Ex-post Risk Premia:

Estimation and Inference using Large Cross Sections

Soohun Kim[1]
Georgia Institute of Technology

Georgios Skoulakis[2]
University of British Columbia

This draft: May 2015

[1]Finance Department, Scheller College of Business, Georgia Institute of Technology, Atlanta, GA 30308; Email: soohun.kim@scheller.gatech.edu.

[2]Finance Department, Sauder School of Business, University of British Columbia, Vancouver, BC, Canada V6T 1Z2; Email: georgios.skoulakis@sauder.ubc.ca.

Ex-post Risk Premia:
Estimation and Inference using Large Cross Sections

Abstract

We propose a modification of the two-pass cross-sectional regression approach for estimating ex-post risk premia in linear asset pricing models, suitable for the case of large cross sections and short time series. Employing the regression-calibration method, we provide a beta correction method, which deals with the error-in-variables problem, based on which we construct an N-consistent estimator of ex-post risk premia and develop associated novel asset pricing tests. Empirically, we reject the implications of the CAPM and Fama-French model but also offer new evidence on the usefulness of the HML factor for pricing large cross sections of individual stocks.

KEY WORDS: Large cross section, N-consistent ex-post risk premia estimator, asset pricing tests.

1 Introduction

A common theme of asset pricing models is that differences in average returns across assets should be attributable to differences in exposures to systematic risk. There is a plethora of proposed models in the literature that differ in the types of systematic risk they identify as relevant. Typically, in these models, systematic risk is captured by a small number of pervasive factors and the average return on an asset is a linear function of the factor betas. There is a long line of research, starting with Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973), on the empirical evaluation of such models.

In this paper, we develop a framework for estimating and evaluating asset pricing factor models using large cross sections of individual stock return data over short time horizons, when the factors are returns to traded portfolios. When researchers are interested in testing an asset pricing model, they have to specify the cross section of test assets. One approach, introduced by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) and since followed by many others, is to form a small number of portfolios and use them as test assets. In fact, following the seminal work by Fama and French (1992), it has become standard practice to sort stocks according to some characteristic, such as size or book-to-market, in order to form portfolios that are subsequently used as test assets. However, Lewellen, Nagel, and Shanken (2010) and Daniel and Titman (2012) document that the performance of seemingly successful asset pricing models crucially depends on the choice of test assets.[1] This finding provides motivation for developing asset pricing tests using individual stock data, as originally suggested by Litzenberger and Ramaswamy (1979). We only consider short time horizons so that we can evaluate the implications of asset pricing models locally in time, and, hence, we restrict ourselves to factors that are traded portfolio returns or spreads focusing on ex-post risk premia (see Shanken (1992)).

Our method is a variant of the two-pass cross-sectional regression (henceforth CSR) method, which, being simple and intuitively appealing, is a popular approach in the literature.[2] The two-pass CSR method is subject to the error-in-variables (EIV) problem due to the fact that estimated betas, instead of the true betas, are used in the second pass. While the two-pass CSR risk premia estimator is consistent, as the time-series sample size T tends to infinity and the cross section size N is fixed, the traditional

[1] The method used to form the test portfolios could affect the inference results in undesirable ways. As Roll (1977) points out, in the process of forming portfolios, important mispricing in individual stocks can be averaged out within portfolios, making it harder to reject the wrong model. Lo and MacKinlay (1990) are concerned about the exact opposite error: if stocks are grouped into portfolios with respect to attributes already observed to be related to average returns, the correct model may be rejected too often.

[2] Alternative approaches for estimating and testing asset pricing models include the maximum likelihood method and the generalized method of moments. Such methods, however, do not seem suitable for dealing with large cross sections of individual assets over short horizons.

Fama-MacBeth standard errors are not consistent and a suitable asymptotic bias correction is needed. The associated econometric theory, that deals with the aforementioned EIV problem, was originally developed by Shanken (1992) and subsequently refined by Jagannathan and Wang (1998), among others. However, when T is fixed and N increases to infinity, the EIV problem, due to beta estimation error, is more severe in the sense that the ex-post risk premia estimator itself is inconsistent. In this paper, we employ the so-called regression calibration method to provide a suitable correction to the beta estimators yielding an N-consistent ex-post risk premia estimator. We further obtain its asymptotic distribution and provide an N-consistent estimator of its asymptotic variance-covariance matrix, which we employ to construct novel asset pricing test statistics.

The main focus of the extant methodological literature on the estimation and evaluation of asset pricing models is the case in which T is large while N is small, which is relevant when portfolios, as opposed to individual stocks, are used as test assets.[3] A few recent papers are devoted to the analysis of linear asset pricing factor models when the number of test assets N is large.

Ang, Liu, and Schwarz (2010) argue that using individual stock data, as opposed to forming portfolios, results in risk premia estimators with smaller variance. Their analysis, however, is justified only when T tends to infinity in the sense that, in their setting, the estimators are T-consistent but not N-consistent. Furthermore, they do not address the issue of bias in the risk premia estimates which turns out to be significant when N is large and T is small as our analysis illustrates. Extending the classical Gibbons, Ross, and Shanken (1989) test, Pesaran and Yamagata (2012) propose a number of tests for the zero alpha null hypothesis, while they are not concerned with the implications of the asset pricing model regarding factor risk premia. The feasible versions of their tests are justified when both N and T tend to infinity jointly at suitable rates. Their simulation as well as empirical evidence focuses on the S&P 500 universe of stocks. Gagliardini, Ossola, and Scaillet (2012) extend the two-pass cross-sectional methodology to the case of a conditional factor model incorporating firm characteristics and unbalanced panels. Their asymptotic theory, based on N and T jointly increasing to infinity at suitable rates, facilitates studying time varying risk premia. Chordia, Goyal, and Shanken (2015), building on Shanken (1992), use bias-corrected risk premia estimates in a context with individual stocks and time variation

[3]The long list of related papers includes, among others, Gibbons (1982), Shanken (1985), Connor and Korajczyk (1988), Lehmann and Modest (1988), Gibbons, Ross, and Shanken (1989), Harvey (1989), Lo and MacKinlay (1990), Zhou (1991), Shanken (1992), Connor and Korajczyk (1993), Zhou (1993), Zhou (1994), Berk (1995), Kim (1995), Hansen and Jagannathan (1997), Ghysels (1998), Jagannathan and Wang (1998), Kan and Zhou (1999), Jagannathan and Wang (2002), Chen and Kan (2004), Lewellen and Nagel (2006), Shanken and Zhou (2007), Kan and Robotti (2009), Hou and Kimmel (2010), Lewellen, Nagel, and Shanken (2010), Nagel and Singleton (2011), Ang and Kristensen (2012), Kan, Gospodinov, and Robotti (2013) and Kan, Robotti, and Shanken (2013).

in the betas through macroeconomic variables and firm characteristics. Their focus is the relative contribution of betas and characteristics in explaining cross-sectional differences in conditional expected returns. Two recent papers by Jegadeesh and Noh (2014) and Pukthuanthong, Roll, and Wang (2014) employ an instrumental variable approach to deal with the EIV problem in the risk premia estimation using individual stocks, where the instruments are betas estimated over separate time periods. However, they do not offer estimators of the variance-covariance matrix of the risk premia estimator. Instead, they resort to the original Fama-MacBeth approach for computing standard errors and test statistics and, hence, ignore the EIV problem in the estimation of the variance-covariance matrix (Shanken (1992)).[4]

We contribute to the existing literature by developing a two-pass CSR approach in order to estimate ex-post risk premia and, for the first time, construct associated asset pricing tests, when the number of assets N tends to infinity while the time-series length T is fixed. Recall that the second step of the two-pass procedure is a regression of returns on estimated betas. In the context of the standard linear regression model, it is well known that OLS estimators are consistent as long as a suitable orthogonality condition between the regression shocks and the regressors is satisfied. When N is fixed and T tends to infinity, this condition is satisfied and the two-pass CSR risk premia estimator is T-consistent. The EIV problem due to beta estimation error, however, manifests itself in the computation of standard errors. In contrast, in our context with T fixed and N increasing to infinity, the orthogonality condition is *not* satisfied and, hence, the two-pass CSR estimator is *not* N-consistent, as explained in Section 6 in Shanken (1992). At the heart of our approach is the beta estimate correction that we achieve by employing the regression-calibration approach. Using the corrected betas in the second pass yields N-consistent estimators of the risk premia. We further show that the risk premia estimator asymptotically follows a normal distribution and obtain the asymptotic variance-covariance matrix. Finally, incorporating a cluster structure for idiosyncratic shock correlations, we provide an N-consistent estimator of the asymptotic variance-covariance matrix which we use to develop statistics for testing the implications of the asset pricing model under examination.

This paper is not the first attempt to provide an N-consistent risk premia estimator. Litzenberger and Ramaswamy (1979), Shanken (1992), and Jagannathan, Skoulakis, and Wang (2010) provide related

[4]Specifically, Jegadeesh and Noh (2014) establish consistency results for their instrumental variable risk premia estimator but do not develop a sampling theory. Pukthuanthong, Roll, and Wang (2014) establish two results on the asymptotic distribution of their instrumental variable risk premia estimator: Theorem 3.1 based on N asymptotics and Theorem 3.2 based on joint N and T asymptotics. However, they do not offer an operational version of these theorems in the sense that they do not provide an estimator of the variance-covariance matrix of the risk premia estimator. Both papers use the original Fama-MacBeth approach to compute standard errors and test statistics.

estimators under different sets of assumptions.[5] Even though these estimators have existed for a long time, their asymptotic distribution and estimates of their variance-covariance matrix have not been developed and, hence, no inference tools have been available to date. Our paper fills this gap in the literature. It turns out that the estimator developed in Jagannathan, Skoulakis, and Wang (2010), adapted to our framework, is equivalent to our risk premia estimator. However, our estimation scheme is fundamentally different in that it is based on the regression-calibration approach and achieves the desired orthogonality condition in the second-pass regression by using EIV-corrected betas. As a result, our risk premia estimator has a convenient OLS form that we exploit to obtain its asymptotic distribution, as N tends to infinity, and develop novel asset pricing tests by comparing the cross-sectional risk premia estimates with the ex-post factor realizations.

To gauge the performance of our EIV-corrected estimator and the associated test statistics, we conduct Monte Carlo simulations calibrated to the CAPM and the Fama-French model. Our simulation exercise clearly illustrates the significant bias reduction in the cross-sectional alpha and risk premia estimates achieved by our beta correction approach and the good performance of our asset pricing tests for relevant sample sizes. Empirically, we employ our method to examine the implications of the CAPM and the Fama-French model using monthly data on individual stocks over the period 1946 to 2010. We reject the joint hypotheses on the alpha and the prices of risk for both models. However, we provide some new evidence that the HML factor of Fama and French (1996) can be a useful factor for pricing large cross sections of individual stocks besides portfolios sorted on book-to-market.

The rest of the paper is organized as follows. In Section 2, we develop a beta correction method that yields an N-consistent two-pass CSR risk premia estimator. We further obtain its asymptotic distribution, provide an estimator of its asymptotic variance-covariance matrix and develop new asset pricing tests. In Section 3, we provide Monte Carlo evidence on the finite sample behavior of the EIV-corrected risk premia estimator and the associated tests, while in Section 4 we present our empirical results. Finally, Section 5 concludes. The proofs of the main results are collected in the Appendix. Additional proofs and results are delegated to the Online Appendix.

[5] Litzenberger and Ramaswamy (1979) develop such an estimator under the assumptions that the disturbance variance-covariance matrix is diagonal and known. Theorem 5 in Shanken (1992) relaxes these assumptions and provides an N-consistent estimator of the ex-post risk premia under sufficiently weak cross-sectional dependence between the disturbances. Jagannathan, Skoulakis, and Wang (2010), in subsection 3.7, provide detailed assumptions under which such an N-consistent ex-post risk premia estimator is obtained (see their Theorem 7).

2 Econometric Framework

2.1 Model specification

Consider an economy with N traded assets and K factors. Let $\mathbf{r}_t = (r_{1,t}, \ldots, r_{N,t})'$ be the vector of returns of the N traded assets in excess of the risk-free return and $\mathbf{f}_t = (f_{1,t}, \ldots, f_{K,t})'$ be the vector of factor realizations at time t. We assume that return and factor data are available over times 1 through T, where T is finite and fixed, and formally consider the case in which the number of assets, N, tends to infinity.[6] Given that the time-series sample size is fixed, the uncertainty about the factors and the ex-post price of risk, cannot be resolved, unless the factors belong to the asset return space (see Shanken (1992), page 6).[7] Hence, since we cannot perform asset pricing tests on non-traded factors, we conduct our analysis conditionally on the factor realizations and focus on the case in which all factors are portfolio return spreads.

The data from time 1 to time τ, where $\tau < T$, are used to estimate the betas of all assets, while the data from time $\tau + 1$ to time T are used to test the asset pricing model under consideration. We refer to the periods covering times 1 through τ and $\tau + 1$ through T as the beta estimation and testing periods, respectively. Both τ and T are fixed throughout our analysis. The expectations of the excess return $\mathbf{r}_t$ and the factor $\mathbf{f}_t$ are denoted by $\boldsymbol{\mu}_r = \mathbb{E}[\mathbf{r}_t]$ and $\boldsymbol{\mu}_f = \mathbb{E}[\mathbf{f}_t]$, respectively. Furthermore, the $K \times K$ factor variance-covariance matrix is denoted by $\boldsymbol{\Sigma}_f = \mathbb{E}\left[(\mathbf{f}_t - \boldsymbol{\mu}_f)(\mathbf{f}_t - \boldsymbol{\mu}_f)'\right]$, while the $N \times K$ excess return-factor covariance matrix is denoted by $\boldsymbol{\Sigma}_{rf} = \mathbb{E}\left[(\mathbf{r}_t - \boldsymbol{\mu}_r)(\mathbf{f}_t - \boldsymbol{\mu}_f)'\right]$. The $N \times K$ beta matrix is then defined by $\mathbf{B} \equiv [\ \boldsymbol{\beta}_1 \ \cdots \ \boldsymbol{\beta}_N\ ]' = \boldsymbol{\Sigma}_{rf}\boldsymbol{\Sigma}_f^{-1}$, where $\boldsymbol{\beta}_i$ denotes the beta vector for the i-th asset, $i = 1, \ldots, N$.[8] Given that the factors comprising $\mathbf{f}_t$ belong to the return space, the risk premia vector equals the vector of factor expectations $\boldsymbol{\mu}_f$, and, hence, the corresponding linear beta model implies the pricing equation $\boldsymbol{\mu}_r = \mathbf{B}\boldsymbol{\mu}_f$. Defining $\mathbf{u}_t = \mathbf{r}_t - \mathbf{B}\mathbf{f}_t$ and imposing the pricing restriction, we obtain the following time-series regression representation:

$$\mathbf{r}_t = \mathbf{B}\mathbf{f}_t + \mathbf{u}_t, \text{ with } \quad \mathbb{E}[\mathbf{u}_t] = \mathbf{0}_N, \quad \mathbb{E}[\mathbf{u}_t\mathbf{f}_t'] = \mathbf{0}_{N\times K}. \tag{1}$$

[6] Note that we require a balanced panel of return data. However, this is not a very restrictive assumption given that T is finite and fixed in theory and small in our empirical applications. Specifically, we split the full sample period, which spans up to 70 years, into multiple subperiods, covering 8 to 10 years, over which we apply our method.

[7] In the notation of Shanken (1992), the vector of ex-post risk premia associated with the general factor $\mathbf{F}_1$ is given by $\overline{\boldsymbol{\gamma}}_1 = \boldsymbol{\gamma}_1 + \overline{\mathbf{F}}_1 - \mathbb{E}[\mathbf{F}_1]$. However, it is clear from the expression for $\overline{\boldsymbol{\gamma}}_1$ that, in the context of ex-post risk premia, one cannot test any null hypothesis about $\boldsymbol{\gamma}_1$ since the expectation $\mathbb{E}[\mathbf{F}_1]$ is unknown and so $\overline{\boldsymbol{\gamma}}_1$ is unknown as well. This is why we focus exclusively on traded factors.

[8] We assume that the betas are constant over the T time periods. Note, however, that this is not such a restrictive assumption given that T is relatively small in our empirical applications, ranging from 8 to 10 years.

For any vector time series $\{\mathbf{y}_1, \ldots, \mathbf{y}_\tau, \mathbf{y}_{\tau+1}, \ldots, \mathbf{y}_T\}$, denote the sample averages over the beta estimation period and the testing period by $\bar{\mathbf{y}}_1 = \frac{1}{\tau}\sum_{t=1}^{\tau}\mathbf{y}_t$ and $\bar{\mathbf{y}}_2 = \frac{1}{T-\tau}\sum_{t=\tau+1}^{T}\mathbf{y}_t$, respectively. Over the beta estimation period, covering times 1 through τ, time-series regressions of the N excess returns on the K realized factors are used to provide estimates of the beta matrix $\mathbf{B}$. Over the testing period, covering time periods $t = \tau+1, \ldots, T$, the data generating process in (1) implies that

$$\bar{\mathbf{r}}_2 = \mathbf{1}_N\alpha + \mathbf{B}\boldsymbol{\lambda} + \bar{\mathbf{u}}_2 = \mathbf{X}\boldsymbol{\gamma} + \bar{\mathbf{u}}_2 \tag{2}$$

where $\mathbf{X} = [\ \mathbf{1}_N \quad \mathbf{B}\ ]$, $\boldsymbol{\gamma} = [\ \alpha \quad \boldsymbol{\lambda}'\ ]'$, α is the cross-sectional intercept, which is equal to 0 when the lineal factor pricing model holds, and $\boldsymbol{\lambda} = \bar{\mathbf{f}}_2$. Recall that, since T is finite and fixed, our analysis is conducted conditionally on the factor realizations. Hence, following Shanken (1992), among others, we refer to $\boldsymbol{\lambda}$ as the ex-post price of risk. When the linear factor pricing model holds, the ex-post price of risk $\boldsymbol{\lambda}$ equals the average factor realization over the testing period, namely $\bar{\mathbf{f}}_2$, given that the factors belong to the return space.

The object of our inference is the $(K+1)\times 1$ vector $\boldsymbol{\gamma}$. To estimate $\boldsymbol{\gamma}$, we use a two-pass cross-sectional regression (CSR) approach inspired by the representation in equation (2). If the true beta matrix $\mathbf{B}$ were known, an N-consistent estimator of $\boldsymbol{\gamma}$ could be obtained by regressing the average excess return vector $\bar{\mathbf{r}}_2$ on a vector of ones and the beta matrix $\mathbf{B}$, under the reasonable assumption of zero limiting cross-sectional correlation between the betas and the shocks. However, the beta matrix $\mathbf{B}$ is not known and has to be estimated using the available data. The widely used two-pass CSR method proceeds as follows. Define the $N \times \tau$ excess return matrix $\mathbf{R}_1$ and the $K \times \tau$ factor realization matrix $\mathbf{F}_1$, over the beta estimation period, by

$$\mathbf{R}_1 = [\ \mathbf{r}_1 \quad \cdots \quad \mathbf{r}_\tau\ ] \ \text{ and } \ \mathbf{F}_1 = [\ \mathbf{f}_1 \quad \cdots \quad \mathbf{f}_\tau\ ]. \tag{3}$$

In the first pass, we use data from the beta estimation period, covering times 1 through τ, to obtain the standard time-series regression estimator of the beta matrix $\mathbf{B}$ given by $\widehat{\mathbf{B}} \equiv [\ \widehat{\boldsymbol{\beta}}_1 \quad \cdots \quad \widehat{\boldsymbol{\beta}}_N\ ]' = \widehat{\boldsymbol{\Sigma}}_{rf}\widehat{\boldsymbol{\Sigma}}_f^{-1}$ where $\widehat{\boldsymbol{\Sigma}}_{rf} = \frac{1}{\tau}\sum_{t=1}^{\tau}(\mathbf{r}_t - \bar{\mathbf{r}}_1)(\mathbf{f}_t - \bar{\mathbf{f}}_1)' = \frac{1}{\tau}\mathbf{R}_1\mathbf{J}_\tau\mathbf{F}_1'$ and $\widehat{\boldsymbol{\Sigma}}_f = \frac{1}{\tau}\sum_{t=1}^{\tau}(\mathbf{f}_t - \bar{\mathbf{f}}_1)(\mathbf{f}_t - \bar{\mathbf{f}}_1)' = \frac{1}{\tau}\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'$, where $\mathbf{J}_m = \mathbf{I}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'$ with $\mathbf{I}_m$ and $\mathbf{1}_m$ denoting the $m \times m$ identity matrix and the $m \times 1$ vector of ones, respectively, for any positive integer m.[9]

In the spirit of the standard two-pass CSR procedure, in the second pass, one would regress the

[9]Standard matrix algebra shows that $\mathbf{J}_m$ is a symmetric and idempotent matrix, and that $\text{tr}(\mathbf{J}_m) = m - 1$.

sample average excess return over the testing period $\bar{\mathbf{r}}_2$ on the first-pass estimate $\widehat{\mathbf{X}}$, as suggested by equation (2), to obtain the CSR estimator:

$$\widehat{\gamma} = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^{-1}\widehat{\mathbf{X}}'\bar{\mathbf{r}}_2. \tag{4}$$

where

$$\widehat{\mathbf{X}} = [\ \mathbf{1}_N \quad \widehat{\mathbf{B}}\ ]. \tag{5}$$

However, it turns out that, when the time-series sample size T is fixed, the CSR estimator $\widehat{\gamma}$ is not consistent, as $N \to \infty$, due to the presence of the non-trivial beta estimation error. This is a manifestation of the well-known EIV problem, as pointed out in Section 6 of Shanken (1992).[10] In what follows, we obtain an expression for the asymptotic bias of $\widehat{\gamma}$ and then use the regression-calibration method to develop a correction that yields an N-consistent estimator.

Define the $N \times \tau$ idiosyncratic shock matrix $\mathbf{U}_1$, over the beta estimation period, by

$$\mathbf{U}_1 = [\ \mathbf{u}_1 \quad \cdots \quad \mathbf{u}_\tau\ ] = \left[\ \mathbf{u}_{1,[1]} \quad \cdots \quad \mathbf{u}_{1,[N]}\ \right]', \tag{6}$$

where $\mathbf{u}'_{1,[i]}$ denotes the i-th row of $\mathbf{U}_1$, for $i = 1, \ldots, N$. Similarly, define the $N \times (T - \tau)$ idiosyncratic shock matrix $\mathbf{U}_2$, over the testing period, by

$$\mathbf{U}_2 = [\ \mathbf{u}_{\tau+1} \quad \cdots \quad \mathbf{u}_T\ ] = \left[\ \mathbf{u}_{2,[1]} \quad \cdots \quad \mathbf{u}_{2,[N]}\ \right]', \tag{7}$$

where $\mathbf{u}'_{2,[i]}$ denotes the i-th row of $\mathbf{U}_2$, for $i = 1, \ldots, N$. It is worth noting that $\bar{\mathbf{u}}_2$ and $\mathbf{U}_2$ satisfy the following relationship:

$$\bar{\mathbf{u}}_2 = \frac{1}{T-\tau}\mathbf{U}_2\mathbf{1}_{T-\tau}. \tag{8}$$

Since $\mathbf{R}_1 = \mathbf{B}\mathbf{F}_1 + \mathbf{U}_1$ implies that $\widehat{\mathbf{B}} = (\mathbf{R}_1\mathbf{J}_\tau\mathbf{F}'_1)(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}'_1)^{-1} = \mathbf{B} + (\mathbf{U}_1\mathbf{J}_\tau\mathbf{F}'_1)(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}'_1)^{-1}$, we can express the $N \times K$ beta estimation error matrix $\mathbf{\Xi}$ as follows

$$\mathbf{\Xi} = \widehat{\mathbf{B}} - \mathbf{B} = (\mathbf{U}_1\mathbf{J}_\tau\mathbf{F}'_1)(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}'_1)^{-1} = \mathbf{U}_1\mathbf{G}_1, \tag{9}$$

[10] When the time-series sample size tends to infinity, the CSR estimator is consistent. However, the variance of the CSR estimator is affected and properly computing standard errors requires appropriate corrections. See, among others, Shanken (1992) and Jagannathan and Wang (1998).

where $\mathbf{G}_1$ is the $\tau \times K$ matrix defined by

$$\mathbf{G}_1 = \mathbf{J}_\tau \mathbf{F}_1' \left(\mathbf{F}_1 \mathbf{J}_\tau \mathbf{F}_1'\right)^{-1}. \tag{10}$$

To illustrate the effect of the beta estimation error in the second-pass of the two-pass CSR procedure, we observe that equation (2) can be expressed as $\overline{\mathbf{r}}_2 = \widehat{\mathbf{X}}\boldsymbol{\gamma} + (\mathbf{X} - \widehat{\mathbf{X}})\boldsymbol{\gamma} + \overline{\mathbf{u}}_2$ or

$$\overline{\mathbf{r}}_2 = \widehat{\mathbf{X}}\boldsymbol{\gamma} + \widehat{\boldsymbol{\omega}}, \tag{11}$$

where

$$\widehat{\boldsymbol{\omega}} = (\mathbf{X} - \widehat{\mathbf{X}})\boldsymbol{\gamma} + \overline{\mathbf{u}}_2 = -(\mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda} + \overline{\mathbf{u}}_2. \tag{12}$$

Note that the error term $\widehat{\boldsymbol{\omega}}$ incorporates the idiosyncratic shock term $\overline{\mathbf{u}}_2$ as well as the term $-(\mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda}$ which reflects the beta estimation error. Using this representation and equation (4), we can express the estimator $\widehat{\boldsymbol{\gamma}}$ as follows:

$$\widehat{\boldsymbol{\gamma}} = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^{-1}\widehat{\mathbf{X}}'(\widehat{\mathbf{X}}\boldsymbol{\gamma} + \widehat{\boldsymbol{\omega}}) = \boldsymbol{\gamma} + (\widehat{\mathbf{X}}'\widehat{\mathbf{X}}/N)^{-1}(\widehat{\mathbf{X}}'\widehat{\boldsymbol{\omega}}/N). \tag{13}$$

We show next that, under the following reasonable assumption, the probability limit of $\widehat{\mathbf{X}}'\widehat{\boldsymbol{\omega}}/N$, as $N \to \infty$, is a non-zero matrix and, hence, $\widehat{\boldsymbol{\gamma}}$ is not an N-consistent estimator of $\boldsymbol{\gamma}$. Under the same assumption, we later develop a method for correcting the beta estimator so that the resulting ex-post risk premia estimator is N-consistent.

