}^{a}_{\hat{\beta}_u} = \left(\frac{\hat{B}_u'\hat{B}_u}{N} - \hat{\sigma}^2_u\hat{\mathcal{F}}_u\right)$, $\hat{b} = \mathrm{tr}(F'F\hat{\Sigma}^{a}_{\hat{\beta}_u})$, $\omega_N = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{J_{it}}{T_i}-\frac{1}{T}\right)^2 \mathrm{tr}\left(\mathcal{P}_{i,u}f_tf_t'\mathcal{P}_{i,u}'\right)$, and $Z_{Q_{i,u}} = \left[\left(Q_{i,u}'\otimes Q_{i,u}'\right) - \frac{Q_{i,u}'Q_{i,u}\mathrm{vec}(M_{i,u})'}{T_i-K-1}\right]'$. Finally, consider the centered statistic

$$\mathcal{S}_u = \sqrt{N}\left(\hat{\mathcal{Q}}_u - \hat{\sigma}^2_u(\hat{\tau}+\hat{\gamma}^{*\prime}_{1,u}\hat{\mathcal{F}}_u\hat{\gamma}^{*}_{1,u}) - \hat{\theta}\,\hat{b}\right). \tag{70}$$

Theorem 6 *Under Assumptions 1 to 3 and 4–5 in Appendix A and under $H_0: \; e_i = 0$ for all i, we have*

$$\mathcal{S}_u \xrightarrow{d} \mathcal{N}\left(0, \mathcal{V}_u + \mathcal{W}_u\right), \tag{71}$$

where

$$\mathcal{V}_u = \mathrm{plim}\frac{1}{N}\sum_{i=1}^{N}\tilde{Z}_{Q_{i,u}}'U_{\epsilon}\tilde{Z}_{Q_{i,u}}, \tag{72}$$

$$\mathcal{W}_u = 4\sigma^2\mathrm{plim}\frac{1}{N}\sum_{i=1}^{N}W_i'W_i \tag{73}$$

with

$$\tilde{Z}_{Q_{i,u}} = Z_{Q_{i,u}} + \left(\omega_N\left(\frac{\mathrm{vec}(M_{i,u})}{T_i-K-1}\right) - \sum_{t=1}^{T}\left(\frac{J_{it}}{T_i}-\frac{1}{T}\right)^2 \mathrm{vec}\left(\mathcal{P}_{i,u}f_tf_t'\mathcal{P}_{i,u}'\right)\right) \tag{74}$$

and

$$W_i = \left[(\gamma_{1i,u}^P - \gamma_1^P)' \beta_i Q_{i,u}' - \sum_{t=1}^{T} \left(\frac{J_{it}}{T_i} - \frac{1}{T} \right)^2 \beta_i' f_t f_t' \mathcal{P}_{i,u}' \right]'. \quad (75)$$

Note that when the panel is balanced, Theorem 6 reduces to Theorem 3 since $\frac{J_{it}}{T_i} = \frac{1}{T}$ and $\bar{f}_{i,u} = \bar{f}$, which implies that $\mathcal{W}_u = 0$, $Q_{i,u} = Q$, and $\tilde{Z}_{Q_{i,u}} = Z_{Q_{i,u}} = Z_Q$.

This variance can be consistently estimated. Let $\hat{Z}_{Q_{i,u}} = \left[\left(\hat{Q}_{i,u}' \otimes \hat{Q}_{i,u}' \right) - \frac{\hat{Q}_{i,u}' \hat{Q}_{i,u} \mathrm{vec}(M_{i,u})'}{T_i - K - 1} \right]'$ and $\hat{\tilde{Z}}_{Q_{i,u}} = \hat{Z}_{Q_{i,u}} + \left(\omega_N \left(\frac{\mathrm{vec}(M_{i,u})}{T_i - K - 1} \right) - \sum_{t=1}^{T} \left(\frac{J_{it}}{T_i} - \frac{1}{T} \right)^2 \mathrm{vec} \left(\mathcal{P}_{i,u} f_t f_t' \mathcal{P}_{i,u}' \right) \right)$. Then,

$$\hat{\mathcal{V}}_u = \frac{1}{N} \sum_{i=1}^{N} \hat{\tilde{Z}}_{Q_{i,u}}' \hat{U}_{\epsilon,u} \hat{\tilde{Z}}_{Q_{i,u}} \quad (76)$$

and

$$\begin{aligned} \hat{\mathcal{W}}_u \;=\; & 4\hat{\sigma}_u^2 \frac{1}{N} \sum_{i=1}^{N} \Bigg(\hat{Q}_{i,u}' \hat{Q}_{i,u} (\bar{f}_{i,u} - \bar{f})' \hat{\Sigma}_{\hat{\beta}_{i,u}}^a (\bar{f}_{i,u} - \bar{f}) \\ & + \sum_{t=1}^{T} \left(\frac{J_{it}}{T_i} - \frac{1}{T} \right)^4 \mathrm{tr} \left(f_t f_t' \mathcal{P}_{i,u}' \mathcal{P}_{i,u} f_t f_t' \hat{\Sigma}_{\hat{\beta}_{i,u}}^a \right) \\ & -2 \, \hat{Q}_{i,u}' \mathcal{P}_{i,u} \sum_{t=1}^{T} \left(\frac{J_{it}}{T_i} - \frac{1}{T} \right)^2 f_t f_t' \hat{\Sigma}_{\hat{\beta}_{i,u}}^a (\bar{f}_{i,u} - \bar{f}) \Bigg). \end{aligned} \quad (77)$$

Monte Carlo simulations (not reported to conserve space) show that the parameter and specification tests based on Theorems 5 and 6 have excellent size and power properties even when the number of missing observations is 30-40% of the entire sample.

3 Simulation Evidence

In this section, we undertake a Monte Carlo simulation experiment to study the empirical rejection rates of the specification and t-tests for the OLS bias-adjusted estimator of Shanken (1992). The return generating process under the null of a correctly specified beta-pricing model is given by

$$R_t = \gamma_0 1_N + B \left(\gamma_1 + f_t - E[f] \right) + \epsilon_t, \quad (78)$$

where $\epsilon_t \sim \mathcal{N}(0,\ \Sigma)$. To study the power of the specification test, we generate the returns on the test assets as in (4), that is, we do not impose the beta-pricing restriction.

In all of our simulation experiments, we consider balanced panels with time-series dimension of $T = 36$ and $T = 72$ observations. Specifically, f_t in (78) and in (4) is the excess market return (from Kenneth French's website) from January 2008 to December 2010 for $T = 36$, and the excess market return from January 2008 to December 2013 for $T = 72$. In our simulation designs, the factor realizations are taken as given and kept fixed throughout. This is consistent with the fact that our analysis of the ex-post risk premia is conditional on the realizations of the factors. In addition, $E[f]$ in (78) is set equal to the time-series mean of f_t over the 2008-2010 sample when performing the analysis for $T = 36$ and to the time-series mean of f_t over the 2008-2013 sample when performing the analysis for $T = 72$. To obtain representative values for the parameters γ_0, γ_1, B, and Σ in (78) and (4), we employ a sample of 3000 stocks from CRSP in addition to the excess market return.[6] Based on this balanced panel of 3000 stock returns and the excess market return, for each time-series sample size, we compute the OLS estimates of B, γ_0, and γ_1. Then, we set the B, γ_0, and γ_1 parameters in (78) and in (4) equal to these OLS estimates. The calibration of Σ is a more delicate task and is described in the next subsection. In the simulations, we consider cross-sections of $N = 100$, 500, 1000, and 3000 stocks. All results are based on 10,000 Monte Carlo replications. Our econometric approach, designed for large N and fixed T, should be able to handle this large number of assets over relative short time spans. The rejection rates of the various tests are computed using our asymptotic results in Section 2.

3.1 Percentage Errors and Root Mean Squared Errors of the Estimates

We start from the case in which Σ is a scalar matrix, that is, $\Sigma = \sigma^2 I_T$. In the simulations, we set σ^2 equal to the cross-sectional average (over the 3000 stocks) of the σ_i^2's estimated from the data. Table I reports the percentage error (bias) and root mean squared error (RMSE), all in percent, of the OLS estimator and of the OLS bias-adjusted estimator of Shanken (1992). Panels A and B are for $T = 36$ and $T = 72$, respectively.

Table I about here

[6] Specifically, we download monthly stock data from January 2008 to December 2013 from the CRSP database and apply two filters in the selection of stocks. First, we require that a stock has a Standard Industry Classification (SIC) code (we adopt the 49 industry classifications listed on Kenneth French's website). Second, we keep a stock in our sample only for the months in which its price is at least 3 dollars. The resulting dataset consists of 3065 individual stocks and we randomly select 3000 of them.

Panel A clearly shows that the bias of the OLS estimator is substantial. For $\hat{\gamma}_0$, the bias ranges from 28.8% for $N = 100$ to 22.9% for $N = 3000$, while for $\hat{\gamma}_1$ the bias ranges from -24.8% for $N = 100$ to -17.8% for $N = 3000$. For $\hat{\Gamma}^*$, the bias is small for $N = 100$ (-2.3% for $\hat{\gamma}_0^*$ and 1.8% for $\hat{\gamma}_1^*$) and becomes negligible for $N \geq 500$. As for the RMSE, the typical bias-variance trade-off emerges up to $N = 500$, with the OLS estimator exhibiting a smaller RMSE than the OLS bias-adjusted estimator. When $N > 500$, the RMSE of the OLS bias-adjusted estimator becomes substantially smaller than the one of the OLS estimator. Panel B for $T = 72$ conveys a similar message. As expected from the theoretical analysis, the larger time-series dimension helps in reducing the bias and RMSE associated with the OLS estimator. However, the bias for the OLS estimator is still substantial and ranges from -18.5% for $N = 100$ to -11.7% for $N = 3000$. For the bias-adjusted estimator, the bias becomes negligible even for $N = 100$ when $T = 72$.

Next, we consider the case in which the Σ matrix is either diagonal or full. As emphasized above, our theoretical results hinge upon the assumption that the model disturbances are weakly cross-sectionally correlated. In order to generate shocks under a weak factor structure, we consider the following data generating process (DGP). Define

$$\epsilon^{(1)} = \eta \left(\frac{\sqrt{\theta}}{N^{\delta}} \right) c' + \sqrt{1 - \theta} Z, \tag{79}$$

where η and c are T and N-vectors of i.i.d. standard normal random variables, respectively, Z is a $(T \times N)$ matrix of i.i.d. standard normal random variables, $0 \leq \theta \leq 1$ is a shrinkage parameter that controls the weight assigned to the diagonal and extra-diagonal elements of Σ, and δ is a parameter that controls the strength of the cross-sectional dependence of the shocks (the bigger is δ, the weaker is the dependence). Our $T \times N$ matrix of shocks is then generated as

$$\epsilon = \epsilon^{(1)} \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{bmatrix}^{0.5} \begin{bmatrix} \frac{\theta}{N^{2\delta}} c_1^2 + (1-\theta) & & & \\ & \frac{\theta}{N^{2\delta}} c_2^2 + (1-\theta) & & \\ & & \ddots & \\ & & & \frac{\theta}{N^{2\delta}} c_N^2 + (1-\theta) \end{bmatrix}^{-0.5}, \tag{80}$$

where c_i is the i-th element of c. Given this specification for the shocks, for our theoretical results to hold we require $\delta > 0.25$.

In Table II, we report results for the diagonal case, that is, we set $\theta = 0$ in the above DGP. To obtain representative values of the shock variances, while accounting for the fact that $\hat{\Sigma}$ is ill-conditioned when T is small and N is large, we first estimate the residual variances from the historical data. Then, at each Monte Carlo iteration, we generate a string of Beta(p, q)-distributed random variables with the p and q parameters calibrated to the cross-sectional mean and variance of the $\hat{\sigma}_i^2$'s. This resampling procedure is used to minimize the impact of an ill-conditioned $\hat{\Sigma}$ on the simulation results.

Table II about here

Overall, we find that the OLS estimator exhibits a slightly higher bias compared to the scalar Σ case. The OLS bias-adjusted estimator continues to perform very well in terms of bias for all the time-series and cross-sectional dimensions considered. The RMSEs of both estimators are now a bit higher than in the scalar Σ case, and the OLS bias-adjusted estimator still outperforms the OLS estimator for $N \geq 500$.

Finally, in Tables III and IV, we allow for weak cross-sectional dependence of the model disturbances by setting $\theta = 0.5$ in the above DGP.

Tables III and IV about here

In Table III, we consider the situation in which δ, the parameter that regulates the strength of the cross-sectional dependence, is equal to 0.5. Consistent with our theoretical results, the bias-adjusted estimator continues to perform very well in this scenario. Setting $\delta = 0.25$ in Table IV has only a modest effect on the bias and RMSEs of the two estimators. Overall, the first 4 tables reveal a superiority of the bias-adjusted estimator of Shanken (1992) over the OLS estimator, not only in terms of bias, but also in terms of RMSE when $N > 500$. Furthermore, the bias-adjusted estimator shows little sensitivity to changes in the length of the time series, consistent with the idea that this estimator should perform well for any given T.