Assumption 1 **(i)** *For all* $t = 1, \ldots, T$, *as* $N \to \infty$, $\frac{1}{N}\mathbf{X}'\mathbf{u}_t = \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N} u_{i,t} & \frac{1}{N}\sum_{i=1i}^{N} u_{i,t}\boldsymbol{\beta}_i' \end{bmatrix}' \xrightarrow{p} \mathbf{0}_{K+1}$. **(ii)** *For all* $t, t' = 1, \ldots, T$, *as* $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N} u_{i,t}u_{i,t'} \xrightarrow{p} v\mathbf{1}_{[t=t']}$, *where* v *is a positive constant.* **(iii)** *As* $N \to \infty$, $\overline{\boldsymbol{\beta}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i \to \boldsymbol{\mu}_\beta$. **(iv)** *As* $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)' \to \mathbf{V}_\beta$, *where* $\mathbf{V}_\beta$ *is a symmetric and positive definite matrix.*

Assumption 1(i) states that, at each time t, the cross-sectional average of the shocks $u_{i,t}$ converges to zero, and the limiting cross-sectional correlation between the shocks $u_{i,t}$ and the betas $\boldsymbol{\beta}_i$ is also zero, as the number of assets N tends to ∞. For $t = t'$, Assumption 1(ii) states the limiting cross-sectional variance of shocks $u_{i,t}$ exists and is denoted by v, while, for $t \neq t'$, Assumption 1(ii) states that the limiting cross-sectional correlation between $u_{i,t}$ and $u_{i,t'}$ vanishes.[11] It is worth noting that this assumption does not require the idiosyncratic variance, at the individual asset level, to be constant over time. Indeed, idiosyncratic variances can vary over time as long as the *limiting cross-sectional average* of the squared shocks stays the same over time. This assumption is simply a version of the Law of Large Numbers applied to the cross section of squared residuals. Furthermore, Assumption 1(ii) requires v to be constant over T time periods, ranging from 8 to 10 years in all of our empirical analysis, and so it is not very restrictive.[12] Assumption 1(iii) states that the limiting cross-sectional average of the betas $\boldsymbol{\beta}_i$ exists while Assumption 1(iv) states that the limiting cross-sectional variance of the betas $\boldsymbol{\beta}_i$ exists and is a symmetric and positive definite matrix.

Since $\widehat{\mathbf{X}} = [\ \mathbf{1}_N \quad \widehat{\mathbf{B}}\ ]$, we have $\widehat{\mathbf{X}}'\widehat{\boldsymbol{\omega}}/N = [\ (\mathbf{1}_N'\widehat{\boldsymbol{\omega}}/N)' \quad (\widehat{\mathbf{B}}'\widehat{\boldsymbol{\omega}}/N)'\ ]'$. It follows from equations (12) and (8) that $\mathbf{1}_N'\widehat{\boldsymbol{\omega}}/N = -(\mathbf{1}_N'\mathbf{U}_1/N)\mathbf{G}_1\boldsymbol{\lambda} + \frac{1}{T-\tau}(\mathbf{1}_N'\mathbf{U}_2/N)\mathbf{1}_{T-\tau}$ and so from Assumption 1(i) it follows that, as $N \to \infty$, $\mathbf{1}_N'\widehat{\boldsymbol{\omega}}/N \xrightarrow{p} 0$. Similarly, from equations (12), (9), and (8), we obtain $\widehat{\mathbf{B}}'\widehat{\boldsymbol{\omega}}/N = -(\mathbf{B}'\mathbf{U}_1/N)\mathbf{G}_1\boldsymbol{\lambda} + \frac{1}{T-\tau}(\mathbf{B}'\mathbf{U}_2/N)\mathbf{1}_{T-\tau} - \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_1/N)\mathbf{G}_1\boldsymbol{\lambda} + \frac{1}{T-\tau}\mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/N)\mathbf{1}_{T-\tau}$ which, in light of Assumptions 1(i) and 1(ii), yields that, as $N \to \infty$, $\widehat{\mathbf{B}}'\widehat{\boldsymbol{\omega}}/N \xrightarrow{p} -v\,(\mathbf{G}_1'\mathbf{G}_1)\,\boldsymbol{\lambda}$. Combining these two probability limits yields

$$\widehat{\mathbf{X}}'\widehat{\boldsymbol{\omega}}/N \xrightarrow{p} -\begin{bmatrix} 0 & \mathbf{0}_K' \\ \mathbf{0}_K & v\mathbf{W}_1 \end{bmatrix}\boldsymbol{\gamma}, \tag{14}$$

where

$$\mathbf{W}_1 = \mathbf{G}_1'\mathbf{G}_1. \tag{15}$$

Therefore, the standard least-squares orthogonality condition is violated in the context of equation (11), implying that $\widehat{\boldsymbol{\gamma}}$ is not an N-consistent estimator of $\boldsymbol{\gamma}$. Indeed, the next proposition provides the exact probability limit of $\widehat{\boldsymbol{\gamma}}$, as $N \to \infty$, from which one can assess its asymptotic bias.[13]

[11] The same assumption has been used in other papers in the literature. Specifically, we refer to Theorem 5 in Shanken (1992) and the associated footnote 23 on page 22, and also Theorem 7 in Jagannathan, Skoulakis, and Wang (2010) and the associated Assumption F on page 103.

[12] In an extension of the present framework, we have relaxed the assumption that v is constant over time. Using a minimum distance approach, we provide consistent estimators of the time-series vector of the limiting cross-sectional variances. These estimators can then be used to obtain corrected beta estimators by the regression-calibration method, which are finally used in the second-pass cross-sectional regression. We did not pursue this generalization as it makes very little difference in the empirical results and would unnecessarily complicate the analysis. Further details and empirical results from this approach are available upon request.

[13] The proof of the proposition can be found in the Online Appendix.

Proposition 1 *Under Assumption 1, as $N \to \infty$, we have*

$$\widehat{\gamma} \xrightarrow{p} \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \mathbf{L}_1^{-1}(v\mathbf{W}_1) \\ \mathbf{0}_K & \mathbf{L}_1^{-1}\mathbf{V}_\beta \end{bmatrix} \gamma,$$

where the $K \times K$ matrix $\mathbf{L}_1$ is given by

$$\mathbf{L}_1 = \mathbf{V}_\beta + v\mathbf{W}_1. \tag{16}$$

Note that $\mathbf{V}_\beta$ is positive definite by Assumption 1(iv), v is a positive scalar by Assumption 1(ii), and $\mathbf{W}_1$ is positive definite by definition (15). It follows that $\mathbf{L}_1$ is positive definite and therefore invertible.

Having established that the standard two-pass CSR procedure would lead to an inconsistent estimator of the quantity of interest $\boldsymbol{\gamma}$, we proceed by employing the regression-calibration approach in order to provide a correction term that leads to an N-consistent estimator of $\boldsymbol{\gamma}$. Regression-calibration is one of the various methods used to handle regressor measurement errors in regression models. The beta estimation error in our context is an example of such measurement error. The regressor measurement error issue, also referred to as the EIV problem, has been recognized as an important topic and studied thoroughly in the economics literature. While several applications of the regression-calibration method have appeared in the statistics literature,[14] the regression-calibration method has not been used in the economics or finance literature. The likely reason is that the method requires information on the variance of the measurement error which can be obtained in certain circumstances using various approaches such as replication, validation, or instrumental variables. In what follows, we show how to estimate the cross-sectional variance of the beta estimation error in our application of the regression-calibration method to testing asset pricing models using a large cross section of assets. To the best of our knowledge, this paper presents the first application of this approach to handle the EIV problem encountered in CSR asset pricing tests.

The main idea behind the regression-calibration approach is to replace the matrix $\widehat{\mathbf{X}}$ in the regression (11) by a "calibrated" matrix $\check{\mathbf{X}}$, suitably selected so that the limiting cross-sectional orthogonality condition, necessary for N-consistency, is satisfied. We do so by introducing a correction matrix $\mathbf{C}$ and

[14] Prentice (1982) introduces the idea to the study of the proportional hazard model while Armstrong (1985) applies the procedure to general linear models. Carroll and Stefanski (1990) and Gleser (1990) extend the application of the method to nonlinear models. Hardin, Schmiediche, and Carroll (2003) show how to implement the method in Stata to fit generalized linear models. Carroll, Ruppert, Stefanski, and Crainiceanu (2006) present a textbook account of the method with several applications and recently developed computation tools.

defining the corrected version of $\widehat{\mathbf{B}}$ as follows:[15]

$$\breve{\mathbf{B}} = \mathbf{1}_N \widehat{\boldsymbol{\mu}}'_\beta + (\widehat{\mathbf{B}} - \mathbf{1}_N \widehat{\boldsymbol{\mu}}'_\beta)(\mathbf{I}_K - \mathbf{C}), \tag{17}$$

where $\widehat{\boldsymbol{\mu}}_\beta = \sum_{i=1}^N \widehat{\boldsymbol{\beta}}_i / N = \widehat{\mathbf{B}}' \mathbf{1}_N / N$ is a natural N-consistent estimator of $\boldsymbol{\mu}_\beta$. Indeed, this is the case since

$$\widehat{\boldsymbol{\mu}}_\beta = \widehat{\mathbf{B}}' \mathbf{1}_N / N = (\mathbf{B} + \boldsymbol{\Xi})' \mathbf{1}_N / N = \overline{\boldsymbol{\beta}} + \mathbf{G}'_1 (\mathbf{U}'_1 \mathbf{1}_N / N) \xrightarrow{p} \boldsymbol{\mu}_\beta, \tag{18}$$

where we use equation (9) and invoke Assumptions 1(i) and 1(iii).

To express equation (11) in terms of the corrected regressor matrix $\breve{\mathbf{B}}$, we note that $\widehat{\mathbf{B}} = \breve{\mathbf{B}} + (\widehat{\mathbf{B}} - \mathbf{1}_N \widehat{\boldsymbol{\mu}}'_\beta)\mathbf{C}$ which yields

$$\overline{\mathbf{r}}_2 = \mathbf{1}_N \alpha + (\widehat{\mathbf{B}} - \mathbf{U}_1 \mathbf{G}_1)\boldsymbol{\lambda} + \overline{\mathbf{u}}_2 = \mathbf{1}_N \alpha + (\breve{\mathbf{B}} + (\widehat{\mathbf{B}} - \mathbf{1}_N \widehat{\boldsymbol{\mu}}'_\beta)\mathbf{C} - \mathbf{U}_1 \mathbf{G}_1)\boldsymbol{\lambda} + \overline{\mathbf{u}}_2 = \breve{\mathbf{X}} \boldsymbol{\gamma} + \breve{\boldsymbol{\omega}}, \tag{19}$$

where $\breve{\mathbf{X}} = [\ \mathbf{1}_N \quad \breve{\mathbf{B}}\ ]$ and

$$\breve{\boldsymbol{\omega}} = ((\widehat{\mathbf{B}} - \mathbf{1}_N \widehat{\boldsymbol{\mu}}'_\beta)\mathbf{C} - \mathbf{U}_1 \mathbf{G}_1)\boldsymbol{\lambda} + \overline{\mathbf{u}}_2. \tag{20}$$

The next step in deriving the corrected risk premia estimator that achieves N-consistency, facilitated by the following lemma, is to identify the correction matrix $\mathbf{C}$ so that the orthogonality condition $\breve{\mathbf{X}}' \breve{\boldsymbol{\omega}} / N \xrightarrow{p} \mathbf{0}_K$ is satisfied.[16]

Lemma 2 *Under Assumption 1, as* $N \to \infty$, $\mathbf{1}'_N \breve{\boldsymbol{\omega}} / N \xrightarrow{p} 0$ *and*

$$\breve{\mathbf{B}}' \breve{\boldsymbol{\omega}} / N \xrightarrow{p} (\mathbf{I}_K - \mathbf{C})' \left[(\mathbf{V}_\beta + v \mathbf{W}_1) \mathbf{C} - v \mathbf{W}_1 \right] \boldsymbol{\lambda},$$

where the $K \times K$ *matrix* $\mathbf{W}_1$ *is defined by (15).*

It follows from the preceding lemma that, to achieve the desired orthogonality condition $\breve{\mathbf{X}}' \breve{\boldsymbol{\omega}} / N \xrightarrow{p}$

[15] The corrected beta estimator $\breve{\mathbf{B}}$ is an example of a shrinkage estimator. Such estimators have appeared in the literature in various contexts. In particular, Vasicek (1973) develops a Bayesian shrinkage approach for improving beta estimation. There are important differences between our estimator and the estimator in Vasicek (1973). Since the primary goal in Vasicek (1973) is to obtain improved beta estimators, the shrinkage factor depends on the particular asset. If the estimation error for a given asset is significant, then the shrinkage intensity towards the cross-sectional beta estimator average increases, according to the Vasicek (1973) approach. In contrast, the focus of our methodology is not the improvement of the beta estimators per se. Instead, our goal is to modify the beta estimator so that the resulting CSR estimator in the second pass is N-consistent. As a result, our estimator uses the same shrinkage intensity across all assets.

[16] The proof of the lemma can be found in the Online Appendix.

$\mathbf{0}_{K+1}$, we need to choose

$$\mathbf{C} = v\left(\mathbf{V}_\beta + v\mathbf{W}_1\right)^{-1}\mathbf{W}_1 = v\mathbf{L}_1^{-1}\mathbf{W}_1, \tag{21}$$

where the $\mathbf{L}_1$ is defined by (16).

The correction factor $\mathbf{C}$ has a nice intuitive interpretation. It is the "ratio" between $v\mathbf{W}_1$ and $\mathbf{L}_1 = \mathbf{V}_\beta + v\mathbf{W}_1$. Note that $v\mathbf{W}_1$ can be interpreted as the noise variance since, according to Assumption 1(ii), we have $\mathbf{\Xi}'\mathbf{\Xi}/N \xrightarrow{p} v\mathbf{W}_1$, as $N \to \infty$. Moreover, it follows from the proof of Proposition 1 that, as $N \to \infty$, $(\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')'(\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')/N \to \mathbf{L}_1$, and therefore $\mathbf{L}_1$ can be interpreted as the signal variance. The correction factor $\mathbf{C}$ thus can thought of as the noise variance to signal variance "ratio". When there is significant beta estimation error, that is noise, the correction factor becomes more pronounced and shrinks the beta estimator towards the cross-sectional average beta in order to mitigate the effect of the estimation error.

While the cross-sectional orthogonality condition is satisfied with our choice of $\mathbf{C}$, the resulting corrected estimator is not feasible since the matrix $\mathbf{C}$ is an unknown population quantity. In the following subsection, we provide an N-consistent estimator of $\mathbf{C}$ and establish the N-consistency of the corresponding feasible estimator of the parameter of interest $\boldsymbol{\gamma}$.

2.2 Feasible N-consistent CSR estimator

Recall that the correction matrix is given by $\mathbf{C} = v\mathbf{L}_1^{-1}\mathbf{W}_1$. Since the matrix $\mathbf{W}_1$ is observed after the beta estimation period, to obtain an N-consistent estimator of $\mathbf{C}$, it suffices to obtain N-consistent estimators of v and $\mathbf{L}_1$.

To develop an N-consistent estimator of v, we start by observing that Assumption 1(ii) implies that $\mathbf{U}_1'\mathbf{U}_1/N \xrightarrow{p} v\mathbf{I}_\tau$ and so $\frac{1}{\tau}\mathrm{tr}(\mathbf{U}_1'\mathbf{U}_1/N) \xrightarrow{p} v$, where tr denotes the trace operator. This suggests that $\mathrm{tr}(\widehat{\mathbf{U}}_1'\widehat{\mathbf{U}}_1/N)$ can be used to develop an estimator of v, where $\widehat{\mathbf{U}}_1$ is the $N \times \tau$ matrix of time-series residuals, given by

$$\widehat{\mathbf{U}}_1 = [\ \widehat{\mathbf{u}}_1 \quad \cdots \quad \widehat{\mathbf{u}}_\tau\ ] \equiv \left[\ \widehat{\mathbf{u}}_{1,[1]} \quad \cdots \quad \widehat{\mathbf{u}}_{1,[N]}\ \right]' \tag{22}$$

where

$$\widehat{\mathbf{u}}_t = (\mathbf{r}_t - \overline{\mathbf{r}}_1) - \widehat{\mathbf{B}}(\mathbf{f}_t - \overline{\mathbf{f}}_1), \quad t = 1, \ldots, \tau, \tag{23}$$

and $\widehat{\mathbf{u}}'_{1,[i]}$ is the i-th row of $\widehat{\mathbf{U}}_1$, for $i = 1, \ldots, N$. The following lemma, accounting for the appropriate degrees-of-freedom adjustment, provides the desired N-consistent estimator of v.[17] It is easily shown that the matrix of residuals $\widehat{\mathbf{U}}_1$ can be expressed as $\widehat{\mathbf{U}}_1 = \mathbf{U}_1\mathbf{H}_1$, where the matrix $\mathbf{H}_1$ is defined by[18]

$$\mathbf{H}_1 = \mathbf{J}_\tau - \mathbf{J}_\tau \mathbf{F}_1' \left(\mathbf{F}_1 \mathbf{J}_\tau \mathbf{F}_1'\right)^{-1} \mathbf{F}_1 \mathbf{J}_\tau. \tag{24}$$

Lemma 3 *Under Assumption 1,*

$$\widehat{v} = \frac{tr(\widehat{\mathbf{U}}_1' \widehat{\mathbf{U}}_1 / N)}{tr(\mathbf{H}_1)} \tag{25}$$

is an N-consistent estimator of v, where the matrix of residuals $\widehat{\mathbf{U}}_1$ is defined in (22) and the matrix $\mathbf{H}_1$ is defined in (24).

Furthermore, it turns out that $\widehat{\mathbf{L}}_1 = \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N$ is an N-consistent estimator of the matrix $\mathbf{L}_1$ defined in (16). The next proposition proves this fact and offers an N-consistent estimator of $\mathbf{C}$.

Proposition 4 *Under Assumption 1, the matrix*

$$\widehat{\mathbf{C}} = \widehat{v}\widehat{\mathbf{L}}_1^{-1}\mathbf{W}_1 \tag{26}$$

is an N-consistent estimator of the matrix $\mathbf{C}$ defined by (21), where

$$\widehat{\mathbf{L}}_1 = \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N. \tag{27}$$

Using the feasible correction matrix $\widehat{\mathbf{C}}$, we define

$$\widetilde{\mathbf{B}} = \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + (\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')(\mathbf{I}_K - \widehat{\mathbf{C}}), \tag{28}$$

which is the feasible version of the corrected second-pass regression matrix $\breve{\mathbf{B}}$. Then, letting

$$\widetilde{\mathbf{X}} = [\ \mathbf{1}_N \quad \widetilde{\mathbf{B}}\ ], \tag{29}$$

[17] The proof of the lemma can be found in the Online Appendix.

[18] Standard matrix algebra shows that $\mathbf{H}_1$ is symmetric and idempotent. Moreover, it follows from the properties of the trace operator that $\mathrm{tr}\,(\mathbf{H}_1) = \tau - K - 1$.

we obtain the feasible version of equation (19) as follows:

$$\bar{\mathbf{r}}_2 = \widetilde{\mathbf{X}}\boldsymbol{\gamma} + \widetilde{\boldsymbol{\omega}}, \tag{30}$$

where

$$\widetilde{\boldsymbol{\omega}} = ((\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')\widehat{\mathbf{C}} - \mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda} + \bar{\mathbf{u}}_2. \tag{31}$$

The following proposition shows that using the regression representation in (30) provides the desired feasible N-consistent estimator of the vector $\boldsymbol{\gamma}$.

Proposition 5 *Under Assumption 1,*

$$\widetilde{\boldsymbol{\gamma}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\bar{\mathbf{r}}_2, \tag{32}$$

where the matrix $\widetilde{\mathbf{X}}$ is defined in (29), is an N-consistent estimator of $\boldsymbol{\gamma}$.

Our estimator $\widetilde{\boldsymbol{\gamma}}$ is related to other estimators that have been developed in the literature. Jagannathan, Skoulakis, and Wang (2010), proceeding in the spirit of Theorem 5 in Shanken (1992), present an N-estimator of the ex-post risk premia. According to Theorem 7 in Jagannathan, Skoulakis, and Wang (2010), this estimator, adapted to our setting with separate estimation and testing periods and using our notation, assumes the form[19] $\boldsymbol{\gamma}^\dagger = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}} - N\widehat{v}\mathbf{M}'\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1}\mathbf{M})^{-1}\widehat{\mathbf{X}}'\bar{\mathbf{r}}_2$, where $\mathbf{M} = \begin{bmatrix} \mathbf{0}_K & \mathbf{I}_K \end{bmatrix}$. According to Lemma 14, presented in the Online Appendix, we have that $(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}' = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}} - N\widehat{v}\mathbf{M}'\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1}\mathbf{M})^{-1}\widehat{\mathbf{X}}'$. It follows that the two estimators, $\widetilde{\boldsymbol{\gamma}}$ and $\boldsymbol{\gamma}^\dagger$, are equivalent. Despite their equivalence, there is an important difference between the two representations. Our estimator $\widetilde{\boldsymbol{\gamma}}$, obtained using the regression-calibration approach, has a convenient OLS form that we exploit in order to obtain its asymptotic distribution and develop novel asset pricing tests that focus on ex-post risk premia based on N asymptotics with a fixed T, as we illustrate next.

2.3 Asymptotic distribution of the N-consistent CSR estimator

The preceding section establishes the N-consistency of the feasible estimator $\widetilde{\boldsymbol{\gamma}}$. In this section, under an additional suitable assumption, we show that $\widetilde{\boldsymbol{\gamma}}$ is asymptotically normal and provide the expression

[19]See Section 3.7, and particularly Theorem 7, in Jagannathan, Skoulakis, and Wang (2010) for details. We also correct a typo in the statement of their Theorem 7 where the matrix $\mathbf{M}$ should be equal to $\begin{bmatrix} \mathbf{0}_K & \mathbf{I}_K \end{bmatrix}$.

for its asymptotic variance-covariance matrix. Recall from (32) and (30) that $\widetilde{\boldsymbol{\gamma}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\overline{\mathbf{r}}_2$ and $\overline{\mathbf{r}}_2 = \widetilde{\mathbf{X}}\boldsymbol{\gamma} + \widetilde{\boldsymbol{\omega}}$, from which we obtain $\sqrt{N}\,(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N)^{-1}(\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N})$. It is shown in the proof of Proposition 5 that, as $N \to \infty$, $\frac{1}{N}\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}} \xrightarrow{p} \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is the invertible $(K+1) \times (K+1)$ matrix defined by (51). Hence, to determine the asymptotic distribution of $\widetilde{\boldsymbol{\gamma}}$, one needs to determine the asymptotic distribution of $\frac{1}{\sqrt{N}}\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}$. It turns out that the latter equals $\sqrt{N}\boldsymbol{\Pi}\mathbf{e}$ plus an asymptotically negligible term, where $\boldsymbol{\Pi}$ is a suitable matrix (for its exact form see equation (70) in the Appendix) and $\mathbf{e}$ is the $\mathcal{T}$-dimensional random vector defined by

$$\mathbf{e} = \begin{bmatrix} \mathbf{U}_1'\mathbf{1}_N/N \\ \mathbf{U}_2'\mathbf{1}_N/N \\ \text{vec}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau) \\ \text{vec}(\mathbf{U}_1'\mathbf{U}_2/N) \\ \text{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/N) \\ \text{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/N) \end{bmatrix}, \tag{33}$$

where $\mathcal{T} = (\tau + K + 1)T$ and vec denotes the column-stacking operator.

The following lemma provides the first step towards obtaining the asymptotic distribution of the estimator $\widetilde{\boldsymbol{\gamma}}$.[20]

Lemma 6 *The vector* $\mathbf{e}$*, defined in equation (33), can be expressed as*

$$\mathbf{e} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_i,$$

where

$$\mathbf{e}_i = \begin{bmatrix} \mathbf{e}_{i,1}' & \mathbf{e}_{i,2}' & \mathbf{e}_{i,3}' & \mathbf{e}_{i,4}' \end{bmatrix}', \tag{34}$$

and the $T \times 1$ *vector* $\mathbf{e}_{i,1}$*, the* $\tau^2 \times 1$ *vector* $\mathbf{e}_{i,2}$*, the* $(T-\tau)\tau \times 1$ *vector* $\mathbf{e}_{i,3}$*, and the* $TK \times 1$ *vector* $\mathbf{e}_{i,4}$ *are given by*

$$\mathbf{e}_{i,1} = \begin{bmatrix} \mathbf{u}_{1,[i]} \\ \mathbf{u}_{2,[i]} \end{bmatrix}, \quad \mathbf{e}_{i,2} = \mathbf{u}_{1,[i]} \otimes \mathbf{u}_{1,[i]} - v\mathbf{i}_\tau, \quad \mathbf{e}_{i,3} = \mathbf{u}_{2,[i]} \otimes \mathbf{u}_{1,[i]}, \quad \mathbf{e}_{i,4} = \begin{bmatrix} \mathbf{u}_{1,[i]} \\ \mathbf{u}_{2,[i]} \end{bmatrix} \otimes (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta),$$

[20] The proof of the lemma can be found in the Online Appendix.

$\otimes$ *denotes the Kronecker product, and* $\mathbf{i}_\tau$ *is the* $\tau^2 \times 1$ *vector defined as*

$$\mathbf{i}_\tau = vec(\mathbf{I}_\tau). \tag{35}$$

Note that Assumption 1 implies that $\mathbf{e} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_i \xrightarrow{p} \mathbf{0}_{\mathcal{T}}$, as $N \to \infty$. In order to obtain the asymptotic distribution of $\widetilde{\boldsymbol{\gamma}}$, we make the following mild assumption, where $\xrightarrow{d}$ denotes convergence in distribution, that postulates that $\mathbf{e}_i$ satisfies a cross-sectional central limit theorem.

Assumption 2 *As* $N \to \infty$, $\sqrt{N}\mathbf{e} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{e}_i \xrightarrow{d} N(\mathbf{0}_{\mathcal{T}}, \mathbf{V}_e)$.

We address the issue of determining and estimating the variance-covariance matrix $\mathbf{V}_e$, under suitable conditions, in the next subsection. The following theorem establishes the asymptotic distribution of the estimator $\widetilde{\boldsymbol{\gamma}}$.

Theorem 7 *Under Assumptions 1 and 2, as* $N \to \infty$,

$$\sqrt{N}\left(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right) \xrightarrow{d} N\left(\mathbf{0}_{K+1}, \mathbf{V}_\gamma\right)$$

where

$$\mathbf{V}_\gamma = \boldsymbol{\Omega}^{-1}\boldsymbol{\Pi}\mathbf{V}_e\boldsymbol{\Pi}'\boldsymbol{\Omega}^{-1}, \tag{36}$$

$\boldsymbol{\Omega} = p\text{-}\lim \frac{1}{N}\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}$ *(see equation (50)), and the matrix* $\boldsymbol{\Pi}$ *is given by (70).*

Naturally, the matrix $\boldsymbol{\Omega}$ is estimated consistently by $\widehat{\boldsymbol{\Omega}} = \frac{1}{N}\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}$. A N-consistent estimator of $\boldsymbol{\Pi}$ is readily obtained through a standard sample analogue argument based on the expression of $\boldsymbol{\Pi}$ in (70). In the following subsection, we obtain an N-consistent estimator of $\mathbf{V}_e$, based on which an N-consistent estimator of $\mathbf{V}_\gamma$ is easily constructed using equation (36).

2.4 Estimation of the variance-covariance matrix $\mathbf{V}_e$

Recall that, according to Assumption 2, $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{e}_i \xrightarrow{d} N(\mathbf{0}_{\mathcal{T}}, \mathbf{V}_e)$. The goal of this subsection is to provide an estimator of the variance-covariance matrix $\mathbf{V}_e$ under suitable conditions. Specifically, we incorporate cross-sectional correlations in the shocks $\mathbf{e}_i$ by using a clustering approach. We assume that there are M_N clusters and that the m-th cluster consists of N_m stocks, for $m = 1, \ldots, M_N$, so that $\sum_{m=1}^{M_N} N_m = N$. For all N, we assume that the cluster sizes N_m, $m = 1, \ldots, M_N$ are bounded.

As $N \to \infty$, the number of clusters, M_N, is assumed to increase so that $\frac{N}{M_N} \to G$, where G is to be interpreted as the average cluster size. For $m = 1, \ldots, M_N$, let I_m be the set of all indices i for which the i-th stock belongs to the m-th cluster, and define the aggregate cluster shocks

$$\boldsymbol{\eta}_m = \begin{bmatrix} \boldsymbol{\eta}'_{m,1} & \boldsymbol{\eta}'_{m,2} & \boldsymbol{\eta}'_{m,3} & \boldsymbol{\eta}'_{m,4} \end{bmatrix}' = \sum_{i \in I_m} \mathbf{e}_i,$$

so that $\boldsymbol{\eta}_{m,k} = \sum_{i \in I_m} \mathbf{e}_{i,k}$, for $k = 1, \ldots, 4$. To facilitate a concise representation of the high-dimensional vectors involved in the subsequent analysis, we introduce some additional notation. For any positive integer p, we define k_p and ℓ_p by

$$k_p = \left[\frac{p-1}{\tau}\right] + 1, \quad \ell_p = p - (k_p - 1)\tau, \tag{37}$$

where $[a]$ denotes the largest integer less than or equal to a. Note that $\ell_p \in \{1, \ldots, \tau\}$ and k_p, defined as above, are the unique such positive integers satisfying $p = (k_p - 1)\tau + \ell_p$. Defining, for $m = 1, \ldots, M_N$,

$$\zeta^{[1]}_{m,r} = \sum\nolimits_{i \in I_m} u_{i,r}, \quad r = 1, \ldots, T, \tag{38}$$

$$\zeta^{[2]}_{m,p} = \sum\nolimits_{i \in I_m} u_{i,k_p} u_{i,\ell_p}, \quad p = 1, \ldots, \tau^2, \tag{39}$$

$$\zeta^{[3]}_{m,p} = \sum\nolimits_{i \in I_m} u_{i,\tau+k_p} u_{i,\ell_p}, \quad p = 1, \ldots, (T-\tau)\tau, \tag{40}$$

$$\boldsymbol{\zeta}^{[4]}_{m,r} = \sum\nolimits_{i \in I_m} u_{i,r} \left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right), \quad r = 1, \ldots, T, \tag{41}$$

we have

$$\boldsymbol{\eta}_{m,1} = \begin{bmatrix} \zeta^{[1]}_{m,1} & \cdots & \zeta^{[1]}_{m,T} \end{bmatrix}',$$

$$\boldsymbol{\eta}_{m,2} = \begin{bmatrix} \zeta^{[2]}_{m,1} & \cdots & \zeta^{[2]}_{m,\tau^2} \end{bmatrix}' - N_m \upsilon \mathbf{i}_\tau,$$

$$\boldsymbol{\eta}_{m,3} = \begin{bmatrix} \zeta^{[3]}_{m,1} & \cdots & \zeta^{[3]}_{m,(T-\tau)\tau} \end{bmatrix}',$$

$$\boldsymbol{\eta}_{m,4} = \begin{bmatrix} \left(\boldsymbol{\zeta}^{[4]}_{m,1}\right)' & \cdots & \left(\boldsymbol{\zeta}^{[4]}_{m,T}\right)' \end{bmatrix}'.$$

Assuming that the shocks are independent across clusters, we postulate that the central limit theorem applies to the random sequence $\{\boldsymbol{\eta}_m : m = 1, \ldots, M_N\}$ so that $\frac{1}{\sqrt{M_N}} \sum_{m=1}^{M_N} \boldsymbol{\eta}_m \xrightarrow{d} N\left(\mathbf{0}_{\mathcal{T}}, \mathbf{V}_\eta\right)$, where

$\mathbf{V}_\eta = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_m \boldsymbol{\eta}_m'$. Hence, we obtain

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{e}_i = \sqrt{\frac{M_N}{N}}\frac{1}{\sqrt{M_N}}\sum_{m=1}^{M_N}\boldsymbol{\eta}_m \xrightarrow{d} N\left(\mathbf{0}_\mathcal{T}, \mathbf{V}_\eta/G\right),$$

and so it follows that $\mathbf{V}_e = \mathbf{V}_\eta/G$. Note that

$$\mathbf{V}_\eta = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{V}_{13} & \mathbf{V}_{14} \\ \mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{V}_{23} & \mathbf{V}_{24} \\ \mathbf{V}_{31} & \mathbf{V}_{32} & \mathbf{V}_{33} & \mathbf{V}_{34} \\ \mathbf{V}_{41} & \mathbf{V}_{42} & \mathbf{V}_{43} & \mathbf{V}_{44} \end{bmatrix}, \tag{42}$$

with

$$\mathbf{V}_{k\ell} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,k}\boldsymbol{\eta}_{m,\ell}', \quad k,\ell = 1,\ldots,4.$$

Since $\mathbf{V}_\eta$ is symmetric, we have $\mathbf{V}_{k\ell} = \mathbf{V}_{\ell k}'$ for all $k,\ell = 1,\ldots,4$. We provide expressions for the matrices $\mathbf{V}_{k\ell}$ for $k,\ell = 1,\ldots,4$ and $k \geq \ell$, under a set of suitable assumptions that we describe next.

Assumption 3 **(i)** *For all $r,s = 1,\ldots,T$, as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,r}^{[1]}\zeta_{m,s}^{[1]} \xrightarrow{p} v_c \mathbf{1}_{[r=s]}$.* **(ii)** *For all $p,q = 1,\ldots,\tau^2$, as $N \to \infty$,*

$$\frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,p}^{[2]}\zeta_{m,q}^{[2]} \xrightarrow{p} \kappa_c^{[1]}\mathbf{1}_{[k_p=\ell_p=k_q=\ell_q]} + \kappa_c^{[2]}\left(\mathbf{1}_{[k_p=k_q\neq\ell_p=\ell_q]} + \mathbf{1}_{[k_p=\ell_q\neq\ell_p=k_q]}\right) + \kappa_c^{[3]}\mathbf{1}_{[k_p=\ell_p\neq k_q=\ell_q]}.$$

(iii) *For all $p,q = 1,\ldots,(T-\tau)\tau$, as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,p}^{[3]}\zeta_{m,q}^{[3]} \xrightarrow{p} \kappa_c^{[2]}\mathbf{1}_{[p=q]}$.* **(iv)** *For all $r,s = 1,\ldots,T$, as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\zeta}_{m,r}^{[4]}\left(\boldsymbol{\zeta}_{m,s}^{[4]}\right)' \xrightarrow{p} \mathbf{C}_{44}\mathbf{1}_{[r=s]}$.*

Assumption 3(i) is similar to Assumption 1(ii), except that it applies to the aggregate cluster shocks $\zeta_{m,r}^{[1]} = \sum_{i\in I_m} u_{i,r}$, and states that, as $N \to \infty$, the limiting cross-sectional variance of $\zeta_{m,r}^{[1]}$ is equal to v_c, while the limiting cross-sectional correlation between $\zeta_{m,r}^{[1]}$ and $\zeta_{m,s}^{[1]}$ vanishes when $r \neq s$.

Assumptions 3(ii) and 3(iii) are statements about the limiting cross-sectional shock moments of total order equal to four. Specifically, Assumption 3(ii) states that, for $r, r' = 1\ldots,\tau$ and $s, s' = 1,\ldots,\tau$, the limiting cross-sectional correlation between the aggregate cluster shock products $\sum_{i\in I_m} u_{i,r}u_{i,s}$ and $\sum_{i\in I_m} u_{i,r'}u_{i,s'}$ is equal to $\kappa_c^{[1]}$ when $r = s = r' = s'$; equal to $\kappa_c^{[2]}$ when $\{r,s\} = \{r',s'\}$ but $r \neq s$; ; equal to $\kappa_c^{[3]}$ when $r = s$ and $r' = s'$ but $r \neq r'$; and equal to 0 otherwise, as $N \to \infty$. Assumption

3(iii) states that, for $r, r' = \tau+1, \ldots, T$ and $s, s' = 1, \ldots, \tau$, the limiting cross-sectional correlation between the aggregate cluster shock products $\sum_{i \in I_m} u_{i,r} u_{i,s}$ and $\sum_{i \in I_m} u_{i,r'} u_{i,s'}$ is equal to $\kappa_c^{[2]}$ when $r = r'$ and $s = s'$; and equal to 0 otherwise, as $N \to \infty$. In this sense, we interpret $\kappa_c^{[1]}$, $\kappa_c^{[2]}$, and $\kappa_c^{[3]}$ as fourth-moment parameters.

In light of Assumption 1(i), which states that the betas $\boldsymbol{\beta}_i$ and the shocks $u_{i,t}$ are cross-sectionally uncorrelated in the limit, for $r, s = 1, \ldots, T$, Assumption 3(iv) states that $\boldsymbol{\zeta}_{m,r}^{[4]} = \sum_{i \in I_m} u_{i,r} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)$ and $\boldsymbol{\zeta}_{m,s}^{[4]} = \sum_{i \in I_m} u_{i,s} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)$ are cross-sectionally uncorrelated, as $N \to \infty$, with $r \neq s$. Moreover, Assumption 3(iv) states that the limiting cross-sectional variance of $\boldsymbol{\zeta}_{m,r}^{[4]}$ is equal to $\mathbf{C}_{44}$, as $N \to \infty$.

Assumption 4 **(i)** *For all* $r = 1, \ldots, T$ *and* $p = 1, \ldots, \tau^2$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \zeta_{m,r}^{[1]} \zeta_{m,p}^{[2]} \xrightarrow{p} \xi_c \mathbf{1}_{[r=k_p=\ell_p]}$. **(ii)** *For all* $r = 1, \ldots, T$ *and* $p = 1, \ldots, (T-\tau)\tau$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \zeta_{m,r}^{[1]} \zeta_{m,p}^{[3]} \xrightarrow{p} 0$. **(iii)** *For all* $r, s = 1, \ldots, T$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \boldsymbol{\zeta}_{m,r}^{[4]} \zeta_{m,s}^{[1]} \xrightarrow{p} \mathbf{c}_{41} \mathbf{1}_{[r=s]}$. **(iv)** *For all* $p = 1, \ldots, \tau^2$ *and* $q = 1, \ldots, (T-\tau)\tau$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \zeta_{m,p}^{[2]} \zeta_{m,q}^{[3]} \xrightarrow{p} 0$. **(v)** *For all* $p = 1, \ldots, \tau^2$ *and* $r = 1, \ldots, T$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \boldsymbol{\zeta}_{m,r}^{[4]} \zeta_{m,p}^{[2]} \xrightarrow{p} \mathbf{c}_{42} \mathbf{1}_{[r=k_p=\ell_p]}$. **(vi)** *For all* $r = 1, \ldots, T$ *and* $p = 1, \ldots, (T-\tau)\tau$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} \boldsymbol{\zeta}_{m,r}^{[4]} \zeta_{m,p}^{[3]} \xrightarrow{p} \mathbf{0}_K$.

Assumptions 4(i) and 4(ii) are statements about the limiting cross-sectional shock moments of total order equal to three. Specifically, they state that, for $r, s = 1, \ldots, T$ and $s' = 1, \ldots, \tau$, the cross-sectional average of the product $\left(\sum_{i \in I_m} u_{i,r}\right) \left(\sum_{i \in I_m} u_{i,s} u_{i,s'}\right)$ converges to ξ_c when $r = s = s'$, and to 0 otherwise. In this sense, we interpret ξ_c as a third-moment parameter.

Assumptions 4(iii), 4(v), and 4(vi) are statements about the limiting cross-sectional relationship between shocks and betas. Specifically, Assumption 4(iii) states that, for $r, s = 1, \ldots, T$, $\boldsymbol{\zeta}_{m,r}^{[4]} = \sum_{i \in I_m} u_{i,r} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)$ and the aggregate cluster shock $\zeta_{m,s}^{[1]} = \sum_{i \in I_m} u_{i,s}$ are cross-sectionally uncorrelated for $r \neq s$, while their limiting cross-sectional correlation converges to $\mathbf{c}_{41}$ for $r = s$, as $N \to \infty$. Moreover, Assumptions 4(v) and 4(vi) state that, for $r, s = 1, \ldots, T$ and $s' = 1, \ldots, \tau$, the limiting cross-sectional correlation of $\sum_{i \in I_m} u_{i,r} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)$ and $\sum_{i \in I_m} u_{i,s} u_{i,s'}$ converges to $\mathbf{c}_{42}$ when $r = s = s'$, and to $\mathbf{0}_K$ otherwise.

Assumption 4(iv) states that, for $r = \tau+1, \ldots, T$, $r' = 1, \ldots, \tau$ and $s, s' = 1, \ldots, \tau$, the aggregate cluster shock products $\sum_{i \in I_m} u_{i,r} u_{i,s}$ and $\sum_{i \in I_m} u_{i,r'} u_{i,s'}$ are cross-sectionally uncorrelated, as $N \to \infty$.

Assumption 5 **(i)** *As* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} N_m^2 \to H$. **(ii)** *For all* $r = 1, \ldots, T$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} N_m \zeta_{m,r}^{[1]} \xrightarrow{p} 0$. **(iii)** *For all* $p = 1, \ldots, \tau^2$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} N_m \zeta_{m,p}^{[2]} \xrightarrow{p} G^2 v \mathbf{1}_{[k_p=\ell_p]}$. **(iv)** *For all* $p = 1, \ldots, (T-\tau)\tau$, *as* $N \to \infty$, $\frac{1}{M_N} \sum_{m=1}^{M_N} N_m \zeta_{m,p}^{[3]} \xrightarrow{p} 0$. **(v)** *For all* $r = 1, \ldots, T$,

$\frac{1}{M_N}\sum_{m=1}^{M_N} N_m \boldsymbol{\zeta}_{m,r}^{[4]} \xrightarrow{p} \mathbf{0}_K$.

Assumption 5(i)-(iv) are statements about the limiting second moment of the cluster size N_m and its interaction with aggregate shock terms $\zeta_{m,r}^{[k]}$, $k = 1, 2, 3$, and 4.

Assumption 5(i) states the average squared cluster size N_m^2 converges to H, as $N \to \infty$.

For $r = 1, \ldots, T$, Assumption 5(ii) states that the cluster size N_m and the aggregate cluster shock $\zeta_{m,r}^{[1]} = \sum_{i \in I_m} u_{i,r}$ are cross-sectionally uncorrelated, as $N \to \infty$.

For $r = 1, \ldots, T$ and $s = 1, \ldots, \tau$, Assumptions 5(iii) and 5(iv) are statements about the limiting cross-sectional uncorrelatedness between the cluster size N_m and the aggregate cluster shock product $\sum_{i \in I_m} u_{i,r} u_{i,s}$, as $N \to \infty$. For $r \neq s$, these assumptions amount to $\frac{1}{M_N}\sum_{m=1}^{M_N} N_m \left(\sum_{i \in I_m} u_{i,r} u_{i,s}\right) \xrightarrow{p} 0$. For $r = s$, according to Assumption 5(iii), the probability limit of $\frac{1}{M_N}\sum_{m=1}^{M_N} N_m \left(\sum_{i \in I_m} u_{i,r}^2\right)$ is equal to the product of the limit $\frac{1}{M_N}\sum_{m=1}^{M_N} N_m$ and the probability limit of $\frac{1}{M_N}\sum_{m=1}^{M_N} \left(\sum_{i \in I_m} u_{i,r}^2\right)$. Note that the former is equal to $\lim \frac{N}{M_N} = G$ while the latter is equal to p-lim $\frac{1}{M_N}\sum_{i=1}^{N} u_{i,r}^2 = Gv$, according to Assumption 1(ii).

Finally, in light of Assumption 1(i), for $r = 1, \ldots, T$, Assumption 5(v) states that the cluster size N_m and $\boldsymbol{\zeta}_{m,r}^{[4]} = \sum_{i \in I_m} u_{i,r}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta\right)$ are cross-sectionally uncorrelated, as $N \to \infty$.

Proposition 8 *Under Assumptions 3, 4, and 5, the variance-covariance matrix $\mathbf{V}_\eta$, which is defined in (42), is determined by the following matrices:*

$$\mathbf{V}_{11} = v_c \mathbf{I}_T,$$

$$\mathbf{V}_{21} = \xi_c \boldsymbol{\mathcal{M}}, \quad \mathbf{V}_{22} = \kappa_c^{[1]} \boldsymbol{\mathcal{K}}^{[1]} + \kappa_c^{[2]} \boldsymbol{\mathcal{K}}^{[2]} + \kappa_c^{[3]} \boldsymbol{\mathcal{K}}^{[3]} - (2G^2 - H) v^2 \mathbf{i}_\tau \mathbf{i}_\tau',$$

$$\mathbf{V}_{31} = \mathbf{0}_{(T-\tau)\tau \times T}, \quad \mathbf{V}_{32} = \mathbf{0}_{(T-\tau)\tau \times \tau^2}, \quad \mathbf{V}_{33} = \kappa_c^{[2]} \mathbf{I}_{(T-\tau)\tau},$$

$$\mathbf{V}_{41} = \mathbf{I}_T \otimes \mathbf{c}_{41}, \quad \mathbf{V}_{42} = \boldsymbol{\mathcal{M}}' \otimes \mathbf{c}_{42}, \quad \mathbf{V}_{43} = \mathbf{0}_{TK \times (T-\tau)\tau}, \quad \mathbf{V}_{44} = \mathbf{I}_T \otimes \mathbf{C}_{44}.$$

The matrices $\boldsymbol{\mathcal{K}}^{[1]}$, $\boldsymbol{\mathcal{K}}^{[2]}$, $\boldsymbol{\mathcal{K}}^{[3]}$, and $\boldsymbol{\mathcal{M}}$ are defined as follows. For each positive integer p, let $\ell_p \in \{1, \ldots, \tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). The $\tau^2 \times \tau^2$ matrix $\boldsymbol{\mathcal{K}}^{[1]}$ is defined by $\boldsymbol{\mathcal{K}}^{[1]}(p, q) = \mathbf{1}_{[k_p = \ell_p = k_q = \ell_q]}$, for $p, q = 1, \ldots, \tau^2$. The $\tau^2 \times \tau^2$ matrix $\boldsymbol{\mathcal{K}}^{[2]}$ is defined by $\boldsymbol{\mathcal{K}}^{[2]}(p, q) = \mathbf{1}_{[k_p = k_q \neq \ell_p = \ell_q]} + \mathbf{1}_{[k_p = \ell_q \neq \ell_p = k_q]}$, for $p, q = 1, \ldots, \tau^2$. The $\tau^2 \times \tau^2$ matrix $\boldsymbol{\mathcal{K}}^{[3]}$ is defined by $\boldsymbol{\mathcal{K}}^{[3]}(p, q) = \mathbf{1}_{[k_p = \ell_p \neq k_q = \ell_q]}$, for $p, q = 1, \ldots, \tau^2$. Finally, the $\tau^2 \times T$ matrix $\boldsymbol{\mathcal{M}}$ is defined by $\boldsymbol{\mathcal{M}}(p, r) = \mathbf{1}_{[k_p = \ell_p = r]}$, for $p = 1, \ldots, \tau^2$ and $r = 1, \ldots, T$.

In light of Proposition 8, one needs N-consistent estimators of the second-moment parameter v_c, the third-moment parameter ξ_c, the fourth-moment parameters $\kappa_c^{[1]}$, $\kappa_c^{[2]}$, $\kappa_c^{[3]}$, the $K \times 1$ vectors $\mathbf{c}_{41}$ and $\mathbf{c}_{42}$, and the $K \times K$ matrix $\mathbf{C}_{44}$, in order to obtain an N-consistent estimator of $\mathbf{V}_\eta$. We provide such estimators in Lemmas 10, 11, 12, and 13 in the Appendix. The following theorem describes the N-consistent estimator of the variance-covariance matrix $\mathbf{V}_\gamma$ that we obtain by combining these estimators with Proposition 8.

Theorem 9 *Under Assumptions 1, 2, 3, 4, and 5,*

$$\widehat{\mathbf{V}}_\gamma = \widehat{\mathbf{\Omega}}^{-1}\widehat{\mathbf{\Pi}}\widehat{\mathbf{V}}_e\widehat{\mathbf{\Pi}}'\widehat{\mathbf{\Omega}}^{-1} \tag{43}$$

is an N-consistent estimator of the variance-covariance matrix $\mathbf{V}_\gamma$, where $\widehat{\mathbf{\Omega}} = \frac{1}{N}\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}$, $\widehat{\mathbf{\Pi}}$ is the sample analogue of $\mathbf{\Pi}$ (see equation (70)), $\widehat{\mathbf{V}}_e = \frac{\widehat{\mathbf{V}}_\eta}{\widehat{G}}$, $\widehat{G} = \frac{N}{M_N}$, $\widehat{\mathbf{V}}_\eta$ is given by

$$\widehat{\mathbf{V}}_\eta = \begin{bmatrix} \widehat{v}_c\mathbf{I}_T & \widehat{\xi}_c\boldsymbol{\mathcal{M}}' & \mathbf{0}_{T\times(T-\tau)\tau} & \mathbf{I}_T \otimes \widehat{\mathbf{c}}_{41}' \\ \widehat{\xi}_c\boldsymbol{\mathcal{M}} & \widehat{\kappa}_c^{[1]}\boldsymbol{\mathcal{K}}^{[1]} + \widehat{\kappa}_c^{[2]}\boldsymbol{\mathcal{K}}^{[2]} + \widehat{\kappa}_c^{[3]}\boldsymbol{\mathcal{K}}^{[3]} - (2\widehat{G}^2 - \widehat{H})\widehat{v}^2\mathbf{i}_\tau\mathbf{i}_\tau' & \mathbf{0}_{\tau^2\times(T-\tau)\tau} & \boldsymbol{\mathcal{M}} \otimes \widehat{\mathbf{c}}_{42}' \\ \mathbf{0}_{(T-\tau)\tau\times T} & \mathbf{0}_{(T-\tau)\tau\times\tau^2} & \widehat{\kappa}_c^{[2]}\mathbf{I}_{(T-\tau)\tau} & \mathbf{0}_{(T-\tau)\tau\times TK} \\ \mathbf{I}_T \otimes \widehat{\mathbf{c}}_{41} & \boldsymbol{\mathcal{M}}' \otimes \widehat{\mathbf{c}}_{42} & \mathbf{0}_{TK\times(T-\tau)\tau} & \mathbf{I}_T \otimes \widehat{\mathbf{C}}_{44} \end{bmatrix}, \tag{44}$$

$\widehat{H} = \frac{1}{M_N}\sum_{m=1}^{M_N} N_m^2$, $\widehat{v}_c$ is provided by Lemma 10, $\widehat{\xi}_c$ is provided by Lemma 11, $\widehat{\kappa}_c^{[1]}$, $\widehat{\kappa}_c^{[2]}$, $\widehat{\kappa}_c^{[3]}$ are provided by Lemma 12, and $\widehat{\mathbf{c}}_{41}$, $\widehat{\mathbf{c}}_{42}$, $\widehat{\mathbf{C}}_{44}$ are provided by Lemma 13.

Combining Theorems 7 and 9 we can readily obtain statistics for testing asset pricing model implications of interest. The joint hypothesis $[\ \alpha \quad \boldsymbol{\lambda}'\ ]' = [\ 0 \quad \overline{\mathbf{f}}_2'\ ]'$ can be tested using $\chi^2(\boldsymbol{\gamma}) = N(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})'\widehat{\mathbf{V}}_\gamma^{-1}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})$ which asymptotically follows a χ^2 distribution with $K+1$ degrees of freedom. Denoting by $\widehat{v}_\alpha$ the $(1,1)$ element of $\widehat{\mathbf{V}}_\gamma$, we can use $t(\alpha) = \frac{\widetilde{\alpha}}{\sqrt{\widehat{v}_\alpha/N}}$, which asymptotically follows a standard normal distribution, to test the hypothesis $\alpha = 0$. Similarly, for $k = 1, \ldots, K$, denoting by $\widehat{v}_{\lambda,k}$ the $(k+1, k+1)$ element of $\widehat{\mathbf{V}}_\gamma$, we can use $t(\lambda_k) = \frac{\widetilde{\lambda}_k - \overline{f}_{k,2}}{\sqrt{\widehat{v}_{\lambda,k}/N}}$, which also has a standard normal asymptotic distribution, to test the hypothesis $\lambda_k = \overline{f}_{k,2}$.

3 Monte Carlo Simulation Evidence

In this section, we use a Monte Carlo simulation in order to illustrate the importance of the EIV correction in terms of bias reduction in the cross-sectional alpha and risk premia estimates and investigate the finite sample behavior of the χ^2 and t statistics based on EIV-corrected estimator $\widetilde{\boldsymbol{\gamma}}$ and the estimator $\widehat{\mathbf{V}}_\gamma$ of its asymptotic variance-covariance matrix $\mathbf{V}_\gamma$. The simulation is designed to follow closely the empirical investigation that we conduct in Section 4.

The simulation setup is described next. We calibrate the betas as well as the variances of the idiosyncratic shocks so that we can generate the returns of individual assets according to the data generating process (1). To obtain representative values of betas and idiosyncratic shock variances, while reducing beta estimation error, we consider all stocks in the CRSP universe from 1980 to 2012 with price above \$3 and select the 2,000 among those that have the longest time series histories. The shortest sample, among the 2,000 selected stocks, consists of 246 months, and, therefore, the effect of beta estimation error can be considered minimal for the purpose of our simulation.

We consider two linear asset pricing models widely used in the literature: the single-factor CAPM (Sharpe (1964), Lintner (1965), Mossin (1966)) and the three-factor model of Fama and French (1993). The factor realizations are taken as given and kept fixed throughout. That is, consistent with out theoretical developments, the simulation tests that we perform are conditional on the factor realizations. For each of the 2,000 individual stocks, we estimate the pair of beta and idiosyncratic shock variance by regressing the stock return, in excess of the one-month T-bill rate, on a constant and the vector of factor realizations.[21] The joint estimation of betas and shock variances is important as it can capture any potential cross-sectional dependence between betas and idiosyncratic shocks through higher moments.