3.2 Rejection Rates of the t-tests

In Tables V through VIII, we consider the empirical rejection rates of centered t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of time-series and cross-sectional observations using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with covariance matrix calibrated as in Tables I through IV. The t-statistics are compared with the critical values from a standard normal distribution. We consider three t-statistics. For the OLS estimator of the ex-post risk premia, the first t-statistic is the one that uses the traditional Fama-MacBeth standard error (t_{FM}), while the second t-statistic is the one that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992, t_{EIV}). Both of these t-statistics were developed in a large T and fixed N framework. We report them here to determine how misleading inference can be when using these t-statistics in a large N and fixed T setup. Finally, the third t-statistic is the one associated with the OLS bias-adjusted estimator and is based on the asymptotic distribution in part (ii) of our Theorem 1.

Table V about here

Starting from the scalar Σ case, Table V shows that the t-statistics associated with the OLS estimator only slightly overreject the null hypothesis for $N = 100$. However, as N increases, the performance of these t-statistics substantially deteriorates. For example, when $N = 3000$, the rejections rate of the Fama-MacBeth t-statistic associated with $\hat{\gamma}_1$ is either 41.6% for $T = 36$ or 33.3% for $T = 72$ at the 5% nominal level. The strong size distortions of the Fama-MacBeth t-test don't show any improvement when accounting for the EIV bias due to the estimation of the betas in the first stage. In contrast our proposed t-statistic, based on Theorems 1 and 2, performs extremely well for all T and N. A similar picture emerges in the Σ full case (Tables VI and VII), with the rejection rates of our proposed t-test being always aligned with the critical values from a standard normal distribution.

Tables VI and VII about here

In Table VIII, we increase the strength of the cross-sectional dependence of the residuals by setting $\delta = 0.25$.

Table VIII about here

In this situation, we start to notice some slight overrejections for the t-test associated with the OLS bias-adjusted estimator. For example, when $T = 36$ and $N = 3000$, the rejection rate for the t-test associated with $\hat{\gamma}_1^*$ is 6.8% at the 5% level, and when $T = 72$ and $N = 3000$, the rejection rate for the t-test associated with $\hat{\gamma}_1^*$ is 5.8% at the 5% level. Overall, these results suggest that our proposed t-test is relatively well behaved even when moving towards a fairly strong factor structure in the residuals. Furthermore, using the standard tools that were developed in a large T and fixed N framework can lead to strong overrejections of the null hypothesis, with the likely consequence that a factor will be found to be priced even when it does not help explain the cross-sectional variation in individual stock returns.

3.3 Rejection Rates of the Specification Test

In Tables IX and X, we investigate the size and power properties of our specification test based on the results in Theorem 3. Table IX is for $T = 36$, while Table X is for $T = 72$.

Tables IX and X about here

Since the specification test has a standard normal distribution, we consider two-sided p-values in the computation of the rejection rates. The results in the two tables suggest that the rejection rates of our test under the null that the model is correctly specified are excellent for the scalar and diagonal cases. When simulating with Σ full, the specification test is very well sized when $\delta = 0.5$ but it overrejects a bit too much when $\delta = 0.25$. The power properties of our specification test are fairly good when $N = 100$ and excellent when $N \geq 500$. As expected, power increases when the number of assets becomes large and the rejection rates are similar across time-series sample sizes. Overall, these simulation results suggest that the tests should be fairly reliable for the time-series and cross-sectional dimensions encountered in our empirical work.

4 Empirical Analysis

In this section, we empirically investigate the performance of some prominent beta-pricing specifications using individual stock return data. We consider three linear beta-pricing models: (i) the single-factor CAPM, (ii) the three-factor model of Fama and French (1993, FF3), and (iii) the five-factor model recently proposed by Fama and French (2015, FF5). The five factors entering these empirical specifications are the market excess return (*mkt*), the return difference between portfolios of stocks with small and large market capitalizations (*smb*), the return difference between portfolios of stocks with high and low book-to-market ratios (*hml*), the average return on two robust operating profitability portfolios minus the average return on two weak operating profitability portfolios (*rmw*), and the average return on two conservative investment portfolios minus the average return on two aggressive investment portfolios (*cma*). The factor data is from Kenneth French's website. We use individual stock data from the CRSP database and apply the same two filters described in the simulation section of the paper in the selection of the stocks. The data is monthly, from January 1966 until December 2013. We carry out our empirical analysis over three-year and six-year nonoverlapping periods. After filtering the data, the average number of stocks across the 16 three-year periods is 2329 and across the eight six-year periods is 1684. We choose to conduct our analysis over relatively short time spans to minimize the possibility of structural breaks in the betas that could in turn adversely affect the performance of our tests. In addition, focusing on short time periods renders us less exposed to the potential criticism that our analysis does not allow for time-variation in the betas and/or in the risk premia. In any event, we think that it is interesting to investigate how some prominent beta-pricing specifications perform over short time spans when using individual assets instead of portfolios.

We start by analyzing the performance of the various beta-pricing models over the 16 three-year periods.

Table XI about here

Panel A of Table XI shows that there is widespread evidence of model misspecification for the CAPM. In 11 out of 16 cases, the CAPM seems to be at odds with the data at the 5% nominal level. The CAPM fares relatively well over the 1972-1974 period and over the 80's, but quite poorly in more recent times. The number of model rejections is about the same when considering FF3 and FF5 (10 in both cases at the 5% nominal level). Moreover, all the models are rejected by the data at any conventional nominal level during the periods that surround the two most recent recessions of the US economy. As for the eight six-year periods, both CAPM and FF3 are rejected in seven out of eight cases, while FF5 in five of the eight subperiods at the 5% nominal level. All models continue to fare poorly over the two most recent recessions in US history even when employing longer time spans.

Table XII about here

Our finding of widespread model misspecification is not very surprising given that these three models have been repeatedly rejected in the literature using portfolios of stocks. Next, we examine the pricing results based on the Γ^P OLS bias-adjusted estimator. We first consider pricing in the CAPM over the 16 three-year and eight six-year nonoverlapping periods.

Tables XIII and XIV about here

The market factor in the CAPM is priced in 12 out of 16 cases at the 5% nominal level (and in 14 out of 16 cases at the 10% nominal level) when considering three-year periods, and in six out of eight cases at the 5% nominal level when considering six-year periods. The performance of the market factor somewhat decreases when considering FF3. We now have evidence of pricing in 6 out of 16 cases at the 5% nominal level (and in 7 out of 16 cases at the 10% nominal level) when considering three-year periods, and in five out of eight cases at the 5% nominal level when considering six-year periods. The performance of the market factor further decreases in FF5. Over the longer time spans, the parameter estimate associated with the market factor is statistically significant in three out of the eight cases at the 5% level. Statistical significance over the shorter time spans is achieved in only two out of the 16 cases at the 5% level. As for *hml* and *smb*, there is fairly strong evidence of pricing when considering FF3, but statistical significance decreases in the five-factor model. Finally, the profitability and investment factors in FF5 are rarely priced. We have some evidence of nonzero risk premia, at the 5% nominal level, over the 1984-1989, 2002-2004, and 2002-2007 periods for *rmw*, and over the 1987-1989, 1990-1995, and 2002-2007 periods for *cma*. Our empirical findings are consistent with the results in Kim and Skoulakis (2014) who find that *hml* is often a priced factor in the cross-section of individual stock returns over short time spans. Somewhat differently from their paper, we find fairly strong pricing ability also for the market and size factors. Moreover, our results complement the empirical findings in Chordia, Goyal, and Shanken (2015) who, unlike us, carry out the analysis over the full sample period 1963-2013. They find little evidence of pricing for the market and value factors and some evidence of pricing for the size, profitability, and investment factors. In contrast, our analysis suggests that the market, size, and value factors are often priced factors in the cross-section of individual equities over relatively short time spans.

5 Conclusion

We analyze the large N and fixed T properties of the OLS bias-adjusted estimator of Shanken (1992) when an underlying beta-pricing model holds exactly. As far as we know, our study is the first to analytically consider the limiting sampling distribution of the modified Shanken's estimator. We also propose a new model specification test that has desirable size and power properties.

Our empirical analysis employs individual monthly stock returns from the CRSP database over nonoverlapping three and six-year periods from 1966 until 2013. The three prominent beta-pricing specifications that we consider are the CAPM, the three-factor model of Fama and French (1993), and the newly proposed five-factor model of Fama and French (2015). We find strong evidence of model misspecification over the various time periods. Although all the models are in general rejected by the data, we do find some convincing pricing ability for some of the factors. In particular, the market factor seems to be strongly priced in the cross-section of individual stocks, followed by the size and value factors of Fama and Franch (1993). Overall, we find little evidence of pricing for the investment and profitability factors proposed by Fama and French (2015).

Looking to the future, asset-pricing models with non-traded factors as well as a different data frequency could, of course, be examined. In terms of the methodology, although our simulation results are encouraging, the finite-sample properties of the test statistics proposed in this paper should be explored further. In addition, since our empirical results point to substantial model misspecification, it would be useful to derive misspecification-robust standard errors of the parameter estimates in a large N framework. Model comparison under correctly specified and potentially misspecified models would also be an interesting research avenue. Finally, the impact of spurious factors (that is, factors that exhibit very small correlations with the returns on the test assets) on statistical inference could be examined.

Appendix A: Additional Assumptions and Lemmas

All the limits are taken for $N \to \infty$. In addition, the expectation operator used throughout this appendix has to be understood as conditional of F.

Assumption 4. (Idiosyncratic component) We require

(i)

$$\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2 - \sigma^2\right) = o\left(\frac{1}{\sqrt{N}}\right) \tag{A.1}$$

with $0 < \sigma^2 < \infty$.

(ii)

$$\sup_j \sum_{i=1}^{N} \mid \sigma_{ij} \mid \mathbb{1}_{\{i \neq j\}} = o\left(1\right), \tag{A.2}$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function.

(iii)

$$\frac{1}{N}\sum_{i=1}^{N}\mu_{4i} \to \mu_4 \tag{A.3}$$

with $0 < \mu_4 < \infty$ and $\mu_{4i} = E[\epsilon_{it}^4]$.

(iv)

$$\frac{1}{N}\sum_{i=1}^{N}\sigma_i^4 \to \sigma_4 \tag{A.4}$$

with $0 < \sigma_4 < \infty$.

(v)

$$\sup_i E[\epsilon_{it}^4] \leq C < \infty \tag{A.5}$$

for a generic constant C.

(vi)

$$E[\epsilon_{it}^3] = 0. \tag{A.6}$$

(vii)

$$\frac{1}{N}\sum_{i=1}^{N}\kappa_{4,iiii} \to \kappa_4 \tag{A.7}$$

with $0 < |\kappa_4| < \infty$, where $\kappa_{4,iiii} = \kappa_4(\epsilon_{it}, \epsilon_{it}, \epsilon_{it}, \epsilon_{it})$ denotes the fourth-order cumulant of the random variables $\{\epsilon_{it}, \epsilon_{it}, \epsilon_{it}, \epsilon_{it}\}$.

(viii) For every $3 \leq h \leq 8$, all the mixed cumulants of order h are such that

$$\sup_{i_1} \sum_{i_2,\ldots,i_h=1}^{N} |\kappa_{h,i_1 i_2 \ldots i_h}| = o\left(N\right) \tag{A.8}$$

for at least one i_j $(2 \leq j \leq h)$ different than i_1.

Assumption 4 essentially describes the cross-sectional behavior of the model disturbances. Assumption 4(i) limits the cross-sectional heterogeneity of the return conditional variances. Assumption 4(ii) implies that the conditional correlation among asset returns is sufficiently weak. In particular, it implies that the maximum eigenvalue of the conditional covariance of asset returns is bounded, which is a fairly common assumption in factor pricing models such as the Arbitrage Pricing Theory (see, for example, Chamberlain and Rothschild, 1983). In Assumption 4(iii), we simply assume the existence of the limit of the conditional fourth moment average across assets. In Assumption 4(iv), the magnitude of σ_4 reflects the degree of cross-sectional heterogeneity of the conditional variance of asset returns. Assumption 4(v) is a bounded fourth moment condition uniform across assets. Assumption 4(vi) is a convenient symmetry assumption but it is not strictly necessary for our results. It could be relaxed at the cost of a more cumbersome notation. Assumption 4(vii) allows for non Gaussianity of asset returns because $|\kappa_4| > 0$. For example, this assumption is satisfied when the marginal distribution of asset returns is a Student t with degrees of freedom greater than four.

Assumption 5.