Two critical aspects of the simulation are, first, the number of clusters in the stock universe and, second, the correlation structure among stock returns within clusters. For simplicity, in our simulations, we assume that all clusters have the same size and correlations within clusters are constant. In the first simulation exercise, that focuses on bias reduction, we set the number of clusters, M_N, equal to 50 and the pairwise correlation ρ, within each cluster, equal to 0.10. In the second simulation, that focuses on the finite sample properties of the test statistics, we let the number of clusters M_N take the values 50 and 100 and assume that, within each cluster, pairwise correlations are equal to ρ which is assumed to take three values: 0, 0.10, and 0.20. In the empirical application of Section 4, we consider clustering based on the 49 industries according to the SIC classification by Kenneth French.[22] Using this classification, we find that the average correlation within industries is roughly 0.10 based on an industry residual model for the shocks, in the spirit of Ang, Liu, and Schwarz (2010) (see their Appendix F.2). Therefore, the range of correlation values that we consider is empirically relevant.

[21]We use the data for the factors MKT (market excess return), SMB (small size minus big size), and HML (high book to market minus low book to market) from Ken French's data library.

[22]The classification is available at http://mba.tuck.dartmouth.edu/pages /faculty/ken.french/Data-Library/det-49-ind-port.html.

In our first simulation exercise, to emphasize the significant role of the EIV correction in reducing the bias in the estimation of $\boldsymbol{\gamma}$, we compare the bias in the $\boldsymbol{\gamma}$ estimates with and without the EIV correction. Both estimation and testing periods consist of 60 months and, to provide a comprehensive picture, we repeat the exercise over twelve testing periods from 1951-1955 to 2006-2010 based on the observed factor realizations. The idiosyncratic shocks are assumed to follow a normal distribution. The results, reported in Table 1, clearly illustrate the bias reduction gains achieved by using the EIV correction. Specifically, for the CAPM, the absolute bias of the EIV-corrected α estimates, averaged over the twelve periods, is 2.5 annualized basis points. In contrast, without the EIV correction, the average absolute bias is 240.6 annualized basis points. The average absolute bias of the λ_1 estimates is 3.0 and 275.4 annualized basis points, with and without the EIV correction, respectively. Similar results hold for the Fama-French model, for which the average absolute biases, with and without the EIV correction, are: 12.3 and 385.8 annualized basis points for α, 16.1 and 315.2 annualized basis points for λ_1, 15.8 and 276.0 annualized basis points for λ_2, and 15.5 and 296.6 annualized basis points for λ_3.

To gauge the importance of the assumption of normally distributed shocks, we repeat the simulation described above using a Student-t distribution with 6 degrees of freedom. The results, reported in Table 8 in the Online Appendix, are almost identical and, hence, our conclusions regarding the importance of the EIV correction for bias reduction in the estimation of cross-sectional alpha and risk premia are robust to the normality assumption.

In our second simulation exercise, we focus on the behavior of the EIV-corrected estimator $\widetilde{\boldsymbol{\gamma}}$ and the estimator $\widehat{\mathbf{V}}_\gamma$ of its asymptotic variance-covariance matrix $\mathbf{V}_\gamma$, as reflected in the finite sample properties of the χ^2 and t statistics. The beta estimation and the testing periods cover the years 2001 to 2005 and 2006 to 2010, consisting of $\tau = 60$ and $T - \tau = 60$ observations, respectively. Using the given factor realizations and the calibrated pairs of betas and idiosyncratic shock variances, we simulate individual stock returns using the data generating process (1) for $t = 1, \ldots, 120$. First, using the simulated data corresponding to the first five-year period ($t = 1, \ldots, 60$), we estimate the betas and then, using the simulated data corresponding to the second five-year period ($t = 61, \ldots, 120$), we test the implications of the asset pricing model employing the χ^2 and the t statistics described in Section 2. We consider three nominal levels of significance, 1%, 5%, and 10%, and compute the corresponding empirical rejection frequencies. The simulation exercise, which is based on $10,000$ Monte Carlo repetitions, is first performed for shocks following a normal distribution and the results are reported in Table 2. The empirical rejection frequencies are very close to the corresponding nominal levels of significance for both the χ^2 and the

t statistics, illustrating that our estimators perform well for empirically relevant sample sizes. The simulation is repeated for shocks following a Student-t distribution with 6 degrees of freedom and the results, reported in Table 9 in the Online Appendix, are again almost identical. The conclusions hold for both the CAPM and the Fama-French model and under both distributional assumptions.

4 Empirical Application

In this section, we present our empirical results illustrating the EIV correction method developed in Section 2. In our investigation, we mainly focus on two commonly used linear factor models: the single-factor CAPM and the three factor Fama-French model. The data on the three factors, market excess return (MKT), small-minus-big spread portfolio return (SMB), and high-minus-low spread portfolio return (HML), are taken from Kenneth French's website. We use monthly stock data from 1946 to 2012 from the CRSP universe and apply two filters in the selection of stocks. First, we require that a stock has a Standard Industry Classification (SIC) code[23] and, second, we keep a stock in our sample only for the months in which its price is at least 3 dollars. The resulting data set consists of 27,234 individual stocks, with a large cross section of available stocks in any given month. It is worth noting that most of the individual stocks have short histories. After applying the two filters, we find that the average available history is 111 months. Our econometric approach, designed for large N and fixed T, is suitable for handling these characteristics of the available large cross-sectional stock data sets.

Our beta estimation periods consist of five years ($\tau = 60$ months) while we consider testing periods consisting of six years ($T - \tau = 72$ months) and three years ($T - \tau = 36$ months), resulting in 10 and 20 non-overlapping testing periods, respectively. In each case, our cross section consists of all stocks with full histories over both the estimation and testing periods. Our empirical evidence consists of (i) estimates of $\boldsymbol{\gamma} = [\ \alpha \quad \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \quad \cdots \quad \lambda_K\ ]'$ is the vector of ex-post factor risk premia for the CAPM and the Fama-French model, for which we have $K = 1$ and $K = 3$, respectively, and (ii) test statistics for the various implications of each model and their corresponding p-values.

Tables 3 and 5 present the estimation results for the 10 six-year testing periods and the 20 three-year testing periods, respectively. We report, in annualized percentages, the estimates of $\boldsymbol{\gamma}$ with and without the EIV correction as well as the corresponding realized factor averages over the testing period $\bar{\mathbf{f}}_2$. Given that all factors under consideration are portfolio spreads, both linear factor models imply that $\alpha = 0$

[23]We follow the 49 industry definitions established by Kenneth French, available at his website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_49_ind_port.html`.

and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$.

A few observations regarding the estimation results are in order. It was shown, in our simulation experiments, that the EIV correction is important in terms of reducing the bias in the estimation of the cross-sectional alpha and risk premia. Our empirical results also demonstrate that the EIV correction matters in the sense that it results in different cross-sectional alpha and risk premia estimates and the differences are economically significant. Specifically, the estimates obtained without the EIV correction are typically pulled towards zero relative to the EIV-corrected estimates, which is consistent with our theoretical arguments. In the context of the CAPM, the EIV-corrected estimates of α, for the six-year testing periods, are frequently sizeable as seen in Table 3: in five out of 10 cases they are higher than 5 per cent per annum in absolute value. The EIV-corrected estimates of the ex-post market risk premium also differ from the market excess return realization and even have the opposite sign in two of out of 10 cases. Similarly, in the context of the Fama-French model, we also obtain sizeable EIV-corrected estimates of α as documented in Table 3: in five out of 10 cases they are higher than 5 per cent per annum in absolute value. The results for the three-year testing periods, presented in Table 5, point to the same conclusions. For both the CAPM and the Fama-French model, we again obtain sizeable α estimates. In particular, out of 20 cases, 13 and 12 EIV-corrected estimates of α are higher than 5 per cent per annum in absolute value for the CAPM and the Fama-French model, respectively. The statistical significance of these estimates is gauged through suitable test statistics.

Using the asymptotic distribution of the EIV-corrected estimator $\widetilde{\boldsymbol{\gamma}}$, established in Theorem 7, and the N-consistent estimator of its asymptotic variance-covariance matrix $\widehat{\mathbf{V}}_\gamma$, obtained in subsection 2.4, we can formally test the various implications of each model. We present the test statistics and the corresponding p-values for the 10 non-overlapping six-year testing periods and the 20 non-overlapping three-year testing periods in Tables 4 and 6, respectively. We use χ^2 statistics to test for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics to test the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \dots, K$. In our baseline inference results, we use clustering based on the 49 industries according to the SIC classification by Kenneth French. As robustness checks, we also consider clustering with respect to size and book-to-market, using 50 and 30 clusters, and report the results in the Online Appendix.

When six-year testing periods are considered, the joint hypothesis that $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ is rejected for both the CAPM and the Fama-French model, based on the χ^2 statistics. Specifically, according to Table 4, in 9 and 10, out of 10 cases, the corresponding p-values are below 5 per cent for the CAPM and the Fama-French model, respectively. This finding is not very surprising given that these two models have been rejected in the literature using portfolios of stocks. It is worth noting, however, that the hypothesis $\alpha = 0$ is not rejected in 4 and 5 out of the 10 cases for the CAPM and the Fama-French model, respectively, at the conventional 5 per cent level of significance. Moreover, according to our t statistics, in the context of the CAPM, we cannot reject the hypothesis that the ex-post market risk premium is equal to the average market excess returns in 5 out of 10 cases. This finding indicates that the explanatory power of the market beta may vary over time. A similar conclusion is reached for the Fama-French model. We cannot reject the hypothesis that the ex-post price of risk on the SMB (HML) beta is equal to the average SMB (HML) factor realizations in 6 (8) out of 10 cases.

We repeat the above tests using 20 non-overlapping three-year testing periods and report the results in Table 6. The three-year testing period evidence is consistent with the above findings based on the six-year testing periods. The joint hypothesis, based on the χ^2 statistics, is rejected in 18 and 16 out of 20 cases for the CAPM and the Fama-French model, respectively. An important observation, in the context of the Fama-French model, is in order. The hypothesis that the ex-post price of risk on the HML beta is equal to the average HML factor realization is not rejected in 17 out of 20 cases. At the same time, we cannot reject the analogous hypothesis for the MKT and SMB factors in 9 and 12 out of 20 cases.

To illustrate the robustness of our inference results, we also use clustering across two different dimensions besides industry: size and book-to-market. We consider 50 and 30 clusters for each case and report the results in the Online Appendix. The inference results for size clustering and six-year testing periods are reported in Tables 10 and 11, for 50 and 30 clusters, respectively. The corresponding three-year testing period results are reported in Tables 12 and 13, for 50 and 30 clusters, respectively. The results in Tables 10, 11, 12, and 13 are consistent with the results based on industry clustering reported in Tables 3 and 5. Specifically, in the case of six-year testing periods reported in Tables 12 and 13, 65 and 60 p-values, out of 70, point to the same inference at the conventional 10 and 5 per cent levels of significance, respectively. When we use clustering with respect to book-to-market, due to lack of book-to-market data, our sample starts in 1969 and 1966 for the six- and three-year testing periods, respectively, and also contains fewer stocks than the original sample. As a result, the estimates of γ based on this sample are slightly different than the estimates reported in Tables 3 and 5. We report the estimates of γ based on stocks with available book-to-market data in Tables 14 and 17, for six- and three-year testing periods, respectively. Tables 15 and 16 contain the six-year testing period inference results for 50 and 30 book-to-market clusters, respectively. Finally, in Tables 18 and 19, we report the three-year testing period inference results for 50 and 30 book-to-market clusters, respectively. As in the case of size clustering, the results based on book-to-market clustering remain consistent with the industry clustering results.

Recall that, according to the results discussed above, in most cases we cannot reject the hypothesis that the ex-post price of risk on the HML beta is equal to the average HML factor realization. This evidence suggests that HML may be a useful factor for pricing large cross sections of individual stocks besides portfolios sorted on the book-to-market ratio. To shed some light into this dimension, we conduct one additional exercise. Using five-year beta estimation periods from 1946 to 2012, we obtain the EIV-corrected cross-sectional estimates of the CAPM alpha at the monthly, quarterly and semiannual frequency and then run a time-series regression of these estimates on the average realizations of the SMB and HML factors over the corresponding time period. As robustness checks, we also include the momentum factor of Carhart (1997) (MOM) and the liquidity factor of Pastor and Stambaugh (2003) (LIQ).[24] The monthly, quarterly, and semiannual series consist of 744, 248, and 124 observations, respectively. The liquidity factor data are available only from 1968 to 2011. We examine six different factor combinations: (i) SMB, (ii) HML, (iii) SMB and HML, (iv) SMB, HML, and MOM, (v) SMB, HML, and LIQ, and (vi) SMB, HML, MOM, and LIQ. In Table 7, we report the results including point estimates and p-values for the intercept and the slopes as well as the adjusted R^2, for each time-series regression. We employ the Newey and West (1987) method to compute standard errors, with no prewhitening, and select the optimal lag according to Newey and West (1994) .

The evidence from the time-series regressions of the EIV-corrected CAPM alpha estimates on the various factor combinations points to one robust finding. Whenever included in the factor combination, the HML factor is found to be statistically significant and this result is consistent across different frequencies, with all associated p-values being equal to zero (up to three decimal points). This finding suggests that the HML might be capturing some dimension missing by the CAPM. On the contrary, the SMB, MOM, and LIQ factor do not appear to explain the time variation of the EIV-corrected CAPM alpha estimates. The smallest p-value for the SMB factor is 0.125 at the monthly frequency with much higher p-values at the quarterly and semiannual frequencies. The smallest overall p-value for the MOM factor is 0.423 while the p-values for the LIQ factor are also quite high, especially at the quarterly and semiannual frequencies. Collectively, our empirical evidence suggests that the HML factor can be a useful factor for pricing individual stocks in large cross sections.

[24] The momentum factor and the liquidity factor data are obtained from the websites of Kenneth French and Lubos Pastor, respectively.

5 Conclusion

This paper contributes to the empirical asset pricing literature by providing a method for estimating ex-post risk premia and developing novel associated tests for evaluating linear factor models using individual stock data over short time horizons. When the size of the cross section N is large and the time series length T is small and fixed, applying the traditional two-pass cross-sectional regression (CSR) approach is problematic since the orthogonality condition in the second-pass regression of returns on estimated betas is violated due to the error-in-variables problem. As a result, the standard two-pass CSR risk premia estimator is not N-consistent. Employing the regression-calibration approach, we develop a modification of the two-pass CSR method and correct the betas obtained in the first pass so that the desired orthogonality condition in the second pass is satisfied. The resulting risk premia estimator is N-consistent and, importantly, has a convenient OLS form that we exploit to obtain its asymptotic distribution, as N tends to infinity. Moreover, employing a cluster structure for idiosyncratic shock correlations, we obtain an estimator of the asymptotic variance-covariance matrix of our risk premia estimator that, in turn, allows us to build, for the first time, asset pricing tests focusing on ex-post risk premia.

We further provide simulation evidence, based on the CAPM and the Fama-French model, illustrating that there our beta correction approach leads to significant bias reduction in the cross-sectional alpha and ex-post risk premia estimates and that our asset pricing tests perform well for empirically relevant sample sizes. In our empirical application, we examine the CAPM and Fama-French model and provide some interesting findings. While we formally reject the joint hypotheses on the alpha and the risk premia for both models, our evidence suggests that the HML factor may be capturing a dimension missing by the CAPM and can be a useful factor for pricing large cross sections of individual stocks besides portfolios sorted on book-to-market.

The analysis in this paper can be extended along a number of dimensions. For instance, one can model the betas as functions of macroeconomic variables to allow for time variation. Moreover, one can relax the assumption that the return shocks are uncorrelated over time. We leave these important extensions for future work. Finally, the regression-calibration approach could be applied in other settings to deal with the error-in-variables problem.

A Appendix

A.1 Auxiliary Facts

A number of facts from matrix algebra are used frequently in the main text and/or in the subsequent proofs. We collect them here for the convenience of the reader. In terms of notation, tr denotes the trace operator, vec denotes the column-stacking operator, $\otimes$ denotes the Kronecker product, and $\odot$ denotes the Hadamard product.

(F1) For conformable matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$.

(F2) For column vectors $\mathbf{x}$ and $\mathbf{y}$, we have $\text{vec}(\mathbf{xy}') = \mathbf{y} \otimes \mathbf{x}$.

(F3) For conformable matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\text{tr}(\mathbf{A}'\mathbf{B}) = (\text{vec}(\mathbf{A}))' \text{vec}(\mathbf{B})$.

(F4) For conformable matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, we have $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) \text{vec}(\mathbf{B})$.

(F5) For conformable matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, we have $(\mathbf{AC}) \otimes (\mathbf{BD}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D})$.

(F6) For square matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A}) \text{tr}(\mathbf{B})$.

(F7) For $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\mathbf{1}'_m (\mathbf{A} \odot \mathbf{B}) \mathbf{1}_n = \text{tr}(\mathbf{AB}')$.

(F8) For $m \times 1$ vectors $\mathbf{a}$ and $\mathbf{c}$ and $n \times 1$ vectors $\mathbf{b}$ and $\mathbf{d}$, we have $(\mathbf{ab}') \odot (\mathbf{cd}') = (\mathbf{a} \odot \mathbf{c})(\mathbf{b} \odot \mathbf{d})'$.

A.2 Proofs of Proposition 4, Proposition 5, Theorem 7, and Proposition 8

Proof of Proposition 4: It follows from the probability limit in (121) and definitions (16) and (27) that

$$\widehat{\mathbf{L}}_1 = \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N \xrightarrow{p} \mathbf{L}_1.$$

Hence, using Lemma 3 and definitions (21) and (26), we obtain $\widehat{\mathbf{C}} \xrightarrow{p} \mathbf{C}$, as $N \to \infty$. ■

Proof of Proposition 5: It follows from equation (31) that

$$\widetilde{\boldsymbol{\omega}} = (\mathbf{J}_N\widehat{\mathbf{B}}\widehat{\mathbf{C}} - \mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda} + \bar{\mathbf{u}}_2. \tag{45}$$

Equation (30) states $\bar{\mathbf{r}}_2 = \widetilde{\mathbf{X}}\boldsymbol{\gamma} + \widetilde{\boldsymbol{\omega}}$ which yields

$$\widetilde{\boldsymbol{\gamma}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\bar{\mathbf{r}}_2 = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'(\widetilde{\mathbf{X}}\boldsymbol{\gamma} + \widetilde{\boldsymbol{\omega}}) = \boldsymbol{\gamma} + (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N)^{-1}(\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/N). \tag{46}$$

Hence, to complete the proof, it suffices to show that $\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/N \xrightarrow{p} \mathbf{0}_{K+1}$ and that $\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N$ converges in probability to an invertible matrix. Note that the first element of $\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}$ is $\mathbf{1}'_N\widetilde{\boldsymbol{\omega}}$ while the remaining K elements are captured by the vector $\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}$. Proceeding as in the proof of the fact that $\mathbf{1}'_N\breve{\boldsymbol{\omega}}/N \xrightarrow{p} 0$ in the proof of Lemma 2, contained in the Online Appendix, we can easily show that

$$\mathbf{1}'_N\widetilde{\boldsymbol{\omega}}/N \xrightarrow{p} 0. \tag{47}$$

Using the expression in equation (45), and following the pattern in decomposition (104), we can express the product $\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}$ as the sum of five terms as follows:

$$\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}} = \mathbf{k}^{(1)} + \mathbf{k}^{(2)} + \mathbf{k}^{(3)} + \mathbf{k}^{(4)} + \mathbf{k}^{(5)} \tag{48}$$

where

$$\mathbf{k}^{(1)} = -\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}'_N\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda},$$

$$\mathbf{k}^{(2)} = \widehat{\boldsymbol{\mu}}_\beta\mathbf{1}'_N\bar{\mathbf{u}}_2 = \frac{1}{T-\tau}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}'_N\mathbf{U}_2\mathbf{1}_{T-\tau},$$

$$\mathbf{k}^{(3)} = (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}\widehat{\mathbf{C}}\boldsymbol{\lambda},$$

$$\mathbf{k}^{(4)} = -(\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda},$$

$$\mathbf{k}^{(5)} = (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\bar{\mathbf{u}}_2 = \frac{1}{T-\tau}\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2\mathbf{1}_{T-\tau}.$$

The only difference between $\mathbf{k}^{(m)}$ and $\mathbf{h}^{(m)}$, $m = 1, \ldots, 5$ (see equations (105)-(109)) is that $\widehat{\mathbf{C}}$ replaces $\mathbf{C}$. According to Proposition 4, the probability limit of $\widehat{\mathbf{C}}$ defined in (26) is $\mathbf{C}$ defined in (21). Thus, the probability limits of $\mathbf{k}^{(m)}/N$ and $\mathbf{h}^{(m)}/N$, $m = 1, \ldots, 5$, should be identical. According to the proof of Lemma 2, which can be found in the Online Appendix, $\sum_{m=1}^{5} \mathbf{h}^{(m)}/N \xrightarrow{p} \mathbf{0}_K$, given the choice of $\mathbf{C}$ in (21). Hence, we have

$$\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}/N = \sum_{m=1}^{5} \mathbf{k}^{(m)}/N \xrightarrow{p} \mathbf{0}_K. \tag{49}$$

Combining the probability limits in (47) and (49), we obtain that $\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/N \xrightarrow{p} \mathbf{0}_{K+1}$.

Next, we establish that the probability limit of $\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N$ is an invertible matrix. Note that

$$\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}} = [\mathbf{1}_N\ \widetilde{\mathbf{B}}]'[\mathbf{1}_N\ \widetilde{\mathbf{B}}] = \begin{bmatrix} \mathbf{1}_N'\mathbf{1}_N & \mathbf{1}_N'\widetilde{\mathbf{B}} \\ \widetilde{\mathbf{B}}'\mathbf{1}_N & \widetilde{\mathbf{B}}'\widetilde{\mathbf{B}} \end{bmatrix}$$

and $\mathbf{1}_N'\widetilde{\mathbf{B}} = \mathbf{1}_N'\mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \mathbf{1}_N'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}}) = N\widehat{\boldsymbol{\mu}}_\beta'$, where we use the property $\mathbf{J}_N\mathbf{1}_N = \mathbf{0}_N$. Moreover,

$$\begin{aligned}
\widetilde{\mathbf{B}}'\widetilde{\mathbf{B}} &= (\mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}}))'(\mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}})) \\
&= \widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N'\mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}}) + (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}}) \\
&= N\widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}}),
\end{aligned}$$

where we again make use of the property $\mathbf{J}_N\mathbf{1}_N = \mathbf{0}_N$. It then follows from Proposition 4 and the probability limits in (18) and (121), that

$$\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N \xrightarrow{p} \boldsymbol{\Omega}, \tag{50}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \\ \boldsymbol{\mu}_\beta & \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta' + (\mathbf{I}_K - \mathbf{C})'\,\mathbf{L}_1\,(\mathbf{I}_K - \mathbf{C}) \end{bmatrix} = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \\ \boldsymbol{\mu}_\beta & \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta' + \mathbf{V}_\beta\mathbf{L}_1^{-1}\mathbf{V}_\beta \end{bmatrix}, \tag{51}$$

since $\mathbf{I}_K - \mathbf{C} = \mathbf{L}_1^{-1}\mathbf{L}_1 - v\mathbf{L}_1^{-1}\mathbf{W}_1 = \mathbf{L}_1^{-1}(\mathbf{L}_1 - v\mathbf{W}_1) = \mathbf{L}_1^{-1}\mathbf{V}_\beta$. Due to Assumption 1(iv), $\mathbf{V}_\beta$ is positive definite and so invertible. Moreover, $\mathbf{L}_1^{-1}$ is positive definite since $\mathbf{L}_1$ is positive definite. Hence, $\mathbf{V}_\beta\mathbf{L}_1^{-1}\mathbf{V}_\beta$ is positive definite as well. Since $\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'$ is positive semidefinite, it follows that $\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta' + \mathbf{V}_\beta\mathbf{L}_1^{-1}\mathbf{V}_\beta$ is positive definite and hence invertible. Theorem 7.1 in Schott (1997) then implies that $\boldsymbol{\Omega}$ is also invertible. Hence, using equation (46) along with the probability limits in (47), (49) and (50), we obtain that $\widetilde{\boldsymbol{\gamma}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\bar{\mathbf{r}}_2 \xrightarrow{p} \boldsymbol{\gamma}$ and thus the proof of the proposition is complete. ■

<u>Proof of Theorem 7</u>: According to equation (30), we have $\bar{\mathbf{r}}_2 = \widetilde{\mathbf{X}}\boldsymbol{\gamma} + \widetilde{\boldsymbol{\omega}}$. Hence, it follows from definition (32) that

$$\sqrt{N}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) = \sqrt{N}((\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\bar{\mathbf{r}}_2 - \boldsymbol{\gamma}) = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N)^{-1}(\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N}). \tag{52}$$

According to the probability limit in (50), $\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/N \xrightarrow{p} \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is defined in (51). Hence, to obtain the asymptotic distribution of $\sqrt{N}\,(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})$, it suffices to obtain the asymptotic distribution of $\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N}$, or, equivalently, the asymptotic distribution of $\mathbf{1}_N'\widetilde{\boldsymbol{\omega}}/\sqrt{N}$ and $\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N}$. Using the property $\mathbf{J}_N\mathbf{1}_N = \mathbf{0}_N$ and equation (45), $\mathbf{1}_N'\widetilde{\boldsymbol{\omega}}/\sqrt{N}$ can be expressed as follows:

$$\mathbf{1}_N'\widetilde{\boldsymbol{\omega}}/\sqrt{N} = -\boldsymbol{\lambda}'\mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N}) + \frac{1}{T-\tau}\mathbf{1}_{T-\tau}'(\mathbf{U}_2'\mathbf{1}_N/\sqrt{N}). \tag{53}$$

Following the decomposition (48), we obtain

$$\frac{1}{\sqrt{N}}\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}} = \boldsymbol{\ell}^{(1)} + \boldsymbol{\ell}^{(2)} + \boldsymbol{\ell}^{(3)} + \boldsymbol{\ell}^{(4)}$$

where

$$\boldsymbol{\ell}^{(1)} = -\widehat{\boldsymbol{\mu}}_\beta(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N})'\mathbf{G}_1\boldsymbol{\lambda}, \quad \boldsymbol{\ell}^{(2)} = \frac{1}{T-\tau}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_{T-\tau}'(\mathbf{U}_2'\mathbf{1}_N/\sqrt{N}),$$

$$\boldsymbol{\ell}^{(3)} = (\mathbf{I}_K - \widehat{\mathbf{C}})'(\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}\widehat{\mathbf{C}} - \widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda}/\sqrt{N}, \quad \boldsymbol{\ell}^{(4)} = \frac{1}{T-\tau}(\mathbf{I}_K - \widehat{\mathbf{C}})'(\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau}.$$