(i)

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \epsilon_i \xrightarrow{d} \mathcal{N}\left(0_T, \sigma^2 I_T\right). \tag{A.9}$$

(ii)

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathrm{vec}(\epsilon_i \epsilon_i' - \sigma_i^2 I_T) \xrightarrow{d} \mathcal{N}(0_{T^2}, U_\epsilon). \tag{A.10}$$

(iii) For a generic T-vector C_T,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(C_T' \otimes \begin{pmatrix} 1 \\ \beta_i \end{pmatrix}\right) \epsilon_i \xrightarrow{d} \mathcal{N}(0_{K+1}, V_c), \tag{A.11}$$

where $V_c \equiv c\sigma^2\Sigma_X$ and $c = C_T'C_T$. In particular, $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(C_T' \otimes \beta_i\right)\epsilon_i \xrightarrow{d} \mathcal{N}(0_K, V_c^\dagger)$, where $V_c^\dagger \equiv c\sigma^2\Sigma_\beta$.

Lemma 1

(i) Under Assumptions 2 to 4, we have

$$\hat{\sigma}^2 - \sigma^2 = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{A.12}$$

(ii) In addition, under Assumption 5, we have

$$\sqrt{N}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}\left(0,\, u_{\sigma^2}\right). \tag{A.13}$$

Proof

(i) Rewrite $\hat{\sigma}^2 - \sigma^2$ as

$$\begin{aligned}\hat{\sigma}^2 - \sigma^2 &= \left(\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2\right)\\ &= \left(\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right) + o\left(\frac{1}{\sqrt{N}}\right)\end{aligned} \tag{A.14}$$

by Assumption 4(i). Moreover,

$$\begin{aligned}\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 &= \frac{\operatorname{tr}\left(M\epsilon\epsilon'\right)}{N(T-K-1)} - \frac{\operatorname{tr}\left(M\right)}{T-K-1}\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\\ &= \frac{\operatorname{tr}\left(P\left(\sum_{i=1}^{N}\sigma_i^2 I_T - \epsilon\epsilon'\right)\right)}{N\left(T-K-1\right)} + \frac{\operatorname{tr}\left(\epsilon\epsilon'\right) - T\sum_{i=1}^{N}\sigma_i^2}{N(T-K-1)}.\end{aligned} \tag{A.15}$$

As for the second term on the right-hand side of (A.15), we have

$$\begin{aligned}\frac{\operatorname{tr}\left(\epsilon\epsilon'\right) - T\sum_{i=1}^{N}\sigma_i^2}{N(T-K-1)} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\epsilon_{it}^2 - \sigma_i^2\right)}{N(T-K-1)}\\ &= O_p\left(\frac{1}{\sqrt{N}}\frac{\sqrt{T}}{(T-K-1)}\right) = O_p\left(\frac{1}{\sqrt{N}}\right).\end{aligned} \tag{A.16}$$

As for the first term on the right-hand side of (A.15), we have

$$\begin{aligned}\frac{\operatorname{tr}\left(P\left(\sum_{i=1}^{N}\sigma_i^2 I_T - \epsilon\epsilon'\right)\right)}{N\left(T-K-1\right)} &= \frac{\sum_{t=1}^{T}d_t\left(D'D\right)^{-1}D'\left(\sum_{i=1}^{N}\sigma_i^2\iota_t - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)}\\ &= \frac{\sum_{t=1}^{T}p_t\left(\sum_{i=1}^{N}\sigma_i^2\iota_t - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)},\end{aligned} \tag{A.17}$$

where ι_t is a T-vector with 1 in the t-th position and zeros elsewhere, d_t is the t-th row of D, and $p_t = d_t \left(D'D\right)^{-1} D'$. Since (A.17) has zero mean, we only need to consider its variance to determine the rate of convergence. We have

$$
\begin{aligned}
& \operatorname{Var}\left(\frac{\sum_{t=1}^{T} p_t \left(\sum_{i=1}^{N} \sigma_i^2 \iota_t - \sum_{i=1}^{N} \epsilon_i \epsilon_{it}\right)}{N(T-K-1)}\right) \\
= \quad & \frac{1}{N^2(T-K-1)^2} E\left[\sum_{i,j=1}^{N} \sum_{t,s=1}^{T} p_t \left(\sigma_i^2 \iota_t - \epsilon_i \epsilon_{it}\right)\left(\sigma_j^2 \iota_s - \epsilon_j \epsilon_{js}\right)' p_s'\right] \\
= \quad & \frac{1}{N^2(T-K-1)^2} \sum_{i,j=1}^{N} \sum_{t,s=1}^{T} p_t E\left[\left(\sigma_i^2 \iota_t - \epsilon_i \epsilon_{it}\right)\left(\sigma_j^2 \iota_s - \epsilon_j \epsilon_{js}\right)'\right] p_s'. \qquad \text{(A.18)}
\end{aligned}
$$

Moreover, we have

$$
E\left[\left(\sigma_i^2 \iota_t - \epsilon_i \epsilon_{it}\right)\left(\sigma_j^2 \iota_s - \epsilon_j \epsilon_{js}\right)'\right] = E\left[\sigma_i^2 \sigma_j^2 \iota_t \iota_s' + \epsilon_i \epsilon_j' \epsilon_{it} \epsilon_{js} - \sigma_i^2 \iota_t \epsilon_j' \epsilon_{js} - \sigma_j^2 \epsilon_{it} \epsilon_i \iota_s'\right]
$$

$$
= \begin{cases}
\mu_{4i} \iota_t \iota_t' + \sigma_i^4 \left(I_T - 2\iota_t \iota_t'\right) & \text{if} \quad i = j,\ t = s \\
\kappa_{4,iijj} \iota_t \iota_t' + \sigma_{ij}^2 (I_T + \iota_t \iota_t') & \text{if} \quad i \neq j,\ t = s \\
\sigma_i^4 \iota_s \iota_t' & \text{if} \quad i = j,\ t \neq s \\
\sigma_{ij}^2 \iota_s \iota_t' & \text{if} \quad i \neq j,\ t \neq s.
\end{cases} \qquad \text{(A.19)}
$$

It follows that

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{\sum_{t=1}^{T} p_t\left(\sum_{i=1}^{N}\sigma_i^2\iota_t - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)}\right)\\
=\ & \frac{1}{N^2(T-K-1)^2}\sum_{t=1}^{T}\sum_{i=1}^{N} p_t\left(\mu_{4i}\iota_t\iota_t' + \sigma_i^4(I_T - 2\iota_t\iota_t')\right)p_t'\\
&+\frac{1}{N^2(T-K-1)^2}\sum_{t=1}^{T}\sum_{i\neq j} p_t\left(\kappa_{4,iijj}\iota_t\iota_t' + \sigma_{ij}^2\left(I_T + \iota_t\iota_t'\right)\right)p_t'\\
&+\frac{1}{N^2(T-K-1)^2}\sum_{i=1}^{N}\sigma_i^4\sum_{t\neq s} p_t\iota_s\iota_t' p_s'\\
&+\frac{1}{N^2(T-K-1)^2}\sum_{i\neq j}\sigma_{ij}^2\sum_{t\neq s} p_t\iota_s\iota_t' p_s'\\
=\ & O\left(\frac{1}{N}\right) \qquad \text{(A.20)}
\end{aligned}
$$

by Assumptions 4(ii), 4(iii), 4(iv), and 4(viii), which implies that the first term on the right-hand side of (A.15) is $O_p\left(\frac{1}{\sqrt{N}}\right)$. Putting the pieces together concludes the proof of part (i).

(ii) Using Assumption 4(i) and the properties of the vec operator, we can write $\sqrt{N}(\hat{\sigma}^2 - \sigma^2)$ as

$$
\sqrt{N}(\hat{\sigma}^2 - \sigma^2) = \frac{1}{T-K-1}\mathrm{vec}(M)'\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}\left(\epsilon_i\epsilon_i' - \sigma_i^2 I_T\right) + o(1). \qquad \text{(A.21)}
$$

The desired result then follows from using Assumption 5(ii). This concludes the proof of part (ii).

Lemma 2 *Let*

$$
\Lambda = \begin{bmatrix} 0 & 0_K' \\ 0_K & \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}. \qquad \text{(A.22)}
$$

(i) Under Assumptions 1 to 4, we have

$$
\hat{X}'\hat{X} = O_p(N). \qquad \text{(A.23)}
$$

In addition, under Assumption 5, we have

(ii)

$$
\hat{\Sigma}_X \xrightarrow{p} \Sigma_X + \Lambda, \qquad \text{(A.24)}
$$

and

(iii)

$$\frac{(\hat{X}-X)'(\hat{X}-X)}{N} \xrightarrow{p} \Lambda. \tag{A.25}$$

Proof

(i) Consider

$$\hat{X}'\hat{X} = \begin{bmatrix} N & 1_N'\hat{B} \\ \hat{B}'1_N & \hat{B}'\hat{B} \end{bmatrix}. \tag{A.26}$$

Then, we have

$$\hat{B}'1_N = \sum_{i=1}^{N} \hat{\beta}_i = \sum_{i=1}^{N} \beta_i + \mathcal{P}' \sum_{i=1}^{N} \epsilon_i. \tag{A.27}$$

Under Assumptions 2 to 4,

$$\begin{aligned} \text{Var}\left(\sum_{t=1}^{T}\sum_{i=1}^{N} \epsilon_{it}(f_t-\bar{f}) \right) &= \sum_{t,s=1}^{T}\sum_{i,j=1}^{N} (f_t-\bar{f})(f_s-\bar{f})'E[\epsilon_{it}\epsilon_{js}] \\ &\leq \sum_{t=1}^{T}\sum_{i,j=1}^{N} (f_t-\bar{f})(f_t-\bar{f})'|\sigma_{ij}| \\ &= O\left(N\sigma^2 \sum_{t=1}^{T} (f_t-\bar{f})(f_t-\bar{f})' \right) = O\left(NT\right). \end{aligned} \tag{A.28}$$

Using Assumption 1, we have

$$\hat{B}'1_N = O_p\left(N + \left(\tfrac{N}{T}\right)^{\frac{1}{2}} \right) = O_p(N). \tag{A.29}$$

Next, consider

$$\begin{aligned} \hat{B}'\hat{B} &= \sum_{i=1}^{N} \hat{\beta}_i\hat{\beta}_i' \\ &= \sum_{i=1}^{N} \left(\beta_i + \mathcal{P}'\epsilon_i\right)\left({\beta_i}' + {\epsilon_i}'\mathcal{P}\right) \\ &= \sum_{i=1}^{N} \beta_i{\beta_i}' + \mathcal{P}'\left(\sum_{i=1}^{N} \epsilon_i{\epsilon_i}' \right)\mathcal{P} \\ &\quad +\mathcal{P}'\left(\sum_{i=1}^{N} \epsilon_i{\beta_i}' \right) + \left(\sum_{i=1}^{N} \beta_i{\epsilon_i}' \right)\mathcal{P}. \end{aligned} \tag{A.30}$$

By Assumption 1,

$$\sum_{i=1}^{N} \beta_i {\beta_i}' = O(N). \tag{A.31}$$

Using similar arguments as for (A.28),

$$\mathcal{P}' \left(\sum_{i=1}^{N} \epsilon_i {\beta_i}' \right) = O_p \left(\left(\frac{N}{T} \right)^{\frac{1}{2}} \right) \tag{A.32}$$

and

$$\left(\sum_{i=1}^{N} \beta_i {\epsilon_i}' \right) \mathcal{P} = O_p \left(\left(\frac{N}{T} \right)^{\frac{1}{2}} \right). \tag{A.33}$$

For $\mathcal{P}' \left(\sum_{i=1}^{N} \epsilon_i {\epsilon_i}' \right) \mathcal{P}$, consider its central part and take the norm of its expectation. Using Assumptions 2 to 4,

$$\begin{aligned}
& \left\| E \left[\tilde{F}' \left(\sum_{i=1}^{N} \epsilon_i {\epsilon_i}' \right) \tilde{F} \right] \right\| \\
= & \left\| E \left[\sum_{t,s=1}^{T} \sum_{i=1}^{N} (f_t - \bar{f})(f_s - \bar{f})' \epsilon_{it} \epsilon_{is} \right] \right\| \\
\leq & \sum_{t,s=1}^{T} \sum_{i=1}^{N} \|(f_t - \bar{f})(f_s - \bar{f})'\| |E[\epsilon_{it}\epsilon_{is}]| \\
= & \sum_{t=1}^{T} \sum_{i=1}^{N} \left\| \left(f_t - \bar{f}\right) \left(f_t - \bar{f}\right)' \right\| \sigma_i^2 \\
= & O \left(N\sigma^2 \sum_{t=1}^{T} \left\| (f_t - \bar{f})(f_t - \bar{f})' \right\| \right) = O(NT).
\end{aligned} \tag{A.34}$$

Then, we have

$$\mathcal{P}' \left(\sum_{i=1}^{N} \epsilon_i {\epsilon_i}' \right) \mathcal{P} = O_p \left(\frac{N}{T} \right) \tag{A.35}$$

and

$$\hat{B}'\hat{B} = O_p \left(N + \left(\frac{N}{T} \right)^{\frac{1}{2}} + \frac{N}{T} \right) = O_p(N). \tag{A.36}$$

This concludes the proof of part (i).