Using Assumption 2 and the probability limit in (18), we have

$$\boldsymbol{\ell}^{(1)} = -\boldsymbol{\mu}_\beta\boldsymbol{\lambda}'\mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N}) + O_p(1) \tag{54}$$

and

$$\boldsymbol{\ell}^{(2)} = \frac{1}{T-\tau}\boldsymbol{\mu}_\beta\mathbf{1}_{T-\tau}'(\mathbf{U}_2'\mathbf{1}_N/\sqrt{N}) + O_p(1). \tag{55}$$

Next, we obtain the asymptotic behavior of $\boldsymbol{\ell}^{(3)}$. It follows from equations (26) and (27) that

$$\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}\widehat{\mathbf{C}}/\sqrt{N} = (N\widehat{\mathbf{L}}_1)(\widehat{v}\widehat{\mathbf{L}}_1^{-1}\mathbf{W}_1)/\sqrt{N} = \sqrt{N}\widehat{v}\mathbf{W}_1. \tag{56}$$

In addition, it follows from equation (9) that

$$\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda}/\sqrt{N} = (\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} = (\mathbf{B} - \mathbf{1}_N'\boldsymbol{\mu}_\beta)'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} + \boldsymbol{\Xi}'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)\,\mathbf{1}_N'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} \tag{57}$$

and

$$(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} = ((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N})\mathbf{G}_1\boldsymbol{\lambda}. \tag{58}$$

Moreover, equations (15) and (9) imply

$$\boldsymbol{\Xi}'\boldsymbol{\Xi}/\sqrt{N} = \mathbf{G}_1'(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau))\mathbf{G}_1 + \sqrt{N}v\mathbf{W}_1. \tag{59}$$

Furthermore,

$$\mathbf{1}_N'\boldsymbol{\Xi}\boldsymbol{\lambda}/\sqrt{N} = \boldsymbol{\lambda}'\mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N}). \tag{60}$$

It then follows from equations (56), (57), (58), (59), (60), the probability limit in (18), and Assumption 2, that

$$\begin{aligned}\boldsymbol{\ell}^{(3)} \quad &= \quad -(\mathbf{I}_K - \mathbf{C})'\,((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N})\mathbf{G}_1\boldsymbol{\lambda} - (\mathbf{I}_K - \mathbf{C})'\,\mathbf{G}_1'(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau))\mathbf{G}_1\boldsymbol{\lambda} \\ &\quad +\sqrt{N}\,(\widehat{v} - v)\,(\mathbf{I}_K - \mathbf{C})'\,\mathbf{W}_1\boldsymbol{\lambda} + O_p(1).\end{aligned} \tag{61}$$

To obtain the asymptotic behavior of $\boldsymbol{\ell}^{(4)}$, we use equation (9) and note that

$$\begin{aligned}\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2/\sqrt{N} \quad &= \quad (\widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta')'\mathbf{U}_2/\sqrt{N} \\ &= \quad (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N} + \boldsymbol{\Xi}'\mathbf{U}_2/\sqrt{N} - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)(\mathbf{1}_N'\mathbf{U}_2/\sqrt{N}) \\ &= \quad (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N} + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N}) - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)(\mathbf{1}_N'\mathbf{U}_2/\sqrt{N}).\end{aligned}$$

From Assumption 2 and the facts $\widehat{\boldsymbol{\mu}}_\beta \xrightarrow{p} \boldsymbol{\mu}_\beta$ (see (18)) and $\widehat{\mathbf{C}} \xrightarrow{p} \mathbf{C}$ (see Proposition 4), it follows that

$$\boldsymbol{\ell}^{(4)} = \frac{1}{T-\tau}\,(\mathbf{I}_K - \mathbf{C})'\,((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} + \frac{1}{T-\tau}\,(\mathbf{I}_K - \mathbf{C})'\,\mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} + O_p(1). \tag{62}$$

Collecting the results of (54), (55), (61), and (62) yields

$$\begin{aligned}\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N} \quad &= \quad -\boldsymbol{\mu}_\beta\boldsymbol{\lambda}'\mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N}) + \frac{1}{T-\tau}\boldsymbol{\mu}_\beta\mathbf{1}_{T-\tau}'(\mathbf{U}_2'\mathbf{1}_N/\sqrt{N}) \\ &\quad -(\mathbf{I}_K - \mathbf{C})'\,((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N})\mathbf{G}_1\boldsymbol{\lambda} \\ &\quad -(\mathbf{I}_K - \mathbf{C})'\,\mathbf{G}_1'(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau))\mathbf{G}_1\boldsymbol{\lambda} \\ &\quad +\sqrt{N}\,(\widehat{v} - v)\,(\mathbf{I}_K - \mathbf{C})'\,\mathbf{W}_1\boldsymbol{\lambda} \\ &\quad +\frac{1}{T-\tau}(\mathbf{I}_K - \mathbf{C})'((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} \\ &\quad +\frac{1}{T-\tau}\,(\mathbf{I}_K - \mathbf{C})'\,\mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} + O_p(1).\end{aligned} \tag{63}$$

Using fact (F4), we obtain

$$(\mathbf{I}_K - \mathbf{C})'((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N})\mathbf{G}_1\boldsymbol{\lambda} = \left((\mathbf{G}_1\boldsymbol{\lambda})' \otimes (\mathbf{I}_K - \mathbf{C})'\right)\mathrm{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N}), \tag{64}$$

$$(\mathbf{I}_K - \mathbf{C})'\mathbf{G}_1'(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau))\mathbf{G}_1\boldsymbol{\lambda} = \left((\mathbf{G}_1\boldsymbol{\lambda})' \otimes (\mathbf{G}_1\,(\mathbf{I}_K - \mathbf{C}))'\right)\mathrm{vec}(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)), \tag{65}$$

$$(\mathbf{I}_K - \mathbf{C})'\,((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} = \left(\mathbf{1}_{T-\tau}' \otimes (\mathbf{I}_K - \mathbf{C})'\right)\mathrm{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N}), \tag{66}$$

and

$$(\mathbf{I}_K - \mathbf{C})'\,\mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N})\mathbf{1}_{T-\tau} = \left(\mathbf{1}_{T-\tau}' \otimes (\mathbf{G}_1\,(\mathbf{I}_K - \mathbf{C}))'\right)\mathrm{vec}(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N}). \tag{67}$$

Collecting the results in equations (63), (64), (65), (66), (67) and Lemma 16, we obtain

$$\begin{aligned}
\widetilde{\mathbf{B}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N} \quad = \quad & -\boldsymbol{\mu}_\beta\boldsymbol{\lambda}'\mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/\sqrt{N}) + \frac{1}{T-\tau}\boldsymbol{\mu}_\beta\mathbf{1}_{T-\tau}'(\mathbf{U}_2'\mathbf{1}_N/\sqrt{N}) \\
& -\left((\mathbf{G}_1\boldsymbol{\lambda})'\otimes(\mathbf{G}_1(\mathbf{I}_K-\mathbf{C}))'\right)\mathrm{vec}(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)) \\
& +\frac{(\mathbf{I}_K-\mathbf{C})'\mathbf{W}_1\boldsymbol{\lambda}\mathrm{vec}(\mathbf{H}_1)'}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{vec}(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)) \\
& +\frac{1}{T-\tau}\left(\mathbf{1}_{T-\tau}'\otimes(\mathbf{G}_1(\mathbf{I}_K-\mathbf{C}))'\right)\mathrm{vec}(\mathbf{U}_1'\mathbf{U}_2/\sqrt{N}) \\
& -\left((\mathbf{G}_1\boldsymbol{\lambda})'\otimes(\mathbf{I}_K-\mathbf{C})'\right)\mathrm{vec}((\mathbf{B}-\mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/\sqrt{N}) \\
& +\frac{1}{T-\tau}\left(\mathbf{1}_{T-\tau}'\otimes(\mathbf{I}_K-\mathbf{C})'\right)\mathrm{vec}((\mathbf{B}-\mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/\sqrt{N}) + O_p(1). \qquad (68)
\end{aligned}$$

Combining (53) and (68), we obtain that

$$\widetilde{\mathbf{X}}'\widetilde{\boldsymbol{\omega}}/\sqrt{N} = \sqrt{N}\boldsymbol{\Pi}\mathbf{e} + O_p(1), \qquad (69)$$

where $\mathbf{e}$ is defined by (33) and $\boldsymbol{\Pi}$ is the $(K+1)\times(\tau+K+1)T$ matrix defined by

$$\boldsymbol{\Pi} = \left[\begin{array}{cccccc} \boldsymbol{\Pi}^{(1)} & \boldsymbol{\Pi}^{(2)} & \boldsymbol{\Pi}^{(3)} & \boldsymbol{\Pi}^{(4)} & \boldsymbol{\Pi}^{(5)} & \boldsymbol{\Pi}^{(6)} \end{array}\right] \qquad (70)$$

where

$$\boldsymbol{\Pi}^{(1)} = -\left[\begin{array}{c} \boldsymbol{\lambda}'\mathbf{G}_1' \\ \boldsymbol{\mu}_\beta\boldsymbol{\lambda}'\mathbf{G}_1' \end{array}\right], \quad \boldsymbol{\Pi}^{(2)} = \frac{1}{T-\tau}\left[\begin{array}{c} 1 \\ \boldsymbol{\mu}_\beta \end{array}\right]\mathbf{1}_{T-\tau}',$$

$$\boldsymbol{\Pi}^{(3)} = -\left[\begin{array}{c} \mathbf{0}_{\tau^2}' \\ (\mathbf{G}_1\boldsymbol{\lambda})'\otimes(\mathbf{G}_1(\mathbf{I}_K-\mathbf{C}))' \end{array}\right] + \left[\begin{array}{c} \mathbf{0}_{\tau^2}' \\ \frac{(\mathbf{I}_K-\mathbf{C})'\mathbf{W}_1\boldsymbol{\lambda}\mathrm{vec}(\mathbf{H}_1)'}{\mathrm{tr}(\mathbf{H}_1)} \end{array}\right],$$

$$\boldsymbol{\Pi}^{(4)} = \left[\begin{array}{c} \mathbf{0}_{\tau(T-\tau)}' \\ \frac{1}{T-\tau}\left(\mathbf{1}_{T-\tau}'\otimes(\mathbf{G}_1(\mathbf{I}_K-\mathbf{C}))'\right) \end{array}\right], \quad \boldsymbol{\Pi}^{(5)} = -\left[\begin{array}{c} \mathbf{0}_{\tau K}' \\ (\mathbf{G}_1\boldsymbol{\lambda})'\otimes(\mathbf{I}_K-\mathbf{C})' \end{array}\right],$$

$$\boldsymbol{\Pi}^{(6)} = \left[\begin{array}{c} \mathbf{0}_{K(T-\tau)}' \\ \frac{1}{T-\tau}\left(\mathbf{1}_{T-\tau}'\otimes(\mathbf{I}_K-\mathbf{C})'\right) \end{array}\right].$$

Combining (69) with (52) and (50) yields the desired result and completes the proof. ■

Proof of Proposition 8: For $r,s=1,\ldots,T$, the (r,s) element of the $T\times T$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,1}\boldsymbol{\eta}_{m,1}'$ is $\frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,r}^{[1]}\zeta_{m,s}^{[1]}$. Hence, by Assumption 3(i), it follows that $\mathbf{V}_{11} = \text{p-}\lim\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,1}\boldsymbol{\eta}_{m,1}' = v_c\mathbf{I}_T$.

For each positive integer p, let $\ell_p\in\{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p=(k_p-1)\tau+\ell_p$ (see (37)). Then, for $p=1,\ldots,\tau^2$ and $r=1,\ldots,T$, the (p,r) element of the $\tau^2\times T$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,2}\boldsymbol{\eta}_{m,1}'$ is $\frac{1}{M_N}\sum_{m=1}^{M_N}(\zeta_{m,p}^{[2]} - N_m v\mathbf{1}_{[k_p=\ell_p]})\zeta_{m,r}^{[1]}$. Hence, by Assumptions 4(i) and 5(ii), it follows that $\mathbf{V}_{21} = \text{p-}\lim\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,2}\boldsymbol{\eta}_{m,1}' = \xi_c\boldsymbol{\mathcal{M}}$, where the matrix $\boldsymbol{\mathcal{M}}$ is defined in the statement of Proposition 8.

For $p,q=1,\ldots,\tau^2$, the (p,q) element of the $\tau^2\times\tau^2$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,2}\boldsymbol{\eta}_{m,2}'$ is

$$\begin{aligned}
\lambda_{p,q} = & \frac{1}{M_N}\sum_{m=1}^{M_N}(\zeta_{m,p}^{[2]} - N_m v\mathbf{1}_{[k_p=\ell_p]})(\zeta_{m,q}^{[2]} - N_m v\mathbf{1}_{[k_q=\ell_q]}) \\
= & \frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,p}^{[2]}\zeta_{m,q}^{[2]} + v^2\mathbf{1}_{[k_p=\ell_p]}\mathbf{1}_{[k_q=\ell_q]}\frac{1}{M_N}\sum_{m=1}^{M_N}N_m^2 \\
& - v\mathbf{1}_{[k_p=\ell_p]}\frac{1}{M_N}\sum_{m=1}^{M_N}N_m\zeta_{m,q}^{[2]} - v\mathbf{1}_{[k_q=\ell_q]}\frac{1}{M_N}\sum_{m=1}^{M_N}N_m\zeta_{m,p}^{[2]}.
\end{aligned}$$

It follows from Assumptions 3(ii), 5(i), and 5(iii) that the probability limit of $\lambda_{p,q}$, as $N\to\infty$, is

$$\begin{aligned}
& \kappa_c^{[1]}\mathbf{1}_{[k_p=\ell_p=k_q=\ell_q]} + \kappa_c^{[2]}\left(\mathbf{1}_{[k_p=k_q\neq\ell_p=\ell_q]} + \mathbf{1}_{[k_p=\ell_q\neq\ell_p=k_q]}\right) + \kappa_c^{[3]}\mathbf{1}_{[k_p=\ell_p\neq k_q=\ell_q]} \\
& + Hv^2\mathbf{1}_{[k_p=\ell_p]}\mathbf{1}_{[k_q=\ell_q]} - 2G^2v^2\mathbf{1}_{[k_p=\ell_p]}\mathbf{1}_{[k_q=\ell_q]}.
\end{aligned}$$

Hence, $\mathbf{V}_{22} = \text{p-}\lim\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,2}\boldsymbol{\eta}_{m,2}' = \kappa_c^{[1]}\boldsymbol{\mathcal{K}}^{[1]} + \kappa_c^{[2]}\boldsymbol{\mathcal{K}}^{[2]} + \kappa_c^{[3]}\boldsymbol{\mathcal{K}}^{[3]} - (2G^2-H)v^2\mathbf{i}_\tau\mathbf{i}_\tau'$, where the matrices $\boldsymbol{\mathcal{K}}^{[i]}$, $i=1,2,3$, are defined in the statement of Proposition 8.

For $p=1,\ldots,(T-\tau)\tau$, $s=1,\ldots,T$, the (p,s) element of the $(T-\tau)\tau\times T$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N}\boldsymbol{\eta}_{m,3}\boldsymbol{\eta}_{m,1}'$ is $\frac{1}{M_N}\sum_{m=1}^{M_N}\zeta_{m,p}^{[3]}\zeta_{m,s}^{[1]}$.

Hence, by Assumption 4(ii), it follows that $\mathbf{V}_{31} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,3}\boldsymbol{\eta}'_{m,1} = \mathbf{0}_{(T-\tau)\tau\times T}$.

For $p = 1,\ldots,(T-\tau)\tau$ and $q = 1,\ldots,\tau^2$, the (p,q) element of the $(T-\tau)\tau \times \tau^2$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,3}\boldsymbol{\eta}'_{m,2}$ is $\frac{1}{M_N}\sum_{m=1}^{M_N} \zeta^{[3]}_{m,p}(\zeta^{[2]}_{m,q} - N_m v \mathbf{1}_{[k_q=\ell_q]})$. So, Assumptions 4(iv) and 5(iv) imply $\mathbf{V}_{32} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,3}\boldsymbol{\eta}'_{m,2} = \mathbf{0}_{(T-\tau)\tau\times\tau^2}$.

For $p,q = 1,\ldots,(T-\tau)\tau$, the (p,q) element of the $(T-\tau)\tau \times (T-\tau)\tau$ matrix $\frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,3}\boldsymbol{\eta}'_{m,3}$ is given by $\frac{1}{M_N}\sum_{m=1}^{M_N} \zeta^{[3]}_{m,p}\zeta^{[3]}_{m,q}$. Assumption 3(iii) then implies $\mathbf{V}_{33} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,3}\boldsymbol{\eta}'_{m,3} = \kappa^{[2]}_c \mathbf{I}_{(T-\tau)\tau}$.

Moreover, Assumption 4(iii) implies that $\mathbf{V}_{41} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,4}\boldsymbol{\eta}'_{m,1} = \mathbf{I}_T \otimes \mathbf{c}_{41}$, Assumptions 4(v) and 5(v) imply that $\mathbf{V}_{42} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,4}\boldsymbol{\eta}'_{m,2} = \boldsymbol{\mathcal{M}}' \otimes \mathbf{c}_{42}$, and from Assumption 4(vi) we obtain $\mathbf{V}_{43} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,4}\boldsymbol{\eta}'_{m,3} = \mathbf{0}_{TK\times(T-\tau)\tau}$. Finally, Assumption 3(iv) implies that $\mathbf{V}_{44} = \text{p-}\lim \frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\eta}_{m,4}\boldsymbol{\eta}'_{m,4} = \mathbf{I}_T \otimes \mathbf{C}_{44}$. ■

A.3 N-consistent estimation of v_c, ξ_c, $\kappa^{[1]}_c$, $\kappa^{[2]}_c$, and $\kappa^{[3]}_c$

In this subsection, we provide N-consistent estimators of the second-moment parameter v_c, the third-moment parameter ξ_c, and the fourth-moment parameters $\kappa^{[1]}_c$, $\kappa^{[2]}_c$, $\kappa^{[3]}_c$. To do so, we need to introduce some additional notation. Define the cluster selection $M_N \times N$ matrix $\boldsymbol{\mathcal{C}}$ with (m,i) element given by

$$\boldsymbol{\mathcal{C}}(m,i) = \mathbf{1}_{[i\in I_m]}, \quad m = 1,\ldots,M_N, \quad i = 1,\ldots,N. \tag{71}$$

It then follows that the (m,r) element of the $M_N \times \tau$ matrix $\boldsymbol{\mathcal{C}}\mathbf{U}_1$ is $\zeta^{[1]}_{m,r}$, for $m = 1,\ldots,M_N$, $r = 1,\ldots,\tau$. Hence, according to Assumption 3(i), we have that, as $N\to\infty$, $(\boldsymbol{\mathcal{C}}\mathbf{U}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N \xrightarrow{p} v_c\mathbf{I}_\tau$. Employing $(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N$, the natural feasible proxy of $(\boldsymbol{\mathcal{C}}\mathbf{U}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N$, and accounting for estimation error, along the lines of Lemma 3, one can establish the following

Lemma 10 *Under Assumption 3(i),*

$$\widehat{v}_c = \frac{tr((\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N)}{tr(\mathbf{H}_1)}$$

is an N-consistent estimator of v_c, where the matrix of residuals $\widehat{\mathbf{U}}_1$ is defined in (22), the matrix $\mathbf{H}_1$ is defined in (24), and the matrix $\boldsymbol{\mathcal{C}}$ is defined in (71).

Noting that, for $m = 1,\ldots,M_N$ and $r = 1,\ldots,\tau$, the (m,r) element of the $M_N\times\tau$ matrix $\boldsymbol{\mathcal{C}}(\mathbf{U}_1\odot\mathbf{U}_1)$ is $\zeta^{[2]}_{m,(r-1)\tau+r}$ and recalling that the (m,r) element of the $M_N\times\tau$ matrix $\boldsymbol{\mathcal{C}}\mathbf{U}_1$ is $\zeta^{[1]}_{m,r}$, it follows from Assumption 4(i) that, as $N\to\infty$,

$$(\boldsymbol{\mathcal{C}}(\mathbf{U}_1\odot\mathbf{U}_1))'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N \xrightarrow{p} \xi_c\mathbf{I}_\tau. \tag{72}$$

In the next lemma, motivated by the probability limit in (72), we utilize $(\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1\odot\widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N$, the natural feasible proxy of $(\boldsymbol{\mathcal{C}}(\mathbf{U}_1\odot\mathbf{U}_1))'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N$, and account for estimation error to provide an N-consistent estimator of the third-moment parameter ξ_c.

Lemma 11 *Under Assumption 4(i),*

$$\widehat{\xi}_c = \frac{tr((\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1\odot\widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N)}{tr((\mathbf{H}_1\odot\mathbf{H}_1)\,\mathbf{H}_1)}$$

is an N-consistent estimator of the third-moment parameter ξ_c, where $\odot$ denotes the Hadamard product, $\widehat{\mathbf{U}}_1$ is the matrix of residuals defined in (22), the matrix $\mathbf{H}_1$ is defined in (24), and the matrix $\boldsymbol{\mathcal{C}}$ is defined in (71).

<u>Proof of Lemma 11</u>: It follows from (22) that

$$\widehat{\mathbf{U}}_1\odot\widehat{\mathbf{U}}_1 = \begin{bmatrix}\widehat{\mathbf{u}}'_{1,[1]}\\ \vdots \\ \widehat{\mathbf{u}}'_{1,[N]}\end{bmatrix} \odot \begin{bmatrix}\widehat{\mathbf{u}}'_{1,[1]}\\ \vdots \\ \widehat{\mathbf{u}}'_{1,[N]}\end{bmatrix} = \begin{bmatrix}\widehat{\mathbf{u}}'_{1,[1]}\odot\widehat{\mathbf{u}}'_{1,[1]}\\ \vdots \\ \widehat{\mathbf{u}}'_{1,[N]}\odot\widehat{\mathbf{u}}'_{1,[N]}\end{bmatrix}.$$

Consider the $N\times\tau^2$ matrices $\boldsymbol{\mathcal{U}}_1$ and $\widehat{\boldsymbol{\mathcal{U}}}_1$ defined in (75). For each positive integer p, let $\ell_p \in \{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p = (k_p-1)\tau+\ell_p$ (see (37)). Then, the (i,p) element of $\widehat{\boldsymbol{\mathcal{U}}}_1$ is $\widehat{\eta}_{i,p} = \widehat{u}_{i,k_p}\widehat{u}_{i,\ell_p}$, for $i = 1,\ldots,N$ and $p = 1,\ldots,\tau^2$. Hence, the (i,r) element of the $N\times\tau$ matrix $\widehat{\boldsymbol{\mathcal{U}}}_1\boldsymbol{\mathcal{N}}$, where $\boldsymbol{\mathcal{N}}$ is defined in Lemma 15, is $\sum_{p=1}^{\tau^2}\widehat{\eta}_{i,p}\mathbf{1}_{[k_p=\ell_p=r]} = \widehat{u}_{i,r}\widehat{u}_{i,r}$, which is equal to the (i,r) element of the matrix $\widehat{\mathbf{U}}_1\odot\widehat{\mathbf{U}}_1$. It follows that $\widehat{\mathbf{U}}_1\odot\widehat{\mathbf{U}}_1 = \widehat{\boldsymbol{\mathcal{U}}}_1\boldsymbol{\mathcal{N}}$, and so

$$
\begin{aligned}
(\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1 \odot \widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N &= \boldsymbol{\mathcal{N}}'((\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N) \\
&= \boldsymbol{\mathcal{N}}'\left(\sum_{m=1}^{M_n}\left(\sum_{i\in I_m}\widehat{\mathbf{u}}_{1,[i]}\otimes\widehat{\mathbf{u}}_{1,[i]}\right)\left(\sum_{i\in I_m}\widehat{\mathbf{u}}'_{1,[i]}\right)/M_N\right).
\end{aligned}
$$

Using equation (125) and the symmetry of $\mathbf{H}_1$, we obtain

$$
\widehat{\mathbf{u}}_{1,[i]} = \mathbf{H}_1\mathbf{u}_{1,[i]}. \tag{73}
$$

Hence, using fact (F5), we obtain

$$
\left(\sum_{i\in I_m}\widehat{\mathbf{u}}_{1,[i]}\otimes\widehat{\mathbf{u}}_{1,[i]}\right)\left(\sum_{i\in I_m}\widehat{\mathbf{u}}'_{1,[i]}\right) = (\mathbf{H}_1\otimes\mathbf{H}_1)\left(\sum_{i\in I_m}\mathbf{u}_{1,[i]}\otimes\mathbf{u}_{1,[i]}\right)\left(\sum_{i\in I_m}\mathbf{u}'_{1,[i]}\right)\mathbf{H}_1.
$$

It follows that

$$
(\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1 \odot \widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N = \boldsymbol{\mathcal{N}}'(\mathbf{H}_1\otimes\mathbf{H}_1)\left((\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N\right)\mathbf{H}_1.
$$

Recall that the (m,s) element of the matrix $\boldsymbol{\mathcal{C}}\mathbf{U}_1$ is $\zeta^{[1]}_{m,s}$, for $m=1,\dots,M_N$ and $s=1,\dots,\tau$, where $\zeta^{[1]}_{m,s}$ is defined in (38). Further, note that the (m,p) element of the $M_N\times\tau^2$ matrix $\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1 = \boldsymbol{\mathcal{C}}(\mathbf{U}_1\odot\mathbf{U}_1)$ is $\zeta^{[2]}_{m,p}$, for $m=1,\dots,M_N$ and $p=1,\dots,\tau^2$, where $\zeta^{[2]}_{m,p}$ is defined in (39). Hence, it follows that the (p,s) element of the $\tau^2\times\tau$ matrix $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N$ is

$$
\vartheta_{p,s} = \sum_{m=1}^{M_N}\zeta^{[2]}_{m,p}\zeta^{[1]}_{m,s}/M_N, \quad p=1,\dots,\tau^2, \quad s=1,\dots,\tau.
$$

It follows from Assumption 4(i), that $\vartheta_{p,s}\xrightarrow{p}\xi_c\mathbf{1}_{[k_p=\ell_p=s]}$ and so, as $N\to\infty$,

$$
(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N \xrightarrow{p} \xi_c\boldsymbol{\mathcal{N}}. \tag{74}
$$

Thus, $(\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1 \odot \widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N \xrightarrow{p} \xi_c(\boldsymbol{\mathcal{N}}'(\mathbf{H}_1\otimes\mathbf{H}_1)\boldsymbol{\mathcal{N}})\mathbf{H}_1$. Lemma 15 yields that $\boldsymbol{\mathcal{N}}'(\mathbf{H}_1\otimes\mathbf{H}_1)\boldsymbol{\mathcal{N}} = \mathbf{H}_1\odot\mathbf{H}_1 = (\mathbf{H}_1\odot\mathbf{H}_1)$, and so

$$
\widehat{\xi}_c = \frac{\mathrm{tr}((\boldsymbol{\mathcal{C}}(\widehat{\mathbf{U}}_1 \odot \widehat{\mathbf{U}}_1))'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N)}{\mathrm{tr}\,((\mathbf{H}_1\odot\mathbf{H}_1)\mathbf{H}_1)} \xrightarrow{p} \xi_c,
$$

completing the proof of the lemma. ■

To proceed with the estimation of the fourth-moment parameters $\kappa_c^{[1]}$, $\kappa_c^{[2]}$, and $\kappa_c^{[3]}$, we need to define the following $N\times\tau^2$ matrices:

$$
\boldsymbol{\mathcal{U}}_1 = \begin{bmatrix} \mathbf{u}'_{1,[1]}\otimes\mathbf{u}'_{1,[1]} \\ \vdots \\ \mathbf{u}'_{1,[N]}\otimes\mathbf{u}'_{1,[N]} \end{bmatrix} \text{ and } \widehat{\boldsymbol{\mathcal{U}}}_1 = \begin{bmatrix} \widehat{\mathbf{u}}'_{1,[1]}\otimes\widehat{\mathbf{u}}'_{1,[1]} \\ \vdots \\ \widehat{\mathbf{u}}'_{1,[N]}\otimes\widehat{\mathbf{u}}'_{1,[N]} \end{bmatrix}. \tag{75}
$$

For $m=1,\dots,M_N$, $p=1,\dots,\tau^2$, the (m,p) element of the $M_N\times\tau^2$ matrix $\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1$ is $\zeta^{[2]}_{m,p}$, and so Assumption 3(ii) implies that

$$
(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N \xrightarrow{p} \kappa_c^{[1]}\boldsymbol{\mathcal{K}}^{[1]} + \kappa_c^{[2]}\boldsymbol{\mathcal{K}}^{[2]} + \kappa_c^{[3]}\boldsymbol{\mathcal{K}}^{[3]}, \tag{76}
$$

where the matrices $\boldsymbol{\mathcal{K}}^{[i]}$, $i=1,2,3$ are defined in Proposition 8. In the following lemma, motivated by the probability limit in (76), we employ $(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)/M_N$, the natural feasible proxy of $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N$, and account for estimation error to provide N-consistent estimators of the fourth-moment parameters $\kappa_c^{[1]}$, $\kappa_c^{[2]}$, and $\kappa_c^{[3]}$.

Lemma 12 *Define*

$$
M_1(\mathbf{H}_1) = tr(\mathbf{H}_1),\ M_2(\mathbf{H}_1) = tr(\mathbf{H}_1\odot\mathbf{H}_1),\ M_4(\mathbf{H}_1) = tr((\mathbf{H}_1\odot\mathbf{H}_1)(\mathbf{H}_1\odot\mathbf{H}_1)),
$$

where the matrix $\mathbf{H}_1$ *is defined in (24). Moreover, define the* 3×3 *matrix*

$$
\boldsymbol{\mathcal{M}}(\mathbf{H}_1) = \begin{bmatrix} M_2(\mathbf{H}_1) & M_1(\mathbf{H}_1)^2 + M_1(\mathbf{H}_1) - 2M_2(\mathbf{H}_1) & M_1(\mathbf{H}_1) - M_2(\mathbf{H}_1) \\ M_2(\mathbf{H}_1) & 2(M_1(\mathbf{H}_1) - M_2(\mathbf{H}_1)) & M_1(\mathbf{H}_1)^2 - M_2(\mathbf{H}_1) \\ M_4(\mathbf{H}_1) & 2(M_2(\mathbf{H}_1) - M_4(\mathbf{H}_1)) & M_2(\mathbf{H}_1) - M_4(\mathbf{H}_1) \end{bmatrix}. \tag{77}
$$

The determinant of $\boldsymbol{\mathcal{M}}(\mathbf{H}_1)$ is given by

$$det(\boldsymbol{\mathcal{M}}(\mathbf{H}_1)) = (M_1(\mathbf{H}_1) - 1) M_1(\mathbf{H}_1) \left(M_1(\mathbf{H}_1)(M_1(\mathbf{H}_1)+2) M_4(\mathbf{H}_1) - 3M_2(\mathbf{H}_1)^2\right).$$

For $\tau \geq \tau^*(K)$, where $\tau^*(K) = \left[\frac{1}{2} + \frac{1}{2}\sqrt{1+12(K+1)}\right] + K + 2$, we have $det(\boldsymbol{\mathcal{M}}(\mathbf{H}_1)) > 0$. Then, under Assumption 3(ii) and for $\tau \geq \tau^*(K)$,

$$\widehat{\boldsymbol{\kappa}}_c = \boldsymbol{\mathcal{M}}(\mathbf{H}_1)^{-1} \boldsymbol{y}_1$$

is an N-consistent estimator of

$$\boldsymbol{\kappa}_c = \left[\begin{array}{ccc} \kappa_c^{[1]} & \kappa_c^{[2]} & \kappa_c^{[3]} \end{array}\right]', \tag{78}$$

where

$$\boldsymbol{y}_1 = \frac{1}{M_N}\left[\begin{array}{ccc} tr((\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)) & \mathbf{i}_\tau'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)\mathbf{i}_\tau & tr(\boldsymbol{\mathcal{N}}'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)\boldsymbol{\mathcal{N}}) \end{array}\right]', \tag{79}$$

and the matrices $\boldsymbol{\mathcal{C}}$ and $\widehat{\boldsymbol{\mathcal{U}}}_1$ are defined in (71) and (75), respectively. Finally, the matrix $\boldsymbol{\mathcal{N}}$ is defined as follows. For each positive integer p, let $\ell_p \in \{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). Then, $\boldsymbol{\mathcal{N}}$ is the $\tau^2 \times \tau$ matrix defined by $\boldsymbol{\mathcal{N}}(p,r) = \mathbf{1}_{[k_p=\ell_p=r]}$, for $p = 1,\ldots,\tau^2$ and $r = 1,\ldots,\tau$.

It is worth emphasizing that the condition $\tau \geq \tau^*(K)$ is not very restrictive. Specifically, we have $\tau^*(1) = 6$, $\tau^*(2) = 7$, $\tau^*(3) = 9$, $\tau^*(4) = 10$. That is, for a single-factor model the estimation period time-series sample size should be at least 6, while for a three-factor model the estimation period time-series sample size should be at least 9.

Proof of Lemma 12: Using equation (73) and fact (F5), we obtain

$$\widehat{\boldsymbol{\mathcal{U}}}_1 = \boldsymbol{\mathcal{U}}_1(\mathbf{H}_1 \otimes \mathbf{H}_1) \tag{80}$$

which, in light of the symmetry of $\mathbf{H}_1$, implies

$$(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)/M_N = (\mathbf{H}_1 \otimes \mathbf{H}_1)((\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N)(\mathbf{H}_1 \otimes \mathbf{H}_1).$$

The (p,q) element of the $\tau^2 \times \tau^2$ matrix $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N$ is $\sum_{m=1}^{M_N} \zeta_{m,p}^{[2]}\zeta_{m,q}^{[2]}/M_N$. Hence, it follows from Assumption 3(ii) that

$$(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N \xrightarrow{p} \kappa_c^{[1]}\boldsymbol{\mathcal{K}}^{[1]} + \kappa_c^{[2]}\boldsymbol{\mathcal{K}}^{[2]} + \kappa_c^{[3]}\boldsymbol{\mathcal{K}}^{[3]}, \tag{81}$$

and so

$$(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)/M_N \xrightarrow{p} \kappa_c^{[1]}\boldsymbol{\mathcal{D}}^{[1]} + \kappa_c^{[2]}\boldsymbol{\mathcal{D}}^{[2]} + \kappa_c^{[3]}\boldsymbol{\mathcal{D}}^{[3]}, \tag{82}$$

where the matrices $\boldsymbol{\mathcal{D}}^{[i]}$, $i = 1,2,3$ are defined in (127). In light of Lemmas 19, 20, and 21, equation (82) yields

$$\boldsymbol{y}_1 \xrightarrow{p} \boldsymbol{\mathcal{M}}(\mathbf{H}_1)\boldsymbol{\kappa}_c, \tag{83}$$

where $\boldsymbol{\kappa}_c$ is the vector of fourth-moment parameters defined in (78), $\boldsymbol{\mathcal{M}}(\mathbf{H}_1)$ is the 3×3 matrix defined in (77) and $\boldsymbol{y}_1$ is the 3×1 vector defined in (79). Simple algebra shows that the determinant of the 3×3 matrix $\boldsymbol{\mathcal{M}}(\mathbf{H}_1)$ is

$$\det(\boldsymbol{\mathcal{M}}(\mathbf{H}_1)) = (M_1(\mathbf{H}_1) - 1) M_1(\mathbf{H}_1)\left[M_1(\mathbf{H}_1)(M_1(\mathbf{H}_1)+2) M_4(\mathbf{H}_1) - 3M_2(\mathbf{H}_1)^2\right].$$

Next, we show that $M_4(\mathbf{H}_1) \geq \frac{1}{\tau}M_2(\mathbf{H}_1)^2$. Denote by h_{rs} the (r,s) element of the matrix $\mathbf{H}_1$. Then, $M_2(\mathbf{H}_1) = \mathrm{tr}(\mathbf{H}_1 \odot \mathbf{H}_1) = \sum_{r=1}^{\tau} h_{rr}^2$. Moreover, the (r,r) element of the matrix $(\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1)$ is $\sum_{s=1}^{\tau} h_{rs}^2 h_{sr}^2$. Hence, $M_4(\mathbf{H}_1) = \mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1)) = \sum_{r=1}^{\tau}\sum_{s=1}^{\tau} h_{rs}^2 h_{sr}^2 = \sum_{r=1}^{\tau}\sum_{s=1}^{\tau} h_{rs}^4$ due to the symmetry of $\mathbf{H}_1$. It follows that $M_4(\mathbf{H}_1) \geq \sum_{r=1}^{\tau} h_{rr}^4$. The Cauchy-Schwarz inequality implies $M_2(\mathbf{H}_1) = \sum_{r=1}^{\tau} h_{rr}^2 \leq \sqrt{\left(\sum_{r=1}^{\tau} h_{rr}^4\right)\cdot\left(\sum_{r=1}^{\tau} 1\right)} = \sqrt{\tau\left(\sum_{r=1}^{\tau} h_{rr}^4\right)} \leq \sqrt{\tau M_4(\mathbf{H}_1)}$ or equivalently $M_4(\mathbf{H}_1) \geq \frac{1}{\tau}M_2(\mathbf{H}_1)^2$. Hence,

$$M_1(\mathbf{H}_1)(M_1(\mathbf{H}_1)+2) M_4(\mathbf{H}_1) - 3M_2(\mathbf{H}_1)^2 \geq \frac{1}{\tau}M_2(\mathbf{H}_1)^2\left[M_1(\mathbf{H}_1)(M_1(\mathbf{H}_1)+2) - 3\tau\right],$$

and so $\det(\boldsymbol{\mathcal{M}}(\mathbf{H}_1)) > 0$ if $M_1(\mathbf{H}_1) > 1$ and $M_1(\mathbf{H}_1)(M_1(\mathbf{H}_1)+2) - 3\tau > 0$. Since $M_1(\mathbf{H}_1) = \tau - K - 1$, the last inequality is equivalent to $M_1(\mathbf{H}_1)^2 - M_1(\mathbf{H}_1) - 3(K+1) > 0$. Since, for $c > 0$, $x^2 - x - c > 0$ if $x > \frac{1}{2} + \frac{1}{2}\sqrt{1+4c}$, it follows that the above inequality is satisfied when $M_1(\mathbf{H}_1) \geq M_1^*(K) = \left[\frac{1}{2} + \frac{1}{2}\sqrt{1+12(K+1)}\right] + 1$, where $[a]$ denotes the integer part of a. Note that $M_1^*(K)$ is (weakly) increasing in K and $M_1^*(1) = 4$, $M_1^*(2) = 4$, $M_1^*(3) = 5$, and $M_1^*(4) = 5$. It follows that $\det(\boldsymbol{\mathcal{M}}(\mathbf{H}_1)) > 0$ for $\tau \geq \tau^*(K) = M_1^*(K) + K + 1$. It follows from (83), that for $\tau \geq \tau^*(K)$, $\widehat{\boldsymbol{\kappa}}_c = \boldsymbol{\mathcal{M}}(\mathbf{H}_1)^{-1}\boldsymbol{y}_1$ is an N-consistent estimator of $\boldsymbol{\kappa}_c$. ■

A.4 N-consistent estimation of $\mathbf{c}_{41}$, $\mathbf{c}_{42}$, and $\mathbf{C}_{44}$

In this subsection, we provide N-consistent estimators of the $K \times 1$ vectors $\mathbf{c}_{41}$ and $\mathbf{c}_{42}$ and the $K \times K$ matrix $\mathbf{C}_{44}$ (see Assumptions 4(iii), 4(v), and 3(iv)). Define the $N \times (\tau K)$ matrix $\boldsymbol{\mathcal{Z}}_1$ by

$$\boldsymbol{\mathcal{Z}}_1 = \begin{bmatrix} \mathbf{u}'_{1,[1]} \otimes (\boldsymbol{\beta}_1 - \boldsymbol{\mu}_\beta)' \\ \vdots \\ \mathbf{u}'_{1,[N]} \otimes (\boldsymbol{\beta}_N - \boldsymbol{\mu}_\beta)' \end{bmatrix} \tag{84}$$

and note that

$$\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1 = \begin{bmatrix} \boldsymbol{\zeta}^{[4]}_{1,1} & \boldsymbol{\zeta}^{[4]}_{1,2} & \cdots & \boldsymbol{\zeta}^{[4]}_{1,\tau} \\ \boldsymbol{\zeta}^{[4]}_{2,1} & \boldsymbol{\zeta}^{[4]}_{2,2} & \cdots & \boldsymbol{\zeta}^{[4]}_{2,\tau} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\zeta}^{[4]}_{M_N,1} & \boldsymbol{\zeta}^{[4]}_{M_N,2} & \cdots & \boldsymbol{\zeta}^{[4]}_{M_N,\tau} \end{bmatrix}.$$

According to Assumption 4(iii), as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\zeta}^{[4]}_{m,r}\zeta^{[1]}_{m,s} \xrightarrow{p} \mathbf{c}_{41}\mathbf{1}_{[r=s]}$, for $r, s = 1, \ldots, \tau$, and, therefore,

$$(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N \xrightarrow{p} \mathbf{I}_\tau \otimes \mathbf{c}_{41}. \tag{85}$$

According to Assumption 4(v), as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\zeta}^{[4]}_{m,r}\zeta^{[2]}_{m,p} \xrightarrow{p} \mathbf{c}_{42}\mathbf{1}_{[r=k_p=\ell_p]}$, for $r = 1, \ldots, \tau$, $p = 1 \ldots \tau^2$, and, therefore,

$$(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N \xrightarrow{p} \boldsymbol{\mathcal{N}}' \otimes \mathbf{c}_{42}, \tag{86}$$

where $\boldsymbol{\mathcal{U}}_1$ is defined in (75) and the $\tau^2 \times \tau$ matrix $\boldsymbol{\mathcal{N}}$ is defined by

$$\boldsymbol{\mathcal{N}}(p, r) = \mathbf{1}_{[k_p=\ell_p=r]}, \quad p = 1, \ldots, \tau^2, \quad r = 1, \ldots, \tau. \tag{87}$$

Finally, according to Assumption 3(iv), as $N \to \infty$, $\frac{1}{M_N}\sum_{m=1}^{M_N} \boldsymbol{\zeta}^{[4]}_{m,r}(\boldsymbol{\zeta}^{[4]}_{m,s})' \xrightarrow{p} \mathbf{C}_{44}\mathbf{1}_{[r=s]}$, for $r, s = 1, \ldots, \tau$, and, therefore,

$$(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)/M_N \xrightarrow{p} \mathbf{I}_\tau \otimes \mathbf{C}_{44}. \tag{88}$$

The next lemma provides N-consistent estimators of the vectors $\mathbf{c}_{41}$, $\mathbf{c}_{42}$, and the matrix $\mathbf{C}_{44}$ based on the matrices $(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N$, $(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)/M_N$, and $(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)/M_N$, the natural feasible proxies of the matrices $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)/M_N$, $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)/M_N$, and $(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)/M_N$, respectively, where

$$\widehat{\boldsymbol{\mathcal{Z}}}_1 = \begin{bmatrix} \widehat{\mathbf{u}}'_{1,[1]} \otimes (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\mu}}_\beta)' \\ \vdots \\ \widehat{\mathbf{u}}'_{1,[N]} \otimes (\widehat{\boldsymbol{\beta}}_N - \widehat{\boldsymbol{\mu}}_\beta)' \end{bmatrix}. \tag{89}$$

Lemma 13 *Let $\boldsymbol{\Lambda}$ be the $\tau \times \tau$ matrix with (r, s) element equal to $\mathbf{1}_{[r \geq s]}$, for all $r, s = 1, \ldots, \tau$, $\boldsymbol{\mathcal{J}}$ be the $(\tau K) \times (\tau K)$ matrix defined given by $\boldsymbol{\mathcal{J}} = \boldsymbol{\Lambda} \otimes \mathbf{1}_{K \times K}$, $\boldsymbol{\mathcal{I}}$ be the $(\tau K) \times K$ matrix defined by $\boldsymbol{\mathcal{I}} = \mathbf{1}_\tau \otimes \mathbf{I}_K$, $\boldsymbol{\mathcal{L}}$ be the $(\tau K) \times \tau$ matrix defined by $\boldsymbol{\mathcal{L}} = \mathbf{I}_\tau \otimes \mathbf{1}_K$. Define*

$$\boldsymbol{\mathcal{Q}}_{41} = \frac{1}{M_N}(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1) - \widehat{\xi}_c\,(\mathbf{H}_1 \otimes \mathbf{G}_1)'\boldsymbol{\mathcal{N}}\mathbf{H}_1 \tag{90}$$

and

$$\boldsymbol{\mathcal{Q}}_{42} = \frac{1}{M_N}(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1)\boldsymbol{\mathcal{N}} - (\mathbf{H}_1 \otimes \mathbf{G}_1)'\left(\widehat{\kappa}^{[1]}_c\boldsymbol{\mathcal{K}}^{[1]} + \widehat{\kappa}^{[2]}_c\boldsymbol{\mathcal{K}}^{[2]} + \widehat{\kappa}^{[3]}_c\boldsymbol{\mathcal{K}}^{[3]}\right)(\mathbf{H}_1 \otimes \mathbf{H}_1)\boldsymbol{\mathcal{N}}, \tag{91}$$

where the matrices $\mathbf{G}_1$, $\mathbf{H}_1$, $\boldsymbol{\mathcal{C}}$, $\boldsymbol{\mathcal{N}}$, and $\widehat{\boldsymbol{\mathcal{Z}}}_1$ are defined in (10), (24), (71), (87), and (89), respectively, the matrices $\boldsymbol{\mathcal{K}}^{[i]}$, $i = 1, 2, 3$, are defined in Proposition 8, the estimator $\widehat{\xi}_c$ is given by Lemma 11, and the estimators $\widehat{\kappa}^{[i]}_c$, $i = 1, 2, 3$, are given by Lemma 12. Then,

$$\widehat{\mathbf{c}}_{41} = \frac{\boldsymbol{\mathcal{I}}'\,(\boldsymbol{\mathcal{L}} \odot \boldsymbol{\mathcal{Q}}_{41})\,\mathbf{1}_\tau}{tr\,(\mathbf{H}_1)} \quad \textit{and} \quad \widehat{\mathbf{c}}_{42} = \frac{\boldsymbol{\mathcal{I}}'\,(\boldsymbol{\mathcal{L}} \odot \boldsymbol{\mathcal{Q}}_{42})\,\mathbf{1}_\tau}{tr\,((\mathbf{H}_1 \odot \mathbf{H}_1)\,\mathbf{H}_1)}$$

are N-consistent estimators of $\mathbf{c}_{41}$ and $\mathbf{c}_{42}$, respectively. Finally, define

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{44} =& \frac{1}{M_N}(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1) - (\mathbf{H}_1 \otimes \mathbf{G}_1)'(\widehat{\kappa}_c^{[1]}\boldsymbol{\mathcal{K}}^{[1]} + \widehat{\kappa}_c^{[2]}\boldsymbol{\mathcal{K}}^{[2]} + \widehat{\kappa}_c^{[3]}\boldsymbol{\mathcal{K}}^{[3]})(\mathbf{H}_1 \otimes \mathbf{G}_1) \\ & - (\mathbf{H}_1 \otimes \mathbf{I}_K)\left(\boldsymbol{\mathcal{N}}' \otimes \widehat{\mathbf{c}}_{42}\right)(\mathbf{H}_1 \otimes \mathbf{G}_1) - \left[(\mathbf{H}_1 \otimes \mathbf{I}_K)\left(\boldsymbol{\mathcal{N}}' \otimes \widehat{\mathbf{c}}_{42}\right)(\mathbf{H}_1 \otimes \mathbf{G}_1)\right]'. \end{aligned} \tag{92}$$

Then,

$$\widehat{\mathbf{C}}_{44} = \frac{\boldsymbol{\mathcal{I}}'\left(\boldsymbol{\mathcal{J}}' \odot \boldsymbol{\mathcal{Q}}_{44} \odot \boldsymbol{\mathcal{J}}\right)\boldsymbol{\mathcal{I}}}{tr(\mathbf{H}_1)}$$

is an N-consistent estimator of $\mathbf{C}_{44}$.

Proof of Lemma 13: Equation (9) implies $\widehat{\mathbf{B}} = \mathbf{B} + \mathbf{U}_1\mathbf{G}_1$, and so $\widehat{\boldsymbol{\beta}}_i = \boldsymbol{\beta}_i + \mathbf{G}_1'\mathbf{u}_{1,[i]}$, for $i = 1, \ldots, N$. Therefore, using equation (73), we obtain

$$\begin{aligned}\widehat{\mathbf{u}}_{1,[i]} \otimes (\widehat{\boldsymbol{\beta}}_i - \widehat{\boldsymbol{\mu}}_\beta) =& (\mathbf{H}_1\mathbf{u}_{1,[i]}) \otimes (\boldsymbol{\beta}_i + \mathbf{G}_1'\mathbf{u}_{1,[i]} - \widehat{\boldsymbol{\mu}}_\beta) \\ =& (\mathbf{H}_1\mathbf{u}_{1,[i]}) \otimes (\mathbf{I}_K(\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta) + \mathbf{G}_1'\mathbf{u}_{1,[i]} - \mathbf{I}_K(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)) \\ =& (\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathbf{u}_{1,[i]} \otimes (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)) + (\mathbf{H}_1 \otimes \mathbf{G}_1')(\mathbf{u}_{1,[i]} \otimes \mathbf{u}_{1,[i]}) \\ & - (\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathbf{u}_{1,[i]} \otimes (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)),\end{aligned}$$

which, in turn, implies $\widehat{\boldsymbol{\mathcal{Z}}}_1 = \boldsymbol{\mathcal{Z}}_1(\mathbf{H}_1 \otimes \mathbf{I}_K) + \boldsymbol{\mathcal{U}}_1(\mathbf{H}_1 \otimes \mathbf{G}_1) - (\mathbf{U}_1 \otimes (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)')(\mathbf{H}_1 \otimes \mathbf{I}_K)$, where the matrices $\boldsymbol{\mathcal{U}}_1$, $\boldsymbol{\mathcal{Z}}_1$, and $\widehat{\boldsymbol{\mathcal{Z}}}_1$ are defined in equations (75), (84), and (89), respectively.

First, we establish that $\widehat{\mathbf{c}}_{41}$ is an N-consistent estimator of $\mathbf{c}_{41}$. Using equation (125), we obtain $\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1 = \boldsymbol{\mathcal{C}}\mathbf{U}_1\mathbf{H}_1$ and so

$$\begin{aligned}(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1) =& (\mathbf{H}_1 \otimes \mathbf{I}_K)(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)\mathbf{H}_1 + (\mathbf{H}_1 \otimes \mathbf{G}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)\mathbf{H}_1 \\ & - (\mathbf{H}_1 \otimes \mathbf{I}_K)\left(\boldsymbol{\mathcal{C}}\left(\mathbf{U}_1 \otimes (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'\right)\right)'(\boldsymbol{\mathcal{C}}\mathbf{U}_1)\mathbf{H}_1.\end{aligned}$$

Combining the probability limits in (85) and (74) with the fact that $\widehat{\boldsymbol{\mu}}_\beta \xrightarrow{p} \boldsymbol{\mu}_\beta$ yields

$$(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\mathbf{U}}_1)/M_N \xrightarrow{p} (\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathbf{I}_\tau \otimes \mathbf{c}_{41})\mathbf{H}_1 + \xi_c(\mathbf{H}_1 \otimes \mathbf{G}_1)'\boldsymbol{\mathcal{N}}\mathbf{H}_1 = (\mathbf{H}_1 \otimes \mathbf{c}_{41}) + \xi_c(\mathbf{H}_1 \otimes \mathbf{G}_1)'\boldsymbol{\mathcal{N}}\mathbf{H}_1,$$

where the last equality is obtained using fact (F5). Hence, in light of definition (90), using Lemma 11, we obtain

$$\boldsymbol{\mathcal{Q}}_{41} \xrightarrow{p} \mathbf{H}_1 \otimes \mathbf{c}_{41}. \tag{93}$$

Denote by h_{rs} the (r, s) element of $\mathbf{H}_1$, for $r, s = 1, \ldots, \tau$. Using the definition of the matrix $\boldsymbol{\mathcal{L}}$, we then have

$$\begin{aligned}\boldsymbol{\mathcal{L}} \odot (\mathbf{H}_1 \otimes \mathbf{c}_{41}) =& \begin{bmatrix} \mathbf{1}_K & \mathbf{0}_K & \cdots & \mathbf{0}_K \\ \mathbf{0}_K & \mathbf{1}_K & \cdots & \mathbf{0}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_K & \mathbf{0}_K & \cdots & \mathbf{1}_K \end{bmatrix} \odot \begin{bmatrix} h_{11}\mathbf{c}_{41} & h_{12}\mathbf{c}_{41} & \cdots & h_{1\tau}\mathbf{c}_{41} \\ h_{21}\mathbf{c}_{41} & h_{22}\mathbf{c}_{41} & \cdots & h_{2\tau}\mathbf{c}_{41} \\ \vdots & \vdots & \ddots & \vdots \\ h_{\tau 1}\mathbf{c}_{41} & h_{\tau 2}\mathbf{c}_{41} & \cdots & h_{\tau\tau}\mathbf{c}_{41} \end{bmatrix} \\ =& \begin{bmatrix} h_{11}\mathbf{c}_{41} & \mathbf{0}_K & \cdots & \mathbf{0}_K \\ \mathbf{0}_K & h_{22}\mathbf{c}_{41} & \cdots & \mathbf{0}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_K & \mathbf{0}_K & \cdots & h_{\tau\tau}\mathbf{c}_{41} \end{bmatrix},\end{aligned}$$

from which, using the definition of the matrix $\boldsymbol{\mathcal{I}}$, we obtain

$$\begin{aligned}\boldsymbol{\mathcal{I}}'\left(\boldsymbol{\mathcal{L}} \odot (\mathbf{H}_1 \otimes \mathbf{c}_{41})\right)\mathbf{1}_\tau =& \begin{bmatrix} \mathbf{I}_K & \cdots & \mathbf{I}_K \end{bmatrix} \begin{bmatrix} h_{11}\mathbf{c}_{41} & \cdots & \mathbf{0}_K \\ \vdots & \ddots & \vdots \\ \mathbf{0}_K & \cdots & h_{\tau\tau}\mathbf{c}_{41} \end{bmatrix} \mathbf{1}_\tau \\ =& \begin{bmatrix} h_{11}\mathbf{c}_{41} & \cdots & h_{\tau\tau}\mathbf{c}_{41} \end{bmatrix} \mathbf{1}_\tau = \mathrm{tr}(\mathbf{H}_1)\,\mathbf{c}_{41}.\end{aligned}$$

Hence, it follows from (93) that, as $N \to \infty$, $\widehat{\mathbf{c}}_{41} = \frac{\boldsymbol{\mathcal{I}}'(\boldsymbol{\mathcal{L}} \odot \boldsymbol{\mathcal{Q}}_{41})\mathbf{1}_\tau}{\mathrm{tr}(\mathbf{H}_1)} \xrightarrow{p} \mathbf{c}_{41}$.