(ii) Using part (i) and under Assumption 2 to 5, we have

$$N^{-1}\hat{B}'1_N = \frac{1}{N} \sum_{i=1}^{N} \beta_i + O_p \left(\frac{1}{\sqrt{N}} \right) \tag{A.37}$$

and

$$
\begin{aligned}
N^{-1}\hat{B}'\hat{B} &= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i' + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\epsilon_i'\right)\mathcal{P} + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i' + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\epsilon_i' - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 I_T + \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 I_T - \sigma^2 I_T + \sigma^2 I_T\right)\mathcal{P} \\
&\quad + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i' + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\left(\epsilon_i\epsilon_i' - \sigma_i^2 I_T\right)\right)\mathcal{P} + \frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2 - \sigma^2\right)\mathcal{P}'\mathcal{P} + \sigma^2\mathcal{P}'\mathcal{P} \\
&\quad + \mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i' + \sigma^2\mathcal{P}'\mathcal{P} + O_p\left(\frac{1}{\sqrt{N}}\right) + o\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right).
\end{aligned}
\tag{A.38}
$$

Assumption 1 concludes the proof of part (ii).

(iii) Note that

$$
\begin{aligned}
\frac{(\hat{X}-X)'(\hat{X}-X)}{N} &= \frac{1}{N}\begin{bmatrix} 0_N' \\ (\hat{B}-B)' \end{bmatrix}\left[0_N,\ (\hat{B}-B)\right] \\
&= \begin{bmatrix} 0 & 0_K' \\ 0_K & \mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P} \end{bmatrix},
\end{aligned}
\tag{A.39}
$$

where 0_N is an N-vector of zeros. As in part (ii) we can write

$$
\frac{\epsilon\epsilon'}{N} = \frac{1}{N}\sum_{i=1}^{N}\left(\epsilon_i\epsilon_i' - \sigma_i^2 I_T\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2 - \sigma^2\right)\right) I_T + \sigma^2 I_T. \tag{A.40}
$$

Assumptions 4(i) and 5(ii) conclude the proof since

$$
\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P} = \sigma^2\mathcal{P}'\mathcal{P} + O_p\left(\frac{1}{\sqrt{N}}\right) + o\left(\frac{1}{\sqrt{N}}\right). \tag{A.41}
$$

Lemma 3

(i) Under Assumptions 1 to 4, we have

$$
X'\bar{\epsilon} = O_p\left(\sqrt{N}\right). \tag{A.42}
$$

(ii) In addition, under Assumption 5, we have

$$\frac{1}{\sqrt{N}} X' \bar{\epsilon} \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V\right). \tag{A.43}$$

Proof

(i) We have

$$X' \bar{\epsilon} = \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} 1_N' \\ B' \end{bmatrix} \epsilon_t \tag{A.44}$$

and

$$\begin{aligned} \operatorname{Var}\left(\frac{1}{T} \sum_{t=1}^{T} 1_N' \epsilon_t\right) &= \frac{1}{T^2} \sum_{t,s=1}^{T} \sum_{i,j=1}^{N} E[\epsilon_{it} \epsilon_{js}] \\ &\leq \frac{1}{T^2} \sum_{t=1}^{T} \sum_{i,j=1}^{N} |\sigma_{ij}| \\ &= O\left(\frac{NT}{T^2} \sigma^2\right) = O(N). \end{aligned} \tag{A.45}$$

Moreover, using Assumptions 1 and 4(ii),

$$\begin{aligned} \operatorname{Var}\left(\frac{1}{T} \sum_{t=1}^{T} B' \epsilon_t\right) &= \frac{1}{T^2} \sum_{t,s=1}^{T} \sum_{i,j=1}^{N} E[\epsilon_{it} \epsilon_{js}] \beta_i \beta_j' \\ &\leq \frac{1}{T^2} \sum_{t=1}^{T} \sum_{i,j=1}^{N} |\beta_i \beta_j'| |\sigma_{ij}| \\ &= O\left(\frac{NT}{T^2} \sigma^2\right) = O(N). \end{aligned} \tag{A.46}$$

Putting the pieces together, $X' \bar{\epsilon} = O_p\left(\sqrt{N}\right)$. This concludes the proof of part (i).

(ii) We have

$$\begin{aligned} \frac{1}{\sqrt{N}} X' \bar{\epsilon} &= \frac{1}{\sqrt{N}} X' \epsilon' \frac{1_T}{T} \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{1_T'}{T} \otimes \begin{bmatrix} 1 \\ \beta_i \end{bmatrix}\right) \epsilon_i. \end{aligned} \tag{A.47}$$

Assumption 5(iii) concludes the proof of part (ii).

Lemma 4

(i) Under Assumptions 2 to 4, we have

$$(\hat{X}-X)'X\Gamma^P = O_p\left(\sqrt{N}\right). \tag{A.48}$$

(ii) In addition, under Assumption 5, we have

$$\frac{1}{\sqrt{N}}(\hat{X}-X)'X\Gamma^P \xrightarrow{d} \mathcal{N}\left(0_{K+1}, \mathcal{K}\right), \tag{A.49}$$

where

$$\mathcal{K} \equiv \sigma^2 \begin{bmatrix} 0 & 0_K' \\ 0_K & \Gamma^{P\prime}\Sigma_X\Gamma^P(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}. \tag{A.50}$$

Proof

(i) We have

$$(\hat{X}-X)'X\Gamma^P = \begin{bmatrix} 0_N' \\ \mathcal{P}'\epsilon \end{bmatrix} X\Gamma^P. \tag{A.51}$$

Using similar arguments as for (A.28) concludes the proof of part (i).

(ii) Using the properties of the vec operator

$$\begin{aligned} \frac{1}{\sqrt{N}}(\hat{X}-X)'X\Gamma^P &= \frac{1}{\sqrt{N}}\begin{bmatrix} 0 & 0_K' \\ \mathcal{P}'\epsilon 1_N & \mathcal{P}'\epsilon B \end{bmatrix}\begin{bmatrix} \gamma_0 \\ \gamma_1^P \end{bmatrix} \\ &= \frac{1}{\sqrt{N}}\begin{bmatrix} 0 \\ \mathcal{P}'\epsilon X\Gamma^P \end{bmatrix} \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{bmatrix} 0_T' \\ \Gamma^{P\prime}\begin{bmatrix} 1 \\ \beta_i \end{bmatrix} \otimes \mathcal{P}' \end{bmatrix}\epsilon_i. \end{aligned} \tag{A.52}$$

Using Assumption 5(iii) concludes the proof of part (ii).

Lemma 5

(i) Under Assumptions 2 to 4, we have

$$(\hat{X}-X)'\bar{\epsilon} = O_p\left(\sqrt{N}\right). \tag{A.53}$$

(ii) In addition, under Assumption 5, we have

$$\frac{1}{\sqrt{N}}(\hat{X}-X)'\bar{\epsilon} \xrightarrow{d} \mathcal{N}\left(0_{K+1}, \mathcal{W}\right). \tag{A.54}$$

Proof

(i)

$$
\begin{aligned}
(\hat{X}-X)'\bar{\epsilon} &= \begin{bmatrix} 0 \\ \mathcal{P}'\epsilon\bar{\epsilon} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{P}'\epsilon\epsilon'\frac{1_T}{T} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ \mathcal{P}'\left[\left(\epsilon\epsilon' - \sum_{i=1}^N \sigma_i^2 I_T\right) + \left(\sum_{i=1}^N \sigma_i^2 - N\sigma^2\right) I_T\right]\frac{1_T}{T} \end{bmatrix} = O_p(\sqrt{N})
\end{aligned}
\tag{A.55}
$$

by Assumption 4.

(ii)

$$
\begin{aligned}
\frac{1}{\sqrt{N}}(\hat{X}-X)'\bar{\epsilon} &= \frac{1}{\sqrt{N}}\begin{bmatrix} 0 \\ \mathcal{P}'\left[\left(\epsilon\epsilon' - \sum_{i=1}^N \sigma_i^2 I_T\right) + \left(\sum_{i=1}^N \sigma_i^2 - N\sigma^2\right) I_T\right]\frac{1_T}{T} \end{bmatrix} \\
&= \begin{bmatrix} 0'_{T^2} \\ \left(\frac{1'_T}{T}\otimes\mathcal{P}'\right) \end{bmatrix} \frac{1}{\sqrt{N}}\sum_{i=1}^N \operatorname{vec}(\epsilon_i\epsilon_i' - \sigma_i^2 I_T) + o(1).
\end{aligned}
\tag{A.56}
$$

The $o(1)$ term in (A.56) is due to Assumption 4(i). Using Assumption 5(ii) concludes the proof of part (ii).

Lemma 6 *Under Assumption 4 and the identification assumption $\kappa_4 = 0$, we have*

$$
\hat{\sigma}_4 \xrightarrow{p} \sigma_4. \tag{A.57}
$$

Proof

We need to show that (i) $E(\hat{\sigma}_4) \to \sigma_4$ and (ii) $\operatorname{Var}(\hat{\sigma}_4) = O\left(\frac{1}{N}\right)$.

(i) By Assumptions 4(iv), 4(vi), and 4(vii), we have

$$
\begin{aligned}
E\left[\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \hat{\epsilon}_{it}^4\right] &= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N E\left[\hat{\epsilon}_{it}^4\right] \\
&= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \sum_{s_1,s_2,s_3,s_4=1}^T m_{ts_1}m_{ts_2}m_{ts_3}m_{ts_4}E\left[\epsilon_{is_1}\epsilon_{is_2}\epsilon_{is_3}\epsilon_{is_4}\right] \\
&= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \kappa_{4,iiii}\sum_{s=1}^T m_{ts}^4 + 3\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \sigma_i^4\left(\sum_{s=1}^T m_{ts}^2\right)^2 \\
&\to \kappa_4\sum_{t=1}^T\sum_{s=1}^T m_{ts}^4 + 3\sigma_4\sum_{t=1}^T\left(\sum_{s=1}^T m_{ts}^2\right)^2,
\end{aligned}
\tag{A.58}
$$

where $\hat{\epsilon}_{it} = \iota_t' M \epsilon_i$ and $M = [m_{ts}]$ for $t, s = 1, \ldots, T$. Note that

$$
\begin{aligned}
\sum_{s=1}^{T} m_{ts}^2 &= ||m_t||^2 \\
&= i_t' M i_t \\
&= i_t' \left(I_T - D(D'D)^{-1} D' \right) i_t \\
&= 1 - \mathrm{tr}\left(D(D'D)^{-1} D' i_t i_t' \right) \\
&= 1 - \mathrm{tr}\left(P i_t i_t' \right) \\
&= 1 - p_{tt} \\
&= m_{tt}, \qquad \text{(A.59)}
\end{aligned}
$$

where p_{tt} is the (t,t)-element of P. Then, we have

$$
\sum_{t=1}^{T} \left(\sum_{s=1}^{T} m_{ts}^2 \right)^2 = \sum_{t=1}^{T} m_{tt}^2 = \mathrm{tr}\left(M^{(2)} \right). \qquad \text{(A.60)}
$$

By setting $\kappa_4 = 0$, it follows that

$$
E\left[\hat{\sigma}_4 \right] \quad \rightarrow \quad \sigma_4. \qquad \text{(A.61)}
$$

This concludes the proof of part (i).

(ii) As for the variance of $\hat{\sigma}_4$, we have

$$\begin{aligned}
\mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\epsilon}_{it}^4\right) &= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\mathrm{Cov}\left(\hat{\epsilon}_{it}^4,\hat{\epsilon}_{js}^4\right)\\
&= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\sum_{\substack{u_1,u_2,\\u_3,u_4=1}}^{T}\sum_{\substack{v_1,v_2,\\v_3,v_4=1}}^{T} m_{tu_1}m_{tu_2}m_{tu_3}m_{tu_4}m_{sv_1}m_{sv_2}m_{sv_3}m_{sv_4}\\
&\quad\times\mathrm{Cov}\left(\epsilon_{iu_1}\epsilon_{iu_2}\epsilon_{iu_3}\epsilon_{iu_4},\epsilon_{jv_1}\epsilon_{jv_2}\epsilon_{jv_3}\epsilon_{jv_4}\right)\\
&= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\sum_{\substack{u_1,u_2,\\u_3,u_4=1}}^{T}\sum_{\substack{v_1,v_2,\\v_3,v_4=1}}^{T} m_{tu_1}m_{tu_2}m_{tu_3}m_{tu_4}m_{sv_1}m_{sv_2}m_{sv_3}m_{sv_4}\\
&\quad\times\Bigg(\kappa_8\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\right)\\
&\quad+\sum^{(6,2)}\kappa_6\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2}\right)\mathrm{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right)\\
&\quad+\sum^{(4,4)}\kappa_4\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{jv_1},\epsilon_{jv_2}\right)\kappa_4\left(\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_3},\epsilon_{jv_4}\right)\\
&\quad+\sum^{(4,2,2)}\kappa_4\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{jv_1},\epsilon_{jv_2}\right)\mathrm{Cov}\left(\epsilon_{iu_3},\epsilon_{iu_4}\right)\mathrm{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right)\\
&\quad+\sum^{(2,2,2,2)}\mathrm{Cov}\left(\epsilon_{iu_1},\epsilon_{iu_2}\right)\mathrm{Cov}\left(\epsilon_{iu_3},\epsilon_{jv_1}\right)\mathrm{Cov}\left(\epsilon_{iu_4},\epsilon_{jv_2}\right)\mathrm{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right)\Bigg),
\end{aligned} \tag{A.62}$$

where $\kappa_4(\cdot)$, $\kappa_6(\cdot)$, and $\kappa_8(\cdot)$ denote the fourth, sixth, and eighth-order mixed cumulants, respectively. By $\sum^{(\nu_1,\nu_2,\ldots,\nu_k)}$ we denote the sum over all possible partitions of a group of K random variables into k subgroups of size $\nu_1,\nu_2,\ldots,\nu_k$, respectively. As an example, consider $\sum^{(6,2)}$. $\sum^{(6,2)}$ defines the sum over all possible partitions of the group of eight random variables $\{\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\}$ into two subgroups of size six and two, respectively. Moreover, since $E[\epsilon_{it}] = E\left[\epsilon_{it}^3\right] = 0$, we do not need to consider further partitions in the above relation.[7] Then, under Assumptions 4(i), 4(ii), 4(v), and 4(viii), it follows that

$$\mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\epsilon}_{it}^4\right) = O\left(\frac{1}{N}\right) \tag{A.63}$$

and $\mathrm{Var}\left(\hat{\sigma}_4\right) = O\left(\frac{1}{N}\right)$. This concludes the proof of part (ii).

[7] According to the theory on cumulants (Brillinger, 1975), evaluation of $\mathrm{Cov}\left(\epsilon_{iu_1}\epsilon_{iu_2}\epsilon_{iu_3}\epsilon_{iu_4},\epsilon_{jv_1}\epsilon_{jv_2}\epsilon_{jv_3}\epsilon_{jv_4}\right)$ requires to consider the indecomposable partitions of the two sets $\{\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4}\}$, $\{\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\}$, meaning that there must be at least one subset that includes an element of both sets.

Appendix B: Proofs of the Theorems

Proof of Theorem 1

(i) Starting from (16), the modified estimator of Shanken (1992) can be written as

$$\begin{aligned}
\hat{\Gamma}^* &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \frac{\hat{X}'\bar{R}}{N} \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \frac{\hat{X}'}{N} \left[\hat{X}\Gamma^P + \bar{\epsilon} - (\hat{X} - X)\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \left[\frac{\hat{X}'\hat{X}}{N}\Gamma^P + \frac{\hat{X}'}{N}\bar{\epsilon} - \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \left(\frac{\hat{X}'\hat{X}}{N}\right) \left[\Gamma^P + \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1} \frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1} \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right] \\
&= \left[I_{K+1} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1} \hat{\Lambda}\right]^{-1} \left[\Gamma^P + \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1} \frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1} \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right].
\end{aligned} \tag{B.1}$$

Hence:

$$\begin{aligned}
\hat{\Gamma}^* - \Gamma^P &= \left(\frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda}\right)^{-1} \left[\frac{\hat{X}'}{N}\bar{\epsilon} - \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P + \hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \left[\frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'}{N}(\hat{X} - X) - \hat{\Lambda}\right)\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} \left[\frac{\hat{X}'}{N}\bar{\epsilon} - \begin{bmatrix} 1_N' \frac{\epsilon'}{N}\mathcal{P}\gamma_1^P \\ \frac{B'\epsilon'}{N}\mathcal{P}\gamma_1^P + \mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P - \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P \end{bmatrix}\right].
\end{aligned} \tag{B.2}$$

By Lemmas 1(i) and 2(i), $\left(\hat{\Sigma}_X - \hat{\Lambda}\right) = O_p(1)$. In addition, Lemmas 3(i) and 5(i) imply that

$$\begin{aligned}
\frac{\hat{X}'\bar{\epsilon}}{N} &= \frac{1}{N}(\hat{X} - X)'\bar{\epsilon} + \frac{1}{N}X'\bar{\epsilon} \\
&= O_p\left(\frac{1}{\sqrt{N}}\right)
\end{aligned} \tag{B.3}$$

and Assumption 5(i) implies that

$$\mathcal{P}'\sum_{i=1}^{N}\epsilon_i = O_p\left(\sqrt{N}\right). \tag{B.4}$$

Note that

$$\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P - \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P \tag{B.5}$$

can be rewritten as

$$\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 I_T\right)\mathcal{P}\gamma_1^P - \left[(\hat{\sigma}^2 - \sigma^2) - \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2\right)\right](\tilde{F}'\tilde{F})^{-1}\gamma_1^P. \tag{B.6}$$

Assumption 5(ii) implies that

$$\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \frac{\sum_{i=1}^{N}\sigma_i^2}{N}I_T\right)\mathcal{P}\gamma_1^P = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{B.7}$$

Using Lemma 1(i) and Assumption 4(i) concludes the proof of part (i) since $\hat{\sigma}^2 - \sigma^2 = O_p\left(\frac{1}{\sqrt{N}}\right)$ and $\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2 = o\left(\frac{1}{\sqrt{N}}\right)$.

(ii) Starting from (B.2), we have

$$\begin{aligned}
\sqrt{N}(\hat{\Gamma}^* - \Gamma^P) &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\bar{\epsilon}}{\sqrt{N}} - \left(\frac{\hat{X}'}{\sqrt{N}}(\hat{X} - X)\Gamma^P\right) + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\bar{\epsilon}}{\sqrt{N}} - \begin{bmatrix} 1_N' \\ \hat{B}' \end{bmatrix}\left[0_N,\ \frac{\epsilon'\mathcal{P}}{\sqrt{N}}\right]\Gamma^P + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{X'\bar{\epsilon}}{\sqrt{N}} + \frac{1}{\sqrt{N}}\begin{bmatrix} 0_N' \\ \mathcal{P}'\epsilon \end{bmatrix}\frac{\epsilon'1_T}{T} - \frac{1}{\sqrt{N}}\begin{bmatrix} 1_N'\epsilon'\mathcal{P} \\ \hat{B}'\epsilon'\mathcal{P} \end{bmatrix}\gamma_1^P + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\begin{bmatrix} 1_N' \\ B' \end{bmatrix}\frac{\epsilon'1_T}{T\sqrt{N}} + \begin{bmatrix} -1_N'\frac{\epsilon'\mathcal{P}}{\sqrt{N}}\gamma_1^P \\ \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\frac{1_T}{T} - B'\frac{\epsilon'\mathcal{P}}{\sqrt{N}}\gamma_1^P - \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\mathcal{P}\gamma_1^P \end{bmatrix}\right. \\
&\quad \left. + \sqrt{N}\hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\begin{bmatrix} \frac{1_N'}{\sqrt{N}}\epsilon'\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) \\ \frac{B'\epsilon'}{\sqrt{N}}\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) + \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) + \frac{\text{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_1^P \end{bmatrix} \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\begin{bmatrix} \frac{1_N'\epsilon'}{\sqrt{N}}Q \\ \frac{B'\epsilon'}{\sqrt{N}}Q \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}Q + \frac{\text{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_1^P \end{bmatrix}\right] \\
&\equiv \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}(I_1 + I_2).
\end{aligned} \tag{B.8}$$

Using Lemmas 1(i) and 2(ii), we have

$$\left(\hat{\Sigma}_X - \hat{\Lambda}\right) \overset{p}{\to} \left(\begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta + \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix} - \begin{bmatrix} 0 & 0_K' \\ 0_K & \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}\right) = \Sigma_X. \tag{B.9}$$

Consider now the terms I_1 and I_2. Both terms have mean zero and, under Assumption 4(vi), they are asymptotically uncorrelated. Assumptions 1, 4(i), 5(i), and 5(iii) imply that

$$\begin{aligned}
\mathrm{Var}(I_1) &= E\left[\begin{array}{cc} Q'\frac{1}{\sqrt{N}}\sum_{i=1}^N \epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j' Q & Q'\frac{1}{\sqrt{N}}\sum_{i=1}^N \epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j'(Q\otimes\beta_j') \\ \frac{1}{\sqrt{N}}\sum_{i=1}^N (Q'\otimes\beta_i)\epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j' Q & \frac{1}{\sqrt{N}}\sum_{i=1}^N (Q'\otimes\beta_i)\epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j'(Q\otimes\beta_j') \end{array}\right] \\
&= \left[\begin{array}{cc} Q'\frac{1}{N}\sum_{i=1}^N E[\epsilon_i\epsilon_i']Q & Q'\frac{1}{N}\sum_{i=1}^N E[\epsilon_i\epsilon_i'](Q\otimes\beta_i') \\ \frac{1}{N}\sum_{i=1}^N (Q'\otimes\beta_i)E[\epsilon_i\epsilon_i']Q & \frac{1}{N}\sum_{i=1}^N (Q'\otimes\beta_i)E[\epsilon_i\epsilon_i'](Q\otimes\beta_i') \end{array}\right] + o(1) \\
&\to \left[\begin{array}{cc} \sigma^2 Q'Q & \sigma^2 Q'(Q\otimes\mu_\beta') \\ \sigma^2(Q'\otimes\mu_\beta)Q & \sigma^2(Q'Q\otimes\Sigma_\beta) \end{array}\right] \\
&= \sigma^2 Q'Q\Sigma_X = \frac{\sigma^2}{T}\left[1+\gamma_1^{P\prime}\left(\tilde{F}'\tilde{F}/T\right)^{-1}\gamma_1^P\right]\Sigma_X. \qquad \text{(B.10)}
\end{aligned}$$

Next, consider I_2. Since $\mathcal{P}'\frac{1}{\sqrt{N}}\sum_{i=1}^N \sigma_i^2 Q + \frac{1}{T-K-1}\mathrm{tr}\left(M\frac{1}{\sqrt{N}}\sum_{i=1}^N \sigma_i^2\right)\mathcal{P}'\mathcal{P}\gamma_1^P = 0_K$, we have

$$\begin{aligned}
I_2 &= \left[\begin{array}{c} 0 \\ (Q'\otimes P')\mathrm{vec}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i' - \sigma_i^2 I_T)\right) + \frac{1}{T-K-1}\mathrm{tr}\left(M\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\right)\mathcal{P}'\mathcal{P}\gamma_1^P \end{array}\right] \\
&\equiv \left[\begin{array}{c} 0 \\ I_{22} \end{array}\right]. \qquad \text{(B.11)}
\end{aligned}$$

Therefore, $\mathrm{Var}(I_2)$ has the following form:

$$\mathrm{Var}(I_2) = \left[\begin{array}{cc} 0 & 0_K' \\ 0_K & E\left[I_{22}I_{22}'\right] \end{array}\right]. \qquad \text{(B.12)}$$

Under Assumptions 4(i) and 5(ii), we have

$$\begin{aligned}
E\left[I_{22}I_{22}'\right] &= E\left[(Q'\otimes\mathcal{P}')\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^N \mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'(Q\otimes\mathcal{P})\right] \\
&+E\left[(Q'\otimes\mathcal{P}')\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^N \mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right] \\
&+E\left[\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^N \mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'(Q\otimes\mathcal{P})\right] \\
&+E\left[\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^N \mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'\frac{\mathrm{vec}(M)}{T-K-1}\right. \\
&\left.\times\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right] \\
&\to \left[(Q'\otimes\mathcal{P}')+\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\right]U_\epsilon\left[(Q\otimes\mathcal{P})+\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right]. \qquad \text{(B.13)}
\end{aligned}$$

Defining $Z = \left[(Q\otimes\mathcal{P})+\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right]$ concludes the proof of part (ii).

Proof of Theorem 2

By Theorem 1(i), $\hat{\gamma}_1^* \xrightarrow{p} \gamma_1^P$. Lemma 1(i) implies that $\hat{\Lambda}$ is a consistent estimator of Λ. Hence, using Lemma 2(ii), we have that $\left(\hat{\Sigma}_X - \hat{\Lambda}\right) \xrightarrow{p} \Sigma_X$, which implies that $\hat{V} \xrightarrow{p} V$. A consistent estimator of W requires a consistent estimator of the matrix U_ϵ, which can be obtained using Lemma 6. This concludes the proof of Theorem 2.