Second, we establish that $\widehat{\mathbf{c}}_{42}$ is an N-consistent estimator of $\mathbf{c}_{42}$. Using equation (80), we obtain $\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1 = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1(\mathbf{H}_1 \otimes \mathbf{H}_1)$ and so

$$\begin{aligned}(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{Z}}}_1)'(\boldsymbol{\mathcal{C}}\widehat{\boldsymbol{\mathcal{U}}}_1) =& (\mathbf{H}_1 \otimes \mathbf{I}_K)(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{Z}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)(\mathbf{H}_1 \otimes \mathbf{H}_1) + (\mathbf{H}_1 \otimes \mathbf{G}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)(\mathbf{H}_1 \otimes \mathbf{H}_1) \\ & - (\mathbf{H}_1 \otimes \mathbf{I}_K)\left(\boldsymbol{\mathcal{C}}\left(\mathbf{U}_1 \otimes (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'\right)\right)'(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{U}}_1)(\mathbf{H}_1 \otimes \mathbf{H}_1).\end{aligned}$$

Combining the probability limits in (76) and (86) with the fact that $\widehat{\boldsymbol{\mu}}_\beta \xrightarrow{p} \boldsymbol{\mu}_\beta$ yields

$$(\mathcal{C}\widehat{\mathcal{Z}}_1)'(\mathcal{C}\widehat{\mathcal{U}}_1)\mathcal{N}/M_N \xrightarrow{p} (\mathbf{H}_1 \otimes \mathbf{I}_K)\left(\mathcal{N}' \otimes \mathbf{c}_{42}\right)(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathcal{N} + (\mathbf{H}_1 \otimes \mathbf{G}_1)'(\kappa_c^{[1]}\mathcal{K}^{[1]} + \kappa_c^{[2]}\mathcal{K}^{[2]} + \kappa_c^{[3]}\mathcal{K}^{[3]})(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathcal{N}.$$

Using fact (F5), we obtain the equalities $(\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathcal{N}' \otimes \mathbf{c}_{42})(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathcal{N} = ((\mathbf{H}_1\mathcal{N}') \otimes \mathbf{c}_{42})(((\mathbf{H}_1 \otimes \mathbf{H}_1)\mathcal{N}) \otimes 1) = (\mathbf{H}_1\mathcal{N}'(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathcal{N}) \otimes \mathbf{c}_{42} = (\mathbf{H}_1(\mathbf{H}_1 \odot \mathbf{H}_1)) \otimes \mathbf{c}_{42}$, where the last one follows from Lemma 15. Hence, in light of definition (91), Lemma 12 then yields

$$\mathcal{Q}_{42} \xrightarrow{p} (\mathbf{H}_1(\mathbf{H}_1 \odot \mathbf{H}_1)) \otimes \mathbf{c}_{42}. \tag{94}$$

Arguing as above, we obtain

$$\mathcal{I}'(\mathcal{L} \odot ((\mathbf{H}_1(\mathbf{H}_1 \odot \mathbf{H}_1)) \otimes \mathbf{c}_{42}))\mathbf{1}_\tau = \mathrm{tr}(\mathbf{H}_1(\mathbf{H}_1 \odot \mathbf{H}_1))\mathbf{c}_{42} = \mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)\mathbf{H}_1)\mathbf{c}_{42},$$

where the last equality follows from fact (F1). Hence, it follows from (94) that, as $N \to \infty$,

$$\widehat{\mathbf{c}}_{42} = \frac{\mathcal{I}'(\mathcal{L} \odot \mathcal{Q}_{42})\mathbf{1}_\tau}{\mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)\mathbf{H}_1)} \xrightarrow{p} \mathbf{c}_{42}. \tag{95}$$

Finally, we establish that $\widehat{\mathbf{C}}_{44}$ is an N-consistent estimator of $\mathbf{C}_{44}$. Using the probability limits in (86), (88), (81), and the fact that $\widehat{\boldsymbol{\mu}}_\beta \xrightarrow{p} \boldsymbol{\mu}_\beta$, we obtain

$$\begin{aligned}\frac{1}{M_N}(\mathcal{C}\widehat{\mathcal{Z}}_1)'(\mathcal{C}\widehat{\mathcal{Z}}_1) \xrightarrow{p} &(\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathbf{I}_\tau \otimes \mathbf{C}_{44})(\mathbf{H}_1 \otimes \mathbf{I}_K)\\ &+ (\mathbf{H}_1 \otimes \mathbf{G}_1)'(\kappa_c^{[1]}\mathcal{K}^{[1]} + \kappa_c^{[2]}\mathcal{K}^{[2]} + \kappa_c^{[3]}\mathcal{K}^{[3]})(\mathbf{H}_1 \otimes \mathbf{G}_1)\\ &+ (\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathcal{N}' \otimes \mathbf{c}_{42})(\mathbf{H}_1 \otimes \mathbf{G}_1)\\ &+ [(\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathcal{N}' \otimes \mathbf{c}_{42})(\mathbf{H}_1 \otimes \mathbf{G}_1)]'.\end{aligned}$$

Then, it follows from fact (F5) and $\mathbf{H}_1$ being idempotent that $(\mathbf{H}_1 \otimes \mathbf{I}_K)(\mathbf{I}_\tau \otimes \mathbf{C}_{44})(\mathbf{H}_1 \otimes \mathbf{I}_K) = \mathbf{H}_1 \otimes \mathbf{C}_{44}$. Hence, in light of definition (92), invoking the probability limit in (95) and Lemma 12 yields

$$\mathcal{Q}_{44} \xrightarrow{p} \mathbf{H}_1 \otimes \mathbf{C}_{44}. \tag{96}$$

Consider partitioning the $(\tau K) \times (\tau K)$ matrix $\mathbf{H}_1 \otimes \mathbf{C}_{44}$ into submatrices of size $K \times K$ and let $\mathcal{W}_{rs}$ the (r, s) such submatrix of $\mathbf{H}_1 \otimes \mathbf{C}_{44}$, for $r, s = 1, \ldots, \tau$. It follows that $\mathcal{W}_{rs} = h_{rs}\mathbf{C}_{44}$ and so, using the definition of the matrices $\mathcal{J}$ and $\mathcal{I}$, we obtain that

$$\mathcal{J}' \odot (\mathbf{H}_1 \otimes \mathbf{C}_{44}) \odot \mathcal{J} = \begin{bmatrix} \mathcal{W}_{11} & \mathbf{0}_{K\times K} & \cdots & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathcal{W}_{22} & \cdots & \mathbf{0}_{K\times K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \cdots & \mathcal{W}_{\tau\tau} \end{bmatrix}$$

and $\mathcal{I}'(\mathcal{J}' \odot (\mathbf{H}_1 \otimes \mathbf{C}_{44}) \odot \mathcal{J})\mathcal{I} = \sum_{r=1}^{\tau} \mathcal{W}_{rr} = \mathrm{tr}(\mathbf{H}_1)\mathbf{C}_{44}$. Hence, it follows from (96) that, as $N \to \infty$, $\widehat{\mathbf{C}}_{44} = \frac{\mathcal{I}'(\mathcal{J}' \odot \mathcal{Q}_{44} \odot \mathcal{J})\mathcal{I}}{\mathrm{tr}(\mathbf{H}_1)} \xrightarrow{p} \mathbf{C}_{44}$, completing the proof of the lemma. ∎

References

Ang, A., and D. Kristensen, 2012, "Testing Conditional Factor Models," *Journal of Financial Economics*, 106, 132–156.

Ang, A., J. Liu, and K. Schwarz, 2010, "Using Stocks or Portfolios in Tests of Factor Models," Working Paper, Columbia University.

Armstrong, B., 1985, "Measurement Error in Generalized Linear Models," *Communications in Statistics, Series B*, 14, 529–544.

Berk, J. B., 1995, "A Critique of Size-Related Anomalies," *Review of Financial Studies*, 8, 275–286.

Black, F., M. C. Jensen, and M. Scholes, 1972, "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets*, ed. by M. C. Jensen. Praeger, New York, NY, pp. 79–121.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, 2006, *Measurement Error in Nonlinear Models: A Modern Perspective*. CRC, 2nd edn.

Carroll, R. J., and L. A. Stefanski, 1990, "Approximate Quasilikelihood Estimation in Models with Surrogate Predictors," *Journal of the American Statistical Association*, 85, 652–663.

Chen, R., and R. Kan, 2004, "Finite Sample Analysis of Two-Pass Cross-Sectional Regressions," Working Paper, University of Toronto.

Chordia, T., A. Goyal, and J. Shanken, 2015, "Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics," Working Paper, Emory University.

Connor, G., and R. A. Korajczyk, 1988, "Risk and Return in an Equilibrium APT: Application of a New Test Methodology," *Journal of Financial Economics*, 21, 255–289.

——— , 1993, "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Financial Economics*, 48, 1263–1291.

Daniel, K., and S. Titman, 2012, "Testing Factor-Model Explanations of Market Anomalies," *Critical Finance Review*, 1, 103–139.

Fama, E. F., and K. R. French, 1992, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— , 1993, "Common Risk Factors in the Returns on Bonds and Stocks," *Journal of Financial Economics*, 33, 3–56.

——— , 1996, "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51, 55–84.

Fama, E. F., and J. D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Gagliardini, P., E. Ossola, and O. Scaillet, 2012, "Time-Varying Risk Premium in Large Cross-Sectional Equity Datasets," Working Paper, University of Lugano.

Ghysels, E., 1998, "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?," *Journal of Finance*, 53, 549–573.

Gibbons, M. R., 1982, "Multivariate Tests of Financial Models: A New Approach," *Journal of Financial Economics*, 10, 3–27.

Gibbons, M. R., S. A. Ross, and J. Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121–1152.

Gleser, L. J., 1990, *Statistical Analysis of Measurement Error*. American Methematics Society, RI.

Hansen, L. P., and R. Jagannathan, 1997, "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance*, 52, 557–90.

Hardin, J. W., H. Schmiediche, and R. J. Carroll, 2003, "The Regression-Calibration Method for Fitting Generalized Linear Models with Additive Measurement Error," *The Stata Journal*, 3, 361–372.

Harvey, C. R., 1989, “Time-Varying Conditional Covariances in Tests of Asset Pricing Models,” *Journal of Financial Economics*, 24, 289–317.

Hou, K., and R. Kimmel, 2010, “On Estimation of Risk Premia in Linear Factor Models,” Working Paper, Ohio State University.

Jagannathan, R., G. Skoulakis, and Z. Wang, 2010, “The Analysis of the Cross Section of Security Returns,” in *Handbook of Financial Econometrics 2*, ed. by Y. Ait-Sahalia, and L. P. Hansen. Elsevier, UK, pp. 73–134.

Jagannathan, R., and Z. Wang, 1998, “An Asymptotic Theory for Estimating Beta-Pricing Models using Cross-Sectional Regression,” *Journal of Finance*, 53, 1285–1309.

——— , 2002, “Empirical Evaluation of Asset-Pricing Models: A Comparison of the SDF and Beta Methods,” *Journal of Finance*, 57, 2337–2367.

Jegadeesh, N., and J. Noh, 2014, “Empirical Tests of Asset Pricing Models with Individual Stocks,” Working Paper, Emory University.

Kan, R., N. Gospodinov, and C. Robotti, 2013, “Chi-Squared Tests for Evaluation and Comparison of Asset Pricing Models,” *Journal of Econometrics*, 173, 108–125.

Kan, R., and C. Robotti, 2009, “Model Comparison Using the Hansen-Jagannathan Distance,” *Review of Financial Studies*, 22, 3449–3490.

Kan, R., C. Robotti, and J. Shanken, 2013, “Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology,” Journal of Finance, forthcoming.

Kan, R., and G. Zhou, 1999, “A Critique of the Stochastic Discount Factor Methodology,” *Journal of Finance*, 54, 1021–1048.

Kim, D., 1995, “The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns,” *Journal of Finance*, 50, 1605–1634.

Lehmann, B. N., and D. M. Modest, 1988, “The Empirical Foundations of the Arbitrage Pricing Theory,” *Journal of Financial Economics*, 21, 213–254.

Lewellen, J., and S. Nagel, 2006, “The Conditional CAPM Does Not Explain Asset Pricing Anomalies,” *Journal of Financial Economics*, 79, 289–314.

Lewellen, J., S. Nagel, and J. Shanken, 2010, “A Skeptical Appraisal of Asset Pricing Tests,” *Journal of Financial Economics*, 96, 175–194.

Lintner, J., 1965, “The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economics and Statistics*, 47, 13–37.

Litzenberger, R. H., and K. Ramaswamy, 1979, “The Effect of Personal Taxes and Dividends of Capital Asset Prices: The Theory and Evidence,” *Journal of Financial Economics*, 7, 163–196.

Lo, A. W., and A. C. MacKinlay, 1990, “Data-Snooping Biases in Tests of Financial Asset Pricing Models,” *Review of Financial Studies*, 3, 431–467.

Mossin, J., 1966, “Equilibrium in a Capital Asset Market,” *Econometrica*, 34, 768–783.

Nagel, S., and K. Singleton, 2011, “Estimation and Evaluation of Conditional Asset Pricing Models,” *Journal of Finance*, 66, 873–909.

Newey, W., and K. West, 1994, “Automatic Lag Selection in Covariance Matrix Estimation,” *Review of Economic Studies*, 61(4), 631–653.

Newey, W. K., and K. D. West, 1987, “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55, 703–708.

Pastor, L., and R. F. Stambaugh, 2003, “Liquidity Risk and Expected Stock Returns,” *Journal of Political Economy*, 111(3), 642–685.

Pesaran, H., and T. Yamagata, 2012, “Testing CAPM with a Large Number of Assets,” Working Paper, University of Cambridge.

Prentice, R. L., 1982, "Covariate Measurement Errors and Parameter Estimation in Failure Time Regression Models," *Biometrika*, 69, 331–342.

Pukthuanthong, K. N., R. Roll, and J. Wang, 2014, "Resolving the Errors-in-Variables Bias in Risk Premium Estimation," Working Paper, UCLA.

Roll, R., 1977, "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics*, 4, 129–176.

Schott, J., 1997, *Matrix Analysis for Statistics*. John Wiley, New York, NY.

Shanken, J., 1985, "Multivariate Tests of the Zero-Beta CAPM," *Journal of Financial Economics*, 14, 327–348.

——— , 1992, "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 1–33.

Shanken, J., and G. Zhou, 2007, "Estimating and Testing Beta Pricing Models: Alternative Methods and Their Performance in Simulations," *Journal of Financial Economics*, 84(1), 40–86.

Sharpe, W. F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425–442.

Vasicek, O., 1973, "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas," *Journal of Fi*, 28, 1233–1239.

Zhou, G., 1991, "Small Sample Tests of Portfolio Efficiency," *Journal of Financial Economics*, 30, 165–191.

——— , 1993, "Asset Pricing Tests under Alternative Distributions," *Journal of Finance*, 48, 1927–1942.

——— , 1994, "Analytical GMM Tests: Asset Pricing with Time-Varying Risk Premiums," *Review Financial Studies*, 7, 687–709.

Table 1: **Empirical bias in the estimation of γ with normally distributed shocks: the role of the EIV correction.** This table presents simulation results on the bias (in annualized basis points) in the estimation of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \ \cdots \ \ \lambda_K\ ]'$ is the vector of ex-post prices of risk, and K is the number of factors. We estimate $\boldsymbol{\gamma}$ with and without the EIV correction. The shocks u_{it} are assumed to follow a normal distribution. The number of individual stocks, N, is equal to 2,000 and the number of clusters, M_N, is set equal to 50. The pairwise correlation of shocks, assumed to be constant within each cluster, is set equal to 0.10. The simulation is calibrated to two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. The results are based on 10,000 Monte Carlo repetitions.

Estimation Period	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-00	01-05
Testing Period	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-00	01-05	06-10
CAPM												
Without EIV correction												
α	587.8	331.0	398.8	-26.5	-65.7	221.1	138.1	177.2	277.1	572.7	20.1	71.4
λ_1	-671.9	-378.3	-455.8	30.8	76.0	-252.7	-159.1	-202.7	-316.0	-656.3	-23.0	-81.7
With EIV correction												
α	-2.9	-6.9	-3.0	-1.6	1.6	-1.2	-1.4	-0.3	-0.5	-6.4	-2.8	-1.8
λ_1	4.1	8.5	4.0	2.2	-1.0	1.6	0.5	0.4	1.6	6.5	3.2	2.1
Fama-French model												
Without EIV correction												
α	572.1	423.5	698.7	148.1	100.8	243.1	434.7	126.9	579.6	644.9	542.2	114.3
λ_1	-837.3	-461.9	-467.4	192.4	0.9	-26.7	-154.7	-375.8	-348.9	-743.3	-145.8	-26.7
λ_2	485.5	-174.3	-55.3	-492.1	94.5	-466.0	-47.0	515.9	-167.4	210.3	-433.9	-169.6
λ_3	-13.9	178.2	-633.9	-316.2	-357.2	-72.2	-699.4	17.4	-494.1	-132.9	-593.5	-50.8
With EIV correction												
α	-17.3	-39.9	-24.6	-0.1	-2.4	-1.0	-11.9	-0.1	-15.4	-20.2	-12.6	-2.6
λ_1	46.4	51.9	11.6	-25.7	-1.1	-1.0	9.4	13.0	4.1	22.8	3.2	2.8
λ_2	-54.3	32.5	4.4	35.8	0.6	8.6	-4.9	-20.8	9.7	-7.2	8.7	1.8
λ_3	-6.8	-47.1	35.3	22.6	8.3	-2.0	14.5	-5.8	19.3	6.8	16.3	0.7

Table 2: **Empirical rejection frequencies with normally distributed shocks.** This table presents simulation results on the finite-sample performance of the EIV-corrected estimator $\widetilde{\boldsymbol{\gamma}}$ of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \ \cdots \ \ \lambda_K\ ]'$ is the vector of ex-post prices of risk, and K is the number of factors. Reported are empirical p-values (in percentages) for the χ^2 statistic associated with the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ and the t statistics associated with the simple hypotheses $\alpha = 0$ as well as $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. The shocks u_{it} are assumed to follow a normal distribution. We consider two possible values for the number of clusters, M_N, namely 50 and 100. The pairwise correlation of shocks, assumed to be constant within each cluster, is denoted by ρ and takes three values, namely 0, 0.10, and 0.20. Three nominal levels of significance are considered: 1%, 5%, and 10%. The simulation is calibrated to two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. The results are based on 10,000 Monte Carlo repetitions.

ρ	0						0.10						0.20					
M_N	50			100			50			100			50			100		
Nominal Size	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10
CAPM																		
Joint χ^2	1.0	4.7	9.8	1.0	5.2	10.3	1.0	5.0	10.0	0.8	4.9	10.0	0.8	4.9	10.4	1.0	5.2	10.2
$t(\alpha)$	1.0	5.0	10.0	1.1	5.2	10.3	0.9	4.9	9.7	0.9	4.9	9.9	0.9	5.2	10.0	1.1	5.1	10.3
$t(\lambda_1)$	1.1	5.1	10.3	1.0	5.1	9.9	0.9	4.8	10.0	0.9	4.9	9.9	1.0	5.0	10.0	0.9	4.8	9.9
Fama-French model																		
Joint χ^2	1.1	5.0	10.4	1.0	5.1	10.0	1.0	4.9	9.8	0.8	4.7	9.5	1.2	5.0	10.2	1.1	5.1	9.8
$t(\alpha)$	1.1	5.2	10.1	0.8	5.2	10.1	1.0	5.0	9.9	0.9	4.6	9.6	0.9	4.7	9.5	1.1	5.3	10.3
$t(\lambda_1)$	0.9	5.0	10.4	1.0	5.2	10.4	1.0	4.7	9.6	0.8	4.5	9.5	1.0	5.0	10.1	1.1	5.3	10.0
$t(\lambda_2)$	1.0	5.1	10.2	1.0	5.3	10.4	1.1	5.1	9.9	1.0	4.9	9.8	1.1	5.0	9.9	0.9	4.7	9.9
$t(\lambda_3)$	1.1	5.3	10.4	1.0	4.8	10.0	1.1	5.1	9.8	1.0	5.2	10.2	1.0	5.0	10.3	1.0	5.0	9.9

Table 3: **Estimation results for six-year testing periods from 1951 to 2010.** This table presents the estimates of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \ \cdots \ \ \lambda_K\ ]'$ is the vector of ex-post factor risk premia, for two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. We report the estimates with and without the EIV correction as well as the realized factor average over the testing period $\bar{\mathbf{f}}_2$. The linear factor model implies that $\alpha = 0$ and $\boldsymbol{\lambda} = \bar{\mathbf{f}}_2$. We consider 10 non-overlapping six-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. All numbers are reported in annualized percentages.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
CAPM										
Without EIV correction										
α	10.96	15.08	7.18	-4.26	6.15	17.34	5.58	10.16	6.92	6.41
λ_1	3.11	-7.53	7.86	-4.97	12.75	-6.33	3.41	2.72	8.65	2.98
With EIV correction										
α	9.55	18.03	2.32	-0.63	-0.78	19.29	4.08	9.04	3.56	6.03
λ_1	4.29	-10.91	12.48	-8.07	19.01	-8.14	5.06	3.95	12.89	3.53
Realized factor average										
λ_1	16.29	7.17	9.34	-10.19	12.69	5.80	8.29	14.80	0.62	2.62
Fama-French model										
Without EIV correction										
α	11.78	14.97	7.25	-5.35	5.49	15.68	5.53	9.59	6.35	4.53
λ_1	2.75	-7.49	3.54	-3.85	9.27	-7.44	2.97	2.20	5.46	3.40
λ_2	-0.57	2.37	6.27	-2.93	9.58	0.61	1.00	2.13	7.86	3.01
λ_3	2.10	-1.89	3.45	2.09	-1.04	5.71	-1.45	0.02	-2.83	0.78
With EIV correction										
α	11.04	17.33	10.12	-2.39	-0.91	16.37	1.08	8.24	3.22	0.50
λ_1	3.83	-10.50	-3.48	-7.38	14.06	-10.78	6.00	3.13	4.95	4.77
λ_2	-1.47	3.80	12.31	-2.79	11.41	2.54	3.56	2.98	14.56	7.77
λ_3	2.98	-2.88	5.53	4.11	0.57	11.12	3.53	0.73	-1.20	4.60
Realized factor average										
λ_1	16.29	7.17	9.34	-10.19	12.69	5.80	8.29	14.80	0.62	2.62
λ_2	-3.21	0.28	12.20	-9.66	14.11	1.57	-1.64	-4.73	10.66	3.15
λ_3	2.09	2.15	6.28	6.29	2.03	12.46	0.44	3.16	7.63	1.39

Table 4: **Test statistics and p-values for six-year testing periods from 1951 to 2010 using 49 industry clusters.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 10 non-overlapping six-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
					CAPM					
Joint χ^2	37.75	128.92	50.57	6.05	38.30	70.40	5.58	58.08	252.74	36.96
	[0.000]	[0.000]	[0.000]	[0.048]	[0.000]	[0.000]	[0.062]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	3.79	10.72	0.80	-0.24	-0.20	8.32	2.19	6.42	1.90	3.55
	[0.000]	[0.000]	[0.424]	[0.813]	[0.842]	[0.000]	[0.029]	[0.000]	[0.057]	[0.000]
$t(\lambda_1)$	-5.00	-8.73	1.11	0.94	1.88	-7.40	-1.69	-7.61	5.43	0.40
	[0.000]	[0.000]	[0.267]	[0.346]	[0.060]	[0.000]	[0.092]	[0.000]	[0.000]	[0.689]
					Fama-French model					
Joint χ^2	40.97	128.38	23.36	52.92	14.28	71.04	22.38	74.29	90.90	35.83
	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	5.00	10.02	2.32	-0.60	-0.23	7.47	0.15	5.35	1.37	0.20
	[0.000]	[0.000]	[0.020]	[0.547]	[0.819]	[0.000]	[0.880]	[0.000]	[0.169]	[0.842]
$t(\lambda_1)$	-5.98	-9.82	-2.46	0.66	0.38	-8.22	-0.49	-6.36	1.68	1.10
	[0.000]	[0.000]	[0.014]	[0.509]	[0.707]	[0.000]	[0.624]	[0.000]	[0.093]	[0.270]
$t(\lambda_2)$	1.50	2.09	0.04	3.75	-1.99	0.74	1.07	5.80	1.83	1.51
	[0.133]	[0.036]	[0.966]	[0.000]	[0.047]	[0.461]	[0.286]	[0.000]	[0.067]	[0.130]
$t(\lambda_3)$	0.64	-2.45	-0.51	-1.07	-0.70	-0.63	0.26	-1.45	-4.41	1.35
	[0.523]	[0.014]	[0.607]	[0.282]	[0.481]	[0.530]	[0.799]	[0.148]	[0.000]	[0.176]