Proof of Theorem 3

We first establish a simpler, asymptotically equivalent, expression for $\sqrt{N}\left(\frac{\hat{e}^{P\prime}\hat{e}^P}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\right)$. Then, we derive the asymptotic distribution of this approximation. Consider the sample ex-post pricing errors

$$\hat{e}^P = \bar{R} - \hat{X}\hat{\Gamma}^*. \tag{B.14}$$

Starting from $\bar{R} = \hat{X}\Gamma^P + \eta^P$ with $\eta^P = \bar{\epsilon} - (\hat{X} - X)\Gamma^P$, we have

$$\begin{aligned} \hat{e}^P &= \hat{X}\Gamma^P + \bar{\epsilon} - (\hat{X} - X)\Gamma^P - \hat{X}\hat{\Gamma}^* \\ &= \bar{\epsilon} - \hat{X}(\hat{\Gamma}^* - \Gamma^P) - (\hat{X} - X)\Gamma^P. \end{aligned} \tag{B.15}$$

Then,

$$\begin{aligned} \hat{e}^{P\prime}\hat{e}^P &= \bar{\epsilon}'\bar{\epsilon} + \Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)\Gamma^P \\ &\quad -2(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\bar{\epsilon} - 2\Gamma^{P\prime}(\hat{X} - X)'\bar{\epsilon} \\ &\quad +2\Gamma^{P\prime}(\hat{X} - X)'\hat{X}(\hat{\Gamma}^* - \Gamma^P) \\ &\quad +(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\hat{X}(\hat{\Gamma}^* - \Gamma^P). \end{aligned} \tag{B.16}$$

Note that

$$\frac{\bar{\epsilon}'\bar{\epsilon}}{N} = \frac{1}{T^2}1_T'\frac{\epsilon\epsilon'}{N}1_T \xrightarrow{p} \frac{\sigma^2}{T}, \tag{B.17}$$

and, by Lemma 2(iii),

$$\Gamma^{P\prime}\frac{(\hat{X} - X)'(\hat{X} - X)}{N}\Gamma^P = \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \xrightarrow{p} \sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P. \tag{B.18}$$

Using Lemmas 3(i) and 5(i) and Theorem 1, we have

$$\frac{(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\bar{\epsilon}}{N} = \frac{(\hat{\Gamma}^* - \Gamma^P)'(\hat{X} - X)'\bar{\epsilon}}{N} + \frac{(\hat{\Gamma}^* - \Gamma^P)'X'\bar{\epsilon}}{N} = O_p\left(\frac{1}{N}\right) \tag{B.19}$$

and

$$\frac{\Gamma^{P\prime}(\hat{X}-X)'\bar{\epsilon}}{N} \quad = \quad O_p\left(\frac{1}{\sqrt{N}}\right). \tag{B.20}$$

In addition, using Lemmas 2(i), 2(iii), 4(i), and Theorem 1, we have

$$\begin{aligned}\frac{\Gamma^{P\prime}(\hat{X}-X)'\hat{X}(\hat{\Gamma}^{*}-\Gamma^{P})}{N} &= \frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)(\hat{\Gamma}^{*}-\Gamma^{P})}{N}+\frac{\Gamma^{P\prime}(\hat{X}-X)'X(\hat{\Gamma}^{*}-\Gamma^{P})}{N}\\ &= O_p\Big(\frac{1}{\sqrt{N}}\Big)+O_p\left(\frac{1}{N}\right) \qquad (B.21)\end{aligned}$$

and

$$\frac{(\hat{\Gamma}^{*}-\Gamma^{P})'\hat{X}'\hat{X}(\hat{\Gamma}^{*}-\Gamma^{P})}{N} \quad = \quad O_p\left(\frac{1}{N}\right). \tag{B.22}$$

It follows that

$$\frac{\hat{e}^{P\prime}\hat{e}^{P}}{N}\xrightarrow{p}\frac{\sigma^2}{T}+\sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^{P}\equiv\sigma^2 Q'Q. \tag{B.23}$$

Collecting terms and rewriting explicitly only the ones that are $O_p\left(\frac{1}{\sqrt{N}}\right)$, we have

$$\begin{aligned}\frac{\hat{e}^{P\prime}\hat{e}^{P}}{N} &= \frac{\bar{\epsilon}'\bar{\epsilon}}{N} && (B.24)\\ &\quad +\frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)\Gamma^{P}}{N} && (B.25)\\ &\quad -2\,\frac{\Gamma^{P\prime}(\hat{X}-X)'\bar{\epsilon}}{N} && (B.26)\\ &\quad +2\,\frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)(\hat{\Gamma}^{*}-\Gamma^{P})}{N} && (B.27)\\ &\quad +O_p\Big(\frac{1}{N}\Big). && (B.28)\end{aligned}$$

Consider the sum of the three terms in (B.24)-(B.26). Under Assumption 4(i), we have

$$\begin{aligned}&\frac{\bar{\epsilon}'\bar{\epsilon}}{N}+\frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)\Gamma^{P}}{N}-2\frac{\Gamma^{P\prime}(\hat{X}-X)'\bar{\epsilon}}{N}\\ =\;&\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\frac{1_T}{T}+\gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}-2\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}\\ =\;&\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\Big(\frac{1_T}{T}-\mathcal{P}\gamma_1^{P}\Big)-\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}+\gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}\\ =\;&\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}Q-Q'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}\\ =\;&Q'\frac{\epsilon\epsilon'}{N}\frac{1_T}{T}-Q'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^{P}\\ =\;&Q'\frac{\epsilon\epsilon'}{N}Q=Q'\Big(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2 I_T\Big)Q+\sigma^2Q'Q+o\left(\frac{1}{\sqrt{N}}\right), \qquad (B.29)\end{aligned}$$

where the $o\left(\frac{1}{\sqrt{N}}\right)$ term comes from $(\bar{\sigma}^2-\sigma^2)Q'Q$. As for the term in (B.27), define

$$\left(\hat{\Sigma}_X-\hat{\Lambda}\right)^{-1}=\left[\begin{array}{cc}\hat{\Sigma}_{11} & \hat{\Sigma}_{12}\\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22}\end{array}\right], \tag{B.30}$$

where every block of $\left(\hat{\Sigma}_X-\hat{\Lambda}\right)^{-1}$ is $O_p(1)$ by the nonsingularity of Σ_X and Slutsky's theorem. Using the same arguments as for Theorem 2, we have

$$\begin{aligned}
& 2\frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)(\hat{\Gamma}^*-\Gamma^P)}{N} \\
= \; & 2\left[\gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\hat{\Sigma}_{21},\ \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\hat{\Sigma}_{22}\right]\left[\begin{array}{c}\frac{1_N'\epsilon'Q}{N}\\ \frac{B'\epsilon'Q}{N}+Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\end{array}\right] \\
= \; & 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right) \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)+o_p\left(\frac{1}{N}\right) \\
= \; & 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)+o_p\left(\frac{1}{N}\right)+O_p\left(\frac{1}{N}\right),
\end{aligned} \tag{B.31}$$

where the two approximations on the right-hand side of the previous expression refer to

$$\begin{aligned}
& 2(\bar{\sigma}^2-\sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2(\bar{\sigma}^2-\sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2(\bar{\sigma}^2-\sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)=o_p\left(\frac{1}{N}\right)
\end{aligned} \tag{B.32}$$

and

$$\begin{aligned}
& 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)=O_p\left(\frac{1}{N}\right),
\end{aligned} \tag{B.33}$$

respectively. Therefore, we have

$$\begin{aligned}
\frac{\hat{e}^{P\prime}\hat{e}^P}{N} \; = \; & Q'\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)Q+\sigma^2Q'Q \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N}+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N}-\bar{\sigma}^2I_T\right)+O_p\left(\frac{1}{N}\right)+o_p\left(\frac{1}{N}\right)+o\left(\frac{1}{\sqrt{N}}\right).
\end{aligned} \tag{B.34}$$

It follows that

$$\frac{\hat{e}^{P\prime}\hat{e}^P}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q} = Q'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)Q - \left(\hat{\sigma}^2\hat{Q}'\hat{Q} - \sigma^2 Q'Q\right)$$
$$+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N}$$
$$+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\text{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + O_p\left(\frac{1}{N}\right) + o_p\left(\frac{1}{N}\right) + o\left(\frac{1}{\sqrt{N}}\right). \tag{B.35}$$

Note that

$$\hat{\sigma}^2\hat{Q}'\hat{Q} - \sigma^2 Q'Q$$
$$= \frac{1}{T}(\hat{\sigma}^2 - \sigma^2) + \hat{\sigma}^2\hat{\gamma}_1^{*\prime}(\tilde{F}'\tilde{F})^{-1}\hat{\gamma}_1^* - \sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P$$
$$= \frac{1}{T}(\hat{\sigma}^2 - \sigma^2) + (\hat{\sigma}^2 - \sigma^2)\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + 2\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + O_p\left(\frac{1}{N}\right)$$
$$= (\hat{\sigma}^2 - \sigma^2)\left(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right) + 2\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + O_p\left(\frac{1}{N}\right)$$
$$= (\hat{\sigma}^2 - \sigma^2)\left(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right) + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N}$$
$$+2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\text{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + O_p\left(\frac{1}{N}\right) + O_p\left(\frac{1}{N\sqrt{N}}\right), \tag{B.36}$$

where $\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}(\hat{\gamma}_1^* - \gamma_1^P) + 2\,(\hat{\sigma}^2 - \sigma^2)(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P = O_p\left(\frac{1}{N}\right)$ and $(\hat{\sigma}^2 - \sigma^2)(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}(\hat{\gamma}_1^* - \gamma_1^P) = O_p\left(\frac{1}{N\sqrt{N}}\right)$. It follows that

$$\frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}$$
$$= Q'\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big)Q - (\hat{\sigma}^2 - \sigma^2)\Big(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\Big) + O_p\Big(\frac{1}{N\sqrt{N}}\Big) + O_p\Big(\frac{1}{N}\Big) + o\Big(\frac{1}{\sqrt{N}}\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big)$$
$$= \Big[\Big(Q'\otimes Q'\Big) - \frac{Q'Q}{T-K-1}\text{vec}(M)'\Big]\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big)$$
$$= Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big), \tag{B.37}$$

where we have condensed $O_p\Big(\frac{1}{N\sqrt{N}}\Big) + O_p\Big(\frac{1}{N}\Big) + o\Big(\frac{1}{\sqrt{N}}\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big)$ into the single term $o_p\Big(\frac{1}{\sqrt{N}}\Big)$ for simplicity. Hence,

$$\sqrt{N}\Big(\frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\Big) = \sqrt{N}Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p(1), \tag{B.38}$$

implying that the asymptotic distribution of $\sqrt{N}\Big(\frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\Big)$ is equivalent to the asymptotic distribution of $\sqrt{N}Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big)$. Finally, by Assumption 5(ii), we have

$$\sqrt{N}Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) \xrightarrow{d} \mathcal{N}\Big(0, Z_Q'U_\epsilon Z_Q\Big). \tag{B.39}$$

This conclude the proof of Theorem 3.

Appendix C: Form of U_ϵ

Denote by U_ϵ the $T^2 \times T^2$ matrix

$$U_\epsilon = \left[\begin{array}{c:c:c:c:c} U_{11} & \cdots & U_{1t} & \cdots & U_{1T} \\ \hdashline \vdots & \ddots & \vdots & \vdots & \vdots \\ \hdashline U_{t1} & \cdots & U_{tt} & \cdots & U_{tT} \\ \hdashline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hdashline U_{T1} & \cdots & U_{Tt} & \cdots & U_{TT} \end{array}\right]. \tag{C.1}$$

Each block of U_ϵ is a $T \times T$ matrix. The blocks along the main diagonal, denoted by U_{tt}, $t = 1, 2, \ldots, T$, are themselves diagonal matrices with $(\kappa_4 + 2\sigma_4)$ in the (t,t)-th position and σ_4 in the (s,s) position for every $s \neq t$, that is,

$$U_{tt} = \begin{array}{c} \\ \\ \\ \rightarrow \\ t\text{-th row} \\ \\ \\ \end{array} \overset{\begin{array}{c}\downarrow \\ t\text{-th column}\end{array}}{\begin{bmatrix} \sigma_4 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \sigma_4 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & (\kappa_4 + 2\sigma_4) & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & \sigma_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sigma_4 \end{bmatrix}}. \tag{C.2}$$

The blocks outside the main diagonal, denoted by U_{ts}, $s, t = 1, 2, \ldots, T$ with $s \neq t$, are all made of zeros except for the (s,t)-th position that contains σ_4, that is,

$$U_{ts} = \begin{array}{c} \\ \\ \\ \\ \underset{s\text{-th row}}{\rightarrow} \\ \\ \\ \\ \end{array} \overset{\substack{\downarrow \\ t\text{-th column}}}{\begin{bmatrix} 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \sigma_4 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix}}. \tag{C.3}$$

Under Assumption 4 and Lemma 6 in Appendix A, it is easy to show that $\hat{U}_\epsilon$ in Theorem 2 is a consistent plug-in estimator of U_ϵ that only depends on $\hat{\sigma}_4$.

References

Bai, Jushan, and Guofu Zhou, 2015, Fama-MacBeth two-pass regressions: Improving risk premia estimates, *Finance Research Letters*, forthcoming.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Brillinger, David A., 1975, Time Series: Data Analysis and Theory, 3rd ed., Holt, Rinehart & Winston, Toronto.

Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica* 51, 1281–1304.

Chordia, Tarun, Amit Goyal, and Jay Shanken, 2015, Cross-sectional asset pricing with individual stocks: Betas versus characteristics, Working Paper, Emory University.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Gagliardini, Patrik, Elisa Ossola, and Olivier Scaillet, 2015, Time-varying risk premium in large cross-sectional equity datasets, *Econometrica*, forthcoming.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gospodinov, Nikolay, and Cesare Robotti, 2012, Asset pricing theories, models, and tests, in H. Kent Baker, Greg Filbeck, eds.: *Portfolio Theory and Management*, Chapter 3, 46–72, Oxford University Press.

Greene, William, 2008, *Econometric Analysis.* Prentice Hall.

Hou, Kewei, and Robert Kimmel, 2006, On the estimation of risk premia in linear factor models, Working Paper, Ohio State University.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jagannathan, Ravi, and Zhenyu Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance* 53, 1285–1309.

Jagannathan, Ravi, Georgios Skoulakis, and Zhenyu Wang, 2010, The analysis of the cross-section of security returns, in Yacine Ait-Sahalia and Lars Hansen, eds.: *Handbook of Financial Econometrics* (Elsevier, Amsterdam).

Jegadeesh, Narasimhan, and Joonki Noh, 2014, Empirical tests of asset pricing models with individual stocks, Working Paper, Emory University.

Kan, Raymond, and Cesare Robotti, 2012, Evaluation of asset pricing models using two-pass cross-sectional regressions, in Jin-Chuan Duan, James E.Gentle, Wolfgang Haerdle, eds.: *Handbook of Computational Finance*, Chapter 9, 223–251, Springer.

Kan, Raymond, Cesare Robotti, and Jay Shanken, 2013, Pricing model performance and the two-pass cross-sectional regression methodology, *Journal of Finance* 68, 2617–2649.

Kim, Soohun, and Georgios Skoulakis, 2014, Estimating and testing linear factor models using large cross sections: The regression-calibration approach, Working Paper, Georgia Institute of Technology.

Litzenberger, Robert H., and Krishna Ramaswamy, 1979, The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence, *Journal of Financial Economics* 7, 16–195.

Pesaran, Hashem M., and Takashi Yamagata, 2012, Testing CAPM with a large number of assets, Working Paper, Cambridge University.

Pukthuanthong, Kuntara, Richard Roll, and Junbo Wang, 2014, Resolving the errors-in-variables bias in risk premium estimation, Working Paper, California Institute of Technology.

Rubin, Donald B., 1976, Inference and missing data, *Biometrika* 63, 581–592.

Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics* 14, 327–348.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

Shanken, Jay, and Guofu Zhou, 2007, Estimating and testing beta pricing models: Alternative methods and their performance in simulations, *Journal of Financial Economics* 84, 40–86.

Table I
Bias and RMSE of the OLS and OLS Bias-Adjusted Estimators in a One-Factor Model ($\mathbf{\Sigma}$ Scalar)

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	28.8%	26.2%	24.6%	22.9%
Bias($\hat{\gamma}_0^*$)	-2.3%	-0.3%	0.3%	-0.2%
RMSE($\hat{\gamma}_0$)	0.3675	0.1875	0.1427	0.1066
RMSE($\hat{\gamma}_0^*$)	0.4509	0.1892	0.1255	0.0699
Bias($\hat{\gamma}_1$)	-24.8%	-20.0%	-18.8%	-17.8%
Bias($\hat{\gamma}_1^*$)	1.8%	0.1%	-0.2%	0.2%
RMSE($\hat{\gamma}_1$)	0.3539	0.1642	0.1277	0.1000
RMSE($\hat{\gamma}_1^*$)	0.4529	0.1655	0.1098	0.0609
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	11.6%	9.8%	8.7%	7.9%
Bias($\hat{\gamma}_0^*$)	-0.8%	-0.0%	-0.0%	-0.1%
RMSE($\hat{\gamma}_0$)	0.2504	0.1198	0.0877	0.0628
RMSE($\hat{\gamma}_0^*$)	0.2881	0.1165	0.0766	0.0426
Bias($\hat{\gamma}_1$)	-18.5%	-14.1%	-12.4%	-11.7%
Bias($\hat{\gamma}_1^*$)	1.0%	-0.0%	0.2%	0.1%
RMSE($\hat{\gamma}_1$)	0.2437	0.1063	0.0787	0.0597
RMSE($\hat{\gamma}_1^*$)	0.2868	0.1026	0.0674	0.0379

Table II
Bias and RMSE of the OLS and OLS Bias-Adjusted Estimators in a One-Factor Model (Σ Diagonal)

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	30.1%	25.8%	24.8%	23.0%
Bias($\hat{\gamma}_0^*$)	-0.7%	-0.8%	0.4%	-0.1%
RMSE($\hat{\gamma}_0$)	0.4047	0.1976	0.1495	0.1100
RMSE($\hat{\gamma}_0^*$)	0.5027	0.2054	0.1364	0.0763
Bias($\hat{\gamma}_1$)	-25.5%	-19.6%	-18.7%	-17.9%
Bias($\hat{\gamma}_1^*$)	0.9%	0.6%	-0.1%	0.1%
RMSE($\hat{\gamma}_1$)	0.3949	0.1733	0.1339	0.1033
RMSE($\hat{\gamma}_1^*$)	0.5104	0.1815	0.1208	0.0681
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	11.2%	10.0%	8.6%	8.0%
Bias($\hat{\gamma}_0^*$)	-1.2%	0.2%	-0.1%	0.0%
RMSE($\hat{\gamma}_0$)	0.2673	0.1246	0.0899	0.0643
RMSE($\hat{\gamma}_0^*$)	0.3116	0.1223	0.0804	0.0446
Bias($\hat{\gamma}_1$)	-18.1%	-14.3%	-12.3%	-11.8%
Bias($\hat{\gamma}_1^*$)	1.5%	-0.3%	0.3%	-0.0%
RMSE($\hat{\gamma}_1$)	0.2621	0.1112	0.0809	0.0612
RMSE($\hat{\gamma}_1^*$)	0.3120	0.1087	0.0711	0.0400

Table III
Bias and RMSE of the OLS and OLS Bias-Adjusted Estimators in a One-Factor Model (Σ Full, $\delta = 0.5$)

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	28.8%	26.0%	24.6%	22.7%
Bias($\hat{\gamma}_0^*$)	-2.6%	-0.6%	0.3%	-0.4%
RMSE($\hat{\gamma}_0$)	0.4065	0.1960	0.1506	0.1089
RMSE($\hat{\gamma}_0^*$)	0.5081	0.2031	0.1385	0.0760
Bias($\hat{\gamma}_1$)	-24.2%	-19.6%	-18.9%	-17.7%
Bias($\hat{\gamma}_1^*$)	2.7%	0.7%	-0.3%	0.3%
RMSE($\hat{\gamma}_1$)	0.3963	0.1727	0.1352	0.1028
RMSE($\hat{\gamma}_1^*$)	0.5159	0.1806	0.1220	0.0681
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	11.8%	9.4%	8.6%	8.0%
Bias($\hat{\gamma}_0^*$)	-0.5%	-0.5%	-0.1%	-0.0%
RMSE($\hat{\gamma}_0$)	0.2671	0.1227	0.0910	0.0642
RMSE($\hat{\gamma}_0^*$)	0.3099	0.1225	0.0820	0.0447
Bias($\hat{\gamma}_1$)	-19.0%	-13.6%	-12.4%	-11.7%
Bias($\hat{\gamma}_1^*$)	0.5%	0.6%	0.1%	0.1%
RMSE($\hat{\gamma}_1$)	0.2614	0.1104	0.0819	0.0611
RMSE($\hat{\gamma}_1^*$)	0.3096	0.1100	0.0720	0.0405

Table IV

Bias and RMSE of the OLS and OLS Bias-Adjusted Estimators in a One-Factor Model (Σ Full, $\delta = 0.25$)

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	28.8%	26.6%	24.2%	23.5%
Bias($\hat{\gamma}_0^*$)	-2.5%	0.1%	-0.3%	0.5%
RMSE($\hat{\gamma}_0$)	0.4191	0.2053	0.1536	0.1135
RMSE($\hat{\gamma}_0^*$)	0.5254	0.2152	0.1450	0.0809
Bias($\hat{\gamma}_1$)	-24.8%	-19.9%	-18.5%	-18.3%
Bias($\hat{\gamma}_1^*$)	2.0%	0.2%	0.2%	-0.4%
RMSE($\hat{\gamma}_1$)	0.4116	0.1824	0.1380	0.1072
RMSE($\hat{\gamma}_1^*$)	0.5355	0.1935	0.1288	0.0731
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	12.2%	9.7%	8.8%	7.9%
Bias($\hat{\gamma}_0^*$)	-0.1%	-0.2%	0.1%	-0.1%
RMSE($\hat{\gamma}_0$)	0.2795	0.1287	0.0939	0.0645
RMSE($\hat{\gamma}_0^*$)	0.3252	0.1292	0.0853	0.0459
Bias($\hat{\gamma}_1$)	-19.3%	-13.9%	-12.6%	-11.7%
Bias($\hat{\gamma}_1^*$)	0.0%	0.2%	-0.1%	0.2%
RMSE($\hat{\gamma}_1$)	0.2761	0.1155	0.0854	0.0615
RMSE($\hat{\gamma}_1^*$)	0.3279	0.1158	0.0763	0.0416

Table V
Size of *t*-tests in a One-Factor Model (Σ Scalar)

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error, while $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992). Finally, the rejection rates for the t-test associated with the OLS bias-adjusted estimator are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.128	0.074	0.021	0.141	0.078	0.022
500	0.186	0.113	0.040	0.213	0.132	0.047
1000	0.243	0.156	0.059	0.290	0.197	0.075
3000	0.438	0.324	0.153	0.538	0.416	0.219
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.127	0.073	0.020	0.140	0.077	0.022
500	0.185	0.113	0.039	0.211	0.132	0.047
1000	0.243	0.156	0.059	0.289	0.197	0.075
3000	0.437	0.323	0.152	0.537	0.415	0.218
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.097	0.051	0.010	0.100	0.048	0.010
500	0.105	0.053	0.011	0.107	0.055	0.012
1000	0.103	0.052	0.010	0.105	0.054	0.011
3000	0.098	0.051	0.011	0.100	0.049	0.010

Table V (Continued)
Size of t-tests in a One-Factor Model ($\boldsymbol{\Sigma}$ Scalar)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.123	0.063	0.016	0.124	0.066	0.016
500	0.167	0.099	0.030	0.181	0.109	0.033
1000	0.211	0.133	0.041	0.237	0.154	0.053
3000	0.378	0.263	0.109	0.449	0.333	0.150
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.122	0.063	0.015	0.123	0.065	0.016
500	0.166	0.099	0.030	0.181	0.108	0.033
1000	0.210	0.132	0.040	0.236	0.153	0.052
3000	0.377	0.261	0.108	0.448	0.331	0.149
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.096	0.047	0.009	0.100	0.048	0.009
500	0.097	0.049	0.010	0.098	0.049	0.010
1000	0.100	0.047	0.009	0.103	0.048	0.009
3000	0.103	0.054	0.010	0.106	0.054	0.010

Table VI
Size of t-tests in a One-Factor Model (Σ Diagonal)

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error, while $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992). Finally, the rejection rates for the t-test associated with the OLS bias-adjusted estimator are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.122	0.066	0.019	0.125	0.072	0.018
500	0.163	0.104	0.033	0.179	0.112	0.036
1000	0.226	0.141	0.050	0.248	0.166	0.060
3000	0.398	0.292	0.128	0.474	0.362	0.174
		$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$	
100	0.120	0.065	0.018	0.124	0.070	0.017
500	0.163	0.103	0.033	0.179	0.111	0.036
1000	0.225	0.141	0.050	0.247	0.165	0.060
3000	0.397	0.291	0.127	0.473	0.362	0.173
		$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$	
100	0.093	0.045	0.011	0.091	0.044	0.010
500	0.102	0.051	0.010	0.096	0.049	0.011
1000	0.099	0.048	0.009	0.101	0.051	0.009
3000	0.099	0.053	0.012	0.099	0.051	0.010

Table VI (Continued)
Size of t-tests in a One-Factor Model ($\boldsymbol{\Sigma}$ Diagonal)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.115	0.060	0.015	0.121	0.064	0.015
500	0.157	0.089	0.027	0.165	0.096	0.030
1000	0.199	0.121	0.036	0.219	0.137	0.044
3000	0.353	0.250	0.103	0.416	0.302	0.134
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.114	0.059	0.014	0.119	0.063	0.015
500	0.157	0.089	0.027	0.163	0.096	0.029
1000	0.198	0.120	0.036	0.218	0.136	0.044
3000	0.351	0.248	0.102	0.414	0.301	0.132
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.097	0.048	0.010	0.096	0.048	0.007
500	0.095	0.046	0.010	0.093	0.047	0.010
1000	0.097	0.049	0.011	0.095	0.049	0.010
3000	0.103	0.052	0.010	0.102	0.051	0.010

Table VII
Size of t-tests in a One-Factor Model ($\boldsymbol{\Sigma}$ Full, $\boldsymbol{\delta = 0.5}$)

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error, while $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992). Finally, the rejection rates for the t-test associated with the OLS bias-adjusted estimator are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.126	0.069	0.020	0.125	0.070	0.021
500	0.166	0.097	0.030	0.181	0.109	0.034
1000	0.227	0.143	0.049	0.258	0.170	0.063
3000	0.393	0.282	0.123	0.472	0.354	0.168
		$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$	
100	0.124	0.068	0.019	0.123	0.068	0.021
500	0.166	0.096	0.030	0.180	0.109	0.034
1000	0.227	0.142	0.049	0.257	0.170	0.063
3000	0.392	0.281	0.122	0.470	0.353	0.167
		$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$	
100	0.097	0.045	0.012	0.094	0.046	0.011
500	0.094	0.045	0.009	0.095	0.045	0.010
1000	0.106	0.051	0.011	0.102	0.050	0.010
3000	0.100	0.051	0.011	0.100	0.053	0.011

Table VII (Continued)
Size of t-tests in a One-Factor Model (Σ Full, $\delta = 0.5$)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.113	0.062	0.014	0.119	0.061	0.014
500	0.150	0.086	0.025	0.165	0.096	0.029
1000	0.202	0.127	0.041	0.228	0.141	0.047
3000	0.353	0.246	0.102	0.417	0.302	0.137
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.112	0.062	0.014	0.117	0.060	0.014
500	0.149	0.085	0.025	0.164	0.096	0.029
1000	0.201	0.126	0.041	0.227	0.141	0.047
3000	0.352	0.244	0.100	0.415	0.301	0.136
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.094	0.046	0.010	0.091	0.044	0.009
500	0.095	0.047	0.010	0.094	0.050	0.011
1000	0.105	0.052	0.011	0.102	0.052	0.010
3000	0.102	0.052	0.012	0.102	0.053	0.013

Table VIII
Size of t-tests in a One-Factor Model (Σ Full, $\delta = 0.25$)

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error, while $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992). Finally, the rejection rates for the t-test associated with the OLS bias-adjusted estimator are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.125	0.068	0.017	0.124	0.068	0.018
500	0.163	0.095	0.034	0.174	0.109	0.039
1000	0.215	0.131	0.046	0.241	0.155	0.057
3000	0.389	0.280	0.125	0.459	0.343	0.164
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.123	0.067	0.017	0.123	0.067	0.017
500	0.162	0.095	0.033	0.174	0.109	0.039
1000	0.214	0.130	0.046	0.240	0.155	0.057
3000	0.388	0.278	0.124	0.458	0.341	0.163
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.109	0.060	0.015	0.112	0.059	0.015
500	0.115	0.062	0.018	0.117	0.064	0.019
1000	0.122	0.065	0.016	0.119	0.066	0.017
3000	0.121	0.069	0.018	0.124	0.068	0.018

Table VIII (Continued)
Size of t-tests in a One-Factor Model ($\boldsymbol{\Sigma}$ Full, $\delta = 0.25$)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.119	0.060	0.014	0.123	0.066	0.015
500	0.155	0.091	0.025	0.163	0.098	0.030
1000	0.199	0.126	0.042	0.222	0.138	0.050
3000	0.334	0.229	0.092	0.390	0.280	0.124
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.117	0.059	0.014	0.122	0.065	0.015
500	0.155	0.090	0.025	0.162	0.098	0.030
1000	0.198	0.125	0.042	0.222	0.138	0.049
3000	0.333	0.228	0.091	0.388	0.278	0.123
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.108	0.057	0.012	0.110	0.059	0.015
500	0.114	0.062	0.015	0.119	0.065	0.015
1000	0.121	0.063	0.015	0.122	0.067	0.016
3000	0.111	0.057	0.012	0.114	0.058	0.014

Table IX
Rejection Rates of the Specification Test in a One-Factor Model

The table presents the size and power properties of the test of correct model specification. The null hypothesis is that the model is correctly specified. The alternative hypothesis is that the model is misspecified. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2010:12 ($T = 36$). Finally, the rejection rates for the specification test are based on the asymptotic distribution in Theorem 3. The rejection rates of the test are based on two-sided p-values.

	Size			Power		
N	10%	5%	1%	10%	5%	1%
Panel A: Σ Scalar						
100	0.103	0.049	0.009	0.882	0.823	0.675
500	0.098	0.050	0.009	1.000	1.000	0.998
1000	0.101	0.052	0.011	1.000	1.000	1.000
3000	0.101	0.050	0.009	1.000	1.000	1.000
Panel B: Σ Diagonal						
100	0.085	0.037	0.010	0.634	0.529	0.340
500	0.093	0.046	0.010	0.983	0.967	0.894
1000	0.099	0.050	0.009	1.000	1.000	0.996
3000	0.097	0.046	0.011	1.000	1.000	1.000
Panel C: Σ Full ($\delta = 0.5$)						
100	0.084	0.040	0.011	0.639	0.534	0.332
500	0.101	0.050	0.012	0.982	0.965	0.887
1000	0.095	0.049	0.011	1.000	1.000	0.997
3000	0.108	0.056	0.011	1.000	1.000	1.000
Panel D: Σ Full ($\delta = 0.25$)						
100	0.110	0.060	0.021	0.621	0.522	0.336
500	0.145	0.084	0.029	0.977	0.956	0.874
1000	0.145	0.088	0.029	1.000	0.999	0.993
3000	0.146	0.087	0.030	1.000	1.000	1.000

Table X

Rejection Rates of the Specification Test in a One-Factor Model

The table presents the size and power properties of the test of correct model specification. The null hypothesis is that the model is correctly specified. The alternative hypothesis is that the model is misspecified. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12 ($T = 72$). Finally, the rejection rates for the specification test are based on the asymptotic distribution in Theorem 3. The rejection rates of the test are based on two-sided p-values.

	Size			Power		
N	10%	5%	1%	10%	5%	1%
			Panel A: Σ Scalar			
100	0.095	0.045	0.009	0.929	0.891	0.781
500	0.101	0.047	0.009	1.000	1.000	1.000
1000	0.104	0.055	0.010	1.000	1.000	1.000
3000	0.099	0.048	0.010	1.000	1.000	1.000
			Panel B: Σ Diagonal			
100	0.085	0.041	0.010	0.771	0.676	0.480
500	0.098	0.046	0.010	1.000	1.000	0.997
1000	0.101	0.049	0.012	1.000	1.000	1.000
3000	0.102	0.051	0.011	1.000	1.000	1.000
			Panel C: Σ Full ($\delta = 0.5$)			
100	0.085	0.039	0.011	0.770	0.681	0.482
500	0.092	0.046	0.009	1.000	0.999	0.996
1000	0.094	0.049	0.010	1.000	1.000	1.000
3000	0.097	0.047	0.010	1.000	1.000	1.000
			Panel D: Σ Full ($\delta = 0.25$)			
100	0.120	0.063	0.023	0.749	0.660	0.470
500	0.140	0.083	0.029	1.000	0.999	0.994
1000	0.149	0.086	0.030	1.000	1.000	1.000
3000	0.153	0.093	0.034	1.000	1.000	1.000

Table XI
Specification Tests of Various Beta-Pricing Models over Three-Year Periods

The table presents the p-value of the specification test ($p_{\mathcal{S}}$) for CAPM, FF3, and FF5. The null hypothesis is that the model is correctly specified. The results are reported for different nonoverlapping three-year periods from January 1966 until December 2013. We also report the number of stocks (N) in each period. The p-value for the specification test is based on the asymptotic distribution in Theorem 3.

Panel A: CAPM

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
$p_{\mathcal{S}}$	0.001	0.023	0.841	0.000	0.012	0.135	0.213	0.403
N	1139	1236	1024	1183	1319	1309	1579	1973

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
$p_{\mathcal{S}}$	0.007	0.232	0.006	0.000	0.002	0.001	0.000	0.003
N	1961	2908	3377	3180	3397	3875	3647	4153

Panel B: FF3

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
$p_{\mathcal{S}}$	0.001	0.000	0.090	0.000	0.029	0.037	0.085	0.947

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
$p_{\mathcal{S}}$	0.028	0.171	0.012	0.000	0.000	0.446	0.000	0.091

Panel C: FF5

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
$p_{\mathcal{S}}$	0.315	0.000	0.161	0.692	0.131	0.048	0.012	0.040

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
$p_{\mathcal{S}}$	0.043	0.032	0.065	0.042	0.000	0.018	0.003	0.375

Table XII

Specification Tests of Various Beta-Pricing Models over Six-Year Periods

The table presents the p-value of the specification test ($p_{\mathcal{S}}$) for CAPM, FF3, and FF5. The null hypothesis is that the model is correctly specified. The results are reported for different nonoverlapping six-year periods from January 1966 until December 2013. We also report the number of stocks (N) in each period. The p-value for the specification test is based on the asymptotic distribution in Theorem 3.

Panel A: CAPM

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
$p_{\mathcal{S}}$	0.003	0.000	0.000	0.465	0.000	0.000	0.015	0.000
N	865	881	1031	1126	1653	2212	2638	3065

Panel B: FF3

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
$p_{\mathcal{S}}$	0.006	0.000	0.000	0.078	0.000	0.000	0.013	0.000

Panel C: FF5

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
$p_{\mathcal{S}}$	0.271	0.000	0.000	0.065	0.000	0.001	0.209	0.002

Table XIII
t-tests for Various Beta-Pricing Models over Three-Year Periods

The table presents the t-statistics for CAPM, FF3, and FF5. t_x denotes the t-test of statistical significance for the parameter associated with factor x, with standard errors based on the results in Theorems 1 and 2. The null hypothesis is that the factor is not priced. The results are reported for different nonoverlapping three-year periods from January 1966 until December 2013.

Panel A: CAPM

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
t_{mkt}	7.676	-5.504	1.993	14.699	15.961	-1.355	-1.906	5.883

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
t_{mkt}	10.864	1.253	6.399	13.701	-1.655	4.785	6.998	13.342

Panel B: FF3

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
t_{mkt}	1.335	0.367	2.860	4.897	4.328	0.148	1.419	2.902
t_{smb}	11.205	-10.278	-7.743	4.406	3.491	4.188	-4.670	-0.997
t_{hml}	-2.788	-2.991	1.846	1.727	-2.920	2.842	0.922	-6.054

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
t_{mkt}	1.916	0.407	2.262	1.569	-0.397	0.161	2.938	0.988
t_{smb}	2.074	-0.333	-2.628	0.500	4.841	-0.137	1.797	-0.168
t_{hml}	-4.133	-0.402	-3.486	-0.552	3.320	-0.612	-0.213	0.235

Panel C: FF5

	66-68	69-71	72-74	75-77	78-80	81-83	84-86	87-89
t_{mkt}	0.365	1.522	0.518	0.580	0.283	0.152	1.830	2.174
t_{smb}	0.641	-1.853	-1.207	-0.055	0.172	0.211	-1.361	-1.322
t_{hml}	-0.304	-1.420	0.402	0.189	-0.315	0.184	1.617	-3.759
t_{rmw}	-0.262	0.583	-0.292	-0.147	0.221	0.134	1.537	-1.087
t_{cma}	0.100	0.076	0.140	-0.119	-0.166	0.271	-1.803	2.510

	90-92	93-95	96-98	99-01	02-04	05-07	08-10	11-13
t_{mkt}	1.415	0.312	2.028	0.543	0.707	0.053	-0.106	0.250
t_{smb}	1.429	-0.256	-1.363	0.036	4.315	-0.053	-0.050	-0.099
t_{hml}	-2.553	-0.301	-1.594	-0.361	0.305	-0.070	-0.126	-0.187
t_{rmw}	0.236	0.213	0.598	0.014	4.204	0.055	0.127	0.145
t_{cma}	-1.078	0.321	-1.267	0.000	-0.417	-0.069	-0.131	-0.080

Table XIV

***t*-tests for Various Beta-Pricing Models over Six-Year Periods**

The table presents the t-statistics for CAPM, FF3, and FF5. t_x denotes the t-test of statistical significance for the parameter associated with factor x, with standard errors based on the results in Theorems 1 and 2. The null hypothesis is that the factor is not priced. The results are reported for different nonoverlapping six-year periods from January 1966 until December 2013.

Panel A: CAPM

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
t_{mkt}	6.751	-0.075	12.820	-1.355	12.256	11.038	3.973	13.045

Panel B: FF3

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
t_{mkt}	3.305	0.192	1.303	1.953	5.016	4.869	2.032	3.336
t_{smb}	1.600	-0.525	5.716	-5.480	2.175	0.369	4.855	5.078
t_{hml}	-4.346	-0.151	-1.879	0.138	-5.053	-4.104	2.597	0.508

Panel C: FF5

	66-71	72-77	78-83	84-89	90-95	96-01	02-07	08-13
t_{mkt}	0.914	0.035	0.313	1.891	4.566	3.093	3.919	1.366
t_{smb}	0.693	0.396	0.995	-4.501	2.273	-0.074	4.653	1.969
t_{hml}	-0.886	-0.105	-0.348	-0.394	-3.864	-2.521	-2.071	-0.652
t_{rmw}	0.680	-1.431	0.456	2.761	-0.624	-0.115	3.130	0.672
t_{cma}	-0.217	0.104	-0.314	0.960	-5.263	-0.810	-4.155	-0.519