Table 5: **Estimation results for three-year testing periods from 1951 to 2010.** This table presents the estimates of $\boldsymbol{\gamma} = [\ \alpha\ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1\ \ \cdots\ \ \lambda_K\ ]'$ is the vector of ex-post factor risk premia, for two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. We report the estimates with and without the EIV correction as well as the realized factor average over the testing period $\bar{\mathbf{f}}_2$. The linear factor model implies that $\alpha = 0$ and $\boldsymbol{\lambda} = \bar{\mathbf{f}}_2$. We consider 20 non-overlapping three-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. All numbers are reported in annualized percentages.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
CAPM										
Without EIV correction										
α	7.64	10.89	18.20	11.79	6.05	9.12	-1.86	-5.23	14.88	7.36
λ_1	-1.30	10.74	-5.78	-9.86	10.78	4.43	-3.95	-8.98	10.86	6.76
With EIV correction										
α	8.22	5.97	20.58	16.62	-1.11	6.54	1.49	-1.61	8.69	4.52
λ_1	-1.78	15.21	-8.53	-15.27	17.56	6.74	-6.67	-11.95	16.41	9.32
Realized factor average										
λ_1	9.62	22.96	10.93	3.41	12.74	5.93	-3.10	-17.27	13.67	11.71
Fama-French model										
Without EIV correction										
α	6.47	13.96	17.68	12.07	4.65	11.70	-3.52	-6.04	13.11	8.52
λ_1	1.38	8.02	-5.74	-9.58	8.29	-2.14	0.03	-11.25	6.23	4.83
λ_2	-1.83	-1.45	2.67	1.06	4.78	7.54	-5.27	-0.07	11.99	4.47
λ_3	-3.09	6.57	-1.83	-0.99	4.78	1.47	0.51	5.90	0.11	-4.43
With EIV correction										
α	5.39	11.35	18.76	16.09	-1.12	20.09	-9.54	-1.80	8.90	9.29
λ_1	3.04	11.53	-7.88	-14.32	12.25	-14.75	8.16	-18.22	8.24	3.19
λ_2	-2.65	-3.22	5.36	2.12	6.17	14.40	-8.54	2.72	14.80	8.01
λ_3	-3.62	7.25	-4.18	-1.15	7.71	3.63	1.55	11.76	1.63	-9.44
Realized factor average										
λ_1	9.62	22.96	10.93	3.41	12.74	5.93	-3.10	-17.27	13.67	11.71
λ_2	-3.62	-2.80	4.00	-3.45	4.11	20.30	-6.32	-13.01	15.30	12.92
λ_3	-2.72	6.89	1.00	3.31	9.48	3.08	0.40	12.18	11.37	-7.31

Continued on next page

Table 5 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
CAPM										
Without EIV correction										
α	15.99	20.87	8.10	0.08	11.18	11.01	-1.65	17.71	2.81	5.00
λ_1	-4.44	-12.44	0.07	9.26	0.80	2.04	15.12	-0.64	1.85	3.83
With EIV correction										
α	17.52	26.34	8.07	-3.75	10.85	9.63	-7.99	17.85	2.56	3.04
λ_1	-5.83	-17.79	0.11	13.31	1.17	3.67	22.91	-0.84	2.20	5.71
Realized factor average										
λ_1	2.38	9.21	10.00	6.58	10.27	19.34	-3.66	4.89	4.89	0.35
Fama-French model										
Without EIV correction										
α	13.24	18.24	7.95	0.89	10.30	10.77	-0.73	13.55	1.39	4.80
λ_1	-6.60	-10.42	0.72	8.13	0.38	1.52	9.45	0.31	2.16	3.17
λ_2	3.57	-2.50	-0.82	1.16	2.05	1.58	11.01	3.07	2.05	3.19
λ_3	8.27	7.57	0.78	-3.30	0.95	-0.83	-7.62	3.97	0.78	0.60
With EIV correction										
α	13.43	18.36	3.52	-2.24	9.52	9.72	-2.10	4.75	-2.06	2.31
λ_1	-10.90	-11.96	4.05	11.67	0.41	2.18	6.63	1.65	3.08	3.49
λ_2	7.20	-1.94	0.92	1.02	3.13	2.56	20.16	11.90	6.21	6.32
λ_3	15.46	12.50	7.27	-4.09	2.02	-1.02	-9.88	11.10	4.13	4.50
Realized factor average										
λ_1	2.38	9.21	10.00	6.58	10.27	19.34	-3.66	4.89	4.89	0.35
λ_2	8.68	-5.53	-4.60	1.32	-0.52	-8.93	11.37	9.94	-2.95	9.24
λ_3	16.13	8.80	1.59	-0.72	5.54	0.78	7.15	8.10	2.73	0.06

Table 6: **Test statistics and p-values for three-year testing periods from 1951 to 2010 using 49 industry clusters.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 20 non-overlapping three-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
					CAPM					
Joint χ^2	23.92	6.38	99.68	34.20	12.27	65.17	7.04	19.56	66.66	2.65
p-value	[0.000]	[0.041]	[0.000]	[0.000]	[0.002]	[0.000]	[0.030]	[0.000]	[0.000]	[0.266]
$t\,(\alpha)$	2.50	1.85	8.87	5.84	-0.27	2.01	0.43	-0.45	1.76	1.35
p-value	[0.012]	[0.064]	[0.000]	[0.000]	[0.789]	[0.044]	[0.665]	[0.650]	[0.078]	[0.178]
$t\,(\lambda_1)$	-3.67	-2.24	-6.68	-5.56	1.16	0.27	-1.24	1.90	0.64	-0.82
p-value	[0.000]	[0.025]	[0.000]	[0.000]	[0.244]	[0.791]	[0.217]	[0.058]	[0.523]	[0.409]
					Fama-French model					
Joint χ^2	7.01	21.91	77.35	47.96	2.80	60.95	4.10	131.48	17.68	36.06
p-value	[0.135]	[0.000]	[0.000]	[0.000]	[0.592]	[0.000]	[0.392]	[0.000]	[0.001]	[0.000]
$t\,(\alpha)$	1.81	4.61	7.98	6.15	-0.23	4.38	-1.76	-0.44	1.73	3.07
p-value	[0.070]	[0.000]	[0.000]	[0.000]	[0.816]	[0.000]	[0.078]	[0.662]	[0.084]	[0.002]
$t\,(\lambda_1)$	-2.38	-4.51	-7.55	-5.97	-0.09	-3.90	1.95	-0.26	-1.14	-2.93
p-value	[0.017]	[0.000]	[0.000]	[0.000]	[0.929]	[0.000]	[0.052]	[0.794]	[0.255]	[0.003]
$t\,(\lambda_2)$	0.59	-0.29	0.58	2.79	0.80	-2.35	-1.09	8.16	-0.28	-2.38
p-value	[0.555]	[0.770]	[0.561]	[0.005]	[0.427]	[0.019]	[0.276]	[0.000]	[0.781]	[0.017]
$t\,(\lambda_3)$	-0.48	0.26	-1.83	-2.25	-0.97	0.18	0.40	-0.13	-3.53	-0.80
p-value	[0.632]	[0.799]	[0.068]	[0.024]	[0.333]	[0.853]	[0.688]	[0.893]	[0.000]	[0.423]

Continued on next page

Table 6 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
					CAPM					
Joint χ^2	64.76	74.94	14.02	12.52	30.16	58.10	158.59	76.67	1.12	77.80
p-value	[0.000]	[0.000]	[0.001]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]	[0.571]	[0.000]
$t(\alpha)$	5.78	7.53	3.23	-1.17	5.49	4.81	-2.97	6.35	1.06	1.12
p-value	[0.000]	[0.000]	[0.001]	[0.242]	[0.000]	[0.000]	[0.003]	[0.000]	[0.291]	[0.264]
$t(\lambda_1)$	-3.25	-8.53	-3.72	2.13	-4.66	-7.14	8.23	-1.44	-0.85	2.51
p-value	[0.001]	[0.000]	[0.000]	[0.033]	[0.000]	[0.000]	[0.000]	[0.150]	[0.394]	[0.012]
					Fama-French model					
Joint χ^2	25.54	51.01	15.25	17.23	28.20	61.69	103.73	11.23	5.19	28.29
p-value	[0.000]	[0.000]	[0.004]	[0.002]	[0.000]	[0.000]	[0.000]	[0.024]	[0.268]	[0.000]
$t(\alpha)$	4.52	3.92	0.31	-0.65	4.66	3.85	-0.67	1.24	-0.61	0.81
p-value	[0.000]	[0.000]	[0.760]	[0.517]	[0.000]	[0.000]	[0.501]	[0.214]	[0.543]	[0.417]
$t(\lambda_1)$	-4.54	-4.88	-0.85	1.47	-4.13	-5.65	2.88	-0.90	-0.66	1.13
p-value	[0.000]	[0.000]	[0.393]	[0.143]	[0.000]	[0.000]	[0.004]	[0.369]	[0.506]	[0.258]
$t(\lambda_2)$	-0.81	1.74	0.69	-0.17	2.06	5.78	2.93	0.45	2.20	-1.03
p-value	[0.420]	[0.081]	[0.491]	[0.863]	[0.039]	[0.000]	[0.003]	[0.654]	[0.028]	[0.302]
$t(\lambda_3)$	-0.22	0.99	0.31	-1.76	-1.68	-0.63	-6.57	0.86	0.43	0.88
p-value	[0.824]	[0.324]	[0.759]	[0.079]	[0.093]	[0.531]	[0.000]	[0.391]	[0.668]	[0.381]

Table 7: **Time-series regression of EIV-corrected CAPM alpha estimates on the SMB, HML, MOM, and LIQ factors.** This table presents the results of time-series regressions of monthly, quarterly, and semiannual EIV-corrected CAPM alpha estimates on the realized averages over the corresponding period of two factors of Fama and French (1993) (SMB and HML), the momentum factor of Carhart (1997) (MOM), and the liquidity factor of Pastor and Stambaugh (2003) (LIQ). We consider six different factor combinations: (i) SMB, (ii) HML, (iii) SMB and HML, (iv) SMB, HML, and MOM, (v) SMB, HML, and LIQ, and (vi) SMB, HML, MOM, and LIQ. For each regression, we report point estimates and p-values (in square brackets) for the intercept and the slopes as well as the adjusted R-squares. Standard errors are computed using the Newey and West (1987) procedure with optimal lag selected as in Newey and West (1994), with no prewhitening.

		Factor combinations					
		SMB	HML	SMB HML	SMB HML MOM	SMB HML LIQ	SMB HML MOM LIQ
				Monthly			
Intercept	estimate	0.008	0.006	0.006	0.006	0.005	0.005
	p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]
SMB	estimate	-0.126		-0.059	-0.057	-0.001	0.003
	p-value	[0.125]		[0.327]	[0.361]	[0.992]	[0.962]
HML	estimate		0.357	0.344	0.348	0.418	0.427
	p-value		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
MOM	estimate				0.016		0.036
	p-value				[0.724]		[0.423]
LIQ	estimate					0.064	0.065
	p-value					[0.082]	[0.104]
Adjusted R^2		0.015	0.117	0.120	0.119	0.192	0.193
				Quarterly			
Intercept	estimate	0.007	0.006	0.006	0.006	0.005	0.005
	p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.006]	[0.004]
SMB	estimate	-0.061		-0.022	-0.037	0.016	0.018
	p-value	[0.506]		[0.799]	[0.691]	[0.867]	[0.861]
HML	estimate		0.358	0.356	0.347	0.432	0.433
	p-value		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
MOM	estimate				-0.037		0.005
	p-value				[0.605]		[0.950]
LIQ	estimate					0.015	0.015
	p-value					[0.816]	[0.812]
Adjusted R^2		-0.001	0.118	0.114	0.113	0.182	0.177
				Semiannual			
Intercept	estimate	0.007	0.005	0.005	0.006	0.004	0.004
	p-value	[0.000]	[0.001]	[0.003]	[0.002]	[0.036]	[0.057]
SMB	estimate	-0.091		-0.008	-0.021	0.095	0.103
	p-value	[0.499]		[0.945]	[0.858]	[0.373]	[0.378]
HML	estimate		0.515	0.514	0.501	0.634	0.641
	p-value		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
MOM	estimate				-0.056		0.026
	p-value				[0.530]		[0.719]
LIQ	estimate					0.041	0.044
	p-value					[0.747]	[0.713]
Adjusted R^2		-0.003	0.207	0.200	0.197	0.331	0.324

Online Appendix for "Ex-post Risk Premia: Estimation and Inference using Large Cross Sections"

In this Online Appendix, we present (i) the proofs not included in the Appendix; (ii) a number of auxiliary lemmas; (iii) additional simulation results under the assumption that the shocks follow a Student-t distribution; and (iv) robustness checks of the empirical results where we use clustering based on size and book-to-market.

Additional Proofs

Proof of Proposition 1: Since $\widehat{\mathbf{X}} = [\ \mathbf{1}_N \quad \widehat{\mathbf{B}}\ ]$, we have

$$\widehat{\mathbf{X}}'\widehat{\mathbf{X}}/N = \begin{bmatrix} \mathbf{1}_N'\mathbf{1}_N/N & (\widehat{\mathbf{B}}'\mathbf{1}_N/N)' \\ \widehat{\mathbf{B}}'\mathbf{1}_N/N & \widehat{\mathbf{B}}'\widehat{\mathbf{B}}/N \end{bmatrix}.$$

Using equation (9), we obtain

$$\widehat{\mathbf{B}}'\mathbf{1}_N/N = \overline{\boldsymbol{\beta}} + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/N) \xrightarrow{p} \boldsymbol{\mu}_\beta, \tag{97}$$

by invoking Assumptions 1(i) and 1(iii). Next, we express $\widehat{\mathbf{B}}'\widehat{\mathbf{B}}/N$ as follows:

$$\widehat{\mathbf{B}}'\widehat{\mathbf{B}}/N = (\widehat{\mathbf{B}} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'(\widehat{\mathbf{B}} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N + (\widehat{\mathbf{B}}'\mathbf{1}_N/N)\boldsymbol{\mu}_\beta' + \boldsymbol{\mu}_\beta(\widehat{\mathbf{B}}'\mathbf{1}_N/N)' - \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'. \tag{98}$$

Furthermore, using equation (9) again and Assumptions 1(i), 1(ii), and 1(iv), we obtain

$$\begin{aligned} & (\widehat{\mathbf{B}} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'(\widehat{\mathbf{B}} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N \\ = \quad & (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_1/N)\mathbf{G}_1 \\ & + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{B}/N) + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/N)\boldsymbol{\mu}_\beta' + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{B}/N) + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{1}_N/N)\boldsymbol{\mu}_\beta' \\ \xrightarrow{p} \quad & \mathbf{V}_\beta + v\mathbf{W}_1 = \mathbf{L}_1, \end{aligned} \tag{99}$$

according to definition (16). Combining (97), (98), and (99) yields

$$\widehat{\mathbf{X}}'\widehat{\mathbf{X}}/N \xrightarrow{p} \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \\ \boldsymbol{\mu}_\beta & \mathbf{L}_1 + \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta' \end{bmatrix},$$

which, upon using the formula for the inverse of a partitioned matrix (see Theorem 7.1 in Schott (1997)), implies

$$(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}/N)^{-1} \xrightarrow{p} \begin{bmatrix} 1 + \boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1}\boldsymbol{\mu}_\beta & -\boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1} \\ -\mathbf{L}_1^{-1}\boldsymbol{\mu}_\beta & \mathbf{L}_1^{-1} \end{bmatrix}.$$

It then follows from expression (13) and the probability limit in (14) that

$$\begin{aligned} \widehat{\boldsymbol{\gamma}} \quad & \xrightarrow{p} \quad \boldsymbol{\gamma} - \begin{bmatrix} 1 + \boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1}\boldsymbol{\mu}_\beta & -\boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1} \\ -\mathbf{L}_1^{-1}\boldsymbol{\mu}_\beta & \mathbf{L}_1^{-1} \end{bmatrix} \begin{bmatrix} 0 & \mathbf{0}_K' \\ \mathbf{0}_K & v\mathbf{W}_1 \end{bmatrix} \boldsymbol{\gamma} \\ & = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1}(v\mathbf{W}_1) \\ \mathbf{0}_K & \mathbf{I}_K - \mathbf{L}_1^{-1}(v\mathbf{W}_1) \end{bmatrix} \boldsymbol{\gamma} = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta'\mathbf{L}_1^{-1}(v\mathbf{W}_1) \\ \mathbf{0}_K & \mathbf{L}_1^{-1}\mathbf{V}_\beta \end{bmatrix} \boldsymbol{\gamma}, \end{aligned}$$

completing the proof of the proposition. ■

Proof of Lemma 2: First, using (17) and (20), we express $\breve{\mathbf{B}}$ and $\breve{\boldsymbol{\omega}}$ as follows:

$$\breve{\mathbf{B}} = \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \mathbf{J}_N\widehat{\mathbf{B}}\left(\mathbf{I}_K - \mathbf{C}\right), \tag{100}$$

and

$$\breve{\boldsymbol{\omega}} = (\mathbf{J}_N\widehat{\mathbf{B}}\mathbf{C} - \mathbf{U}_1\mathbf{G}_1)\boldsymbol{\lambda} + \overline{\mathbf{u}}_2. \tag{101}$$

Then, we can express $\mathbf{1}_N'\breve{\boldsymbol{\omega}}_T$ as follows:

$$\mathbf{1}_N'\breve{\boldsymbol{\omega}}_T = \mathbf{1}_N'\mathbf{J}_N\widehat{\mathbf{B}}\mathbf{C}\boldsymbol{\lambda} - \mathbf{1}_N'\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda} + \mathbf{1}_N'\overline{\mathbf{u}}_2 = -\mathbf{1}_N'\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda} + \frac{1}{T-\tau}\mathbf{1}_N'\mathbf{U}_2\mathbf{1}_{T-\tau},$$

where the second equality follows from the property $\mathbf{J}_N\mathbf{1}_N = \mathbf{0}_N$ and equation (8). Note that Assumption 1(i) implies

$$\mathbf{1}_N'\mathbf{U}_1/N \xrightarrow{p} \mathbf{0}_\tau' \tag{102}$$

and

$$\mathbf{1}_N'\mathbf{U}_2/N \xrightarrow{p} \mathbf{0}_{T-\tau}'. \tag{103}$$

Thus, as $N \to \infty$, $\mathbf{1}_N'\breve{\boldsymbol{\omega}}_T/N \xrightarrow{p} -\mathbf{0}_\tau'\mathbf{G}_1\boldsymbol{\lambda} = 0$. Using equations (100) and (101) and invoking the property $\mathbf{J}_N\mathbf{1}_N = \mathbf{0}_N$, we can express the product $\breve{\mathbf{B}}'\breve{\boldsymbol{\omega}}_T$ as the sum of five terms as follows:

$$\breve{\mathbf{B}}'\breve{\boldsymbol{\omega}}_T = \mathbf{h}^{(1)} + \mathbf{h}^{(2)} + \mathbf{h}^{(3)} + \mathbf{h}^{(4)} + \mathbf{h}^{(5)}, \tag{104}$$

where

$$\mathbf{h}^{(1)} = -\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N'\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda}, \tag{105}$$

$$\mathbf{h}^{(2)} = \widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N'\overline{\mathbf{u}}_2 = \frac{1}{T-\tau}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N'\mathbf{U}_2\mathbf{1}_{T-\tau}, \tag{106}$$

$$\mathbf{h}^{(3)} = (\mathbf{I}_K - \mathbf{C})'\,\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}\mathbf{C}\boldsymbol{\lambda}, \tag{107}$$

$$\mathbf{h}^{(4)} = -\,(\mathbf{I}_K - \mathbf{C})'\,\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1\mathbf{G}_1\boldsymbol{\lambda}, \tag{108}$$

$$\mathbf{h}^{(5)} = (\mathbf{I}_K - \mathbf{C})'\,\widehat{\mathbf{B}}'\mathbf{J}_N\overline{\mathbf{u}}_2 = \frac{1}{T-\tau}\,(\mathbf{I}_K - \mathbf{C})'\,\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2\mathbf{1}_{T-\tau}. \tag{109}$$

In what follows, we obtain the probability limits of $\mathbf{h}^{(m)}/N$, as $N \to \infty$, for $m = 1, \ldots, 5$. Note that Assumption 1(i) implies that, as $N \to \infty$,

$$(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/N \xrightarrow{p} \mathbf{0}_{K\times\tau} \tag{110}$$

and

$$(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/N \xrightarrow{p} \mathbf{0}_{K\times(T-\tau)}. \tag{111}$$

Moreover, Assumption 1(ii) implies

$$\mathbf{U}_1'\mathbf{U}_1/N \xrightarrow{p} v\mathbf{I}_\tau \tag{112}$$

and

$$\mathbf{U}_1'\mathbf{U}_2/N \xrightarrow{p} \mathbf{0}_{\tau\times(T-\tau)}. \tag{113}$$

Using (103), (102), and (18), we immediately obtain from (105) and (106), respectively, that, as $N \to \infty$,

$$\mathbf{h}^{(1)}/N \xrightarrow{p} \mathbf{0}_K \tag{114}$$

and

$$\mathbf{h}^{(2)}/N \xrightarrow{p} \mathbf{0}_K. \tag{115}$$

Next, we obtain the probability limit of $\mathbf{h}^{(3)}/N$. Using equation (9) we obtain

$$\mathbf{J}_N\widehat{\mathbf{B}} = \widehat{\mathbf{B}} - \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' = (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta') + \boldsymbol{\Xi} - \mathbf{1}_N(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'$$

which yields

$$\begin{aligned} & \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N = (\mathbf{J}_N\widehat{\mathbf{B}})'(\mathbf{J}_N\widehat{\mathbf{B}})/N \\ = \; & (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N + (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\boldsymbol{\Xi}/N + \boldsymbol{\Xi}'(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N \\ & -(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{1}_N(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'/N - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)\mathbf{1}_N'(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')/N \\ & +\boldsymbol{\Xi}'\boldsymbol{\Xi}/N - \boldsymbol{\Xi}'\mathbf{1}_N(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'/N - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)\mathbf{1}_N'\boldsymbol{\Xi}/N + (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'. \end{aligned} \tag{116}$$

Next, note that the probability limit in (110) implies

$$(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\boldsymbol{\Xi}/N = ((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/N)\mathbf{G}_1 \xrightarrow{p} \mathbf{0}_{K\times K}. \tag{117}$$

Furthermore, according to Assumption 1(iii), we have

$$(\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{1}_N/N = \overline{\boldsymbol{\beta}} - \boldsymbol{\mu}_\beta \to \mathbf{0}_K, \tag{118}$$

and the probability limit in (102) implies

$$\mathbf{1}_N'\boldsymbol{\Xi}/N = (\mathbf{1}_N'\mathbf{U}_1/N)\mathbf{G}_1 \xrightarrow{p} \mathbf{0}_K'. \tag{119}$$

Finally, the probability limit in (112) implies

$$\boldsymbol{\Xi}'\boldsymbol{\Xi}/N = \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_1/N)\mathbf{G}_1 \xrightarrow{p} v\mathbf{W}_1, \tag{120}$$

where the matrix $\mathbf{W}_1$ is defined in equation (15). Thus, in light of Assumption 1(iv) and equations (117), (118), (119), and (120), it follows from (116) that, as $N \to \infty$,

$$\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N \xrightarrow{p} \mathbf{V}_\beta + v\mathbf{W}_1. \tag{121}$$

Then, it follows from (107) that, as $N \to \infty$,

$$\mathbf{h}^{(3)}/N \xrightarrow{p} (\mathbf{I}_K - \mathbf{C})'\left(\mathbf{V}_\beta + v\mathbf{W}_1\right)\mathbf{C}\boldsymbol{\lambda}. \tag{122}$$

Next, we obtain the probability limit of $\mathbf{h}^{(4)}/N$. Using equation (9), we obtain

$$\begin{aligned} & \widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1/N = (\mathbf{J}_N\widehat{\mathbf{B}})'\mathbf{U}_1/N = ((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta') + \boldsymbol{\Xi} - \mathbf{1}_N(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)')'\mathbf{U}_1/N \\ = \; & (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1/N + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_1/N) + (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)(\mathbf{1}_N'\mathbf{U}_1/N). \end{aligned}$$

Hence, using (18) and the probability limits in (110), (112) and (102), we obtain

$$\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_1/N \xrightarrow{p} v\mathbf{G}_1'.$$

Then, the probability limit of $\mathbf{h}^{(4)}/N$, as $N \to \infty$, is equal to $-\left(\mathbf{I}_K - \mathbf{C}\right)'\left(v\mathbf{G}_1\right)'\mathbf{G}_1\boldsymbol{\lambda}$, and so

$$\mathbf{h}^{(4)}/N \xrightarrow{p} -(\mathbf{I}_K - \mathbf{C})'(v\mathbf{W}_1)\boldsymbol{\lambda}, \tag{123}$$

where the matrix $\mathbf{W}_1$ is defined in equation (15).

Finally, we obtain the probability limit of $\mathbf{h}^{(5)}/N$. Using equation (9) we obtain

$$\begin{aligned} & \widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2/N = (\mathbf{J}_N\widehat{\mathbf{B}})'\mathbf{U}_2/N = ((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta') + \boldsymbol{\Xi} - \mathbf{1}_N(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)')'\mathbf{U}_2/N \\ = \; & (\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2/N + \mathbf{G}_1'(\mathbf{U}_1'\mathbf{U}_2/N) - (\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)(\mathbf{1}_N'\mathbf{U}_2/N). \end{aligned}$$

Using (18) and the probability limits in (111), (113) and (103), we obtain that, as $N \to \infty$, $\widehat{\mathbf{B}}'\mathbf{J}_N\mathbf{U}_2/N \xrightarrow{p} \mathbf{0}_{K\times(T-\tau)}$ and so

$$\mathbf{h}^{(5)}/N \xrightarrow{p} \mathbf{0}_K. \tag{124}$$

Collecting the results in (114), (115), (122), (123) and (124), we obtain that, as $N \to \infty$,

$$\breve{\mathbf{B}}'\breve{\boldsymbol{\omega}}/N \xrightarrow{p} (\mathbf{I}_K - \mathbf{C})'\left[\left(\mathbf{V}_\beta + v\mathbf{W}_1\right)\mathbf{C} - v\mathbf{W}_1\right]\boldsymbol{\lambda},$$

completing the proof of the lemma. ■

Proof of Lemma 3: Using equations (22) and (23), we obtain $\widehat{\mathbf{U}}_1 = (\mathbf{R}_1 - \bar{\mathbf{r}}_1\mathbf{1}_\tau') - \widehat{\mathbf{B}}(\mathbf{F}_1 - \bar{\mathbf{f}}_1\mathbf{1}_\tau')$ which, in light of equation (9), yields

$$\widehat{\mathbf{U}}_1 = \mathbf{R}_1\mathbf{J}_\tau - \widehat{\mathbf{B}}\mathbf{F}_1\mathbf{J}_\tau = \left(\mathbf{R}_1 - \mathbf{B}\mathbf{F}_1 - (\mathbf{U}_1\mathbf{G}_1)\mathbf{F}_1\right)\mathbf{J}_\tau = \left(\mathbf{U}_1 - \mathbf{U}_1(\mathbf{G}_1\mathbf{F}_1)\right)\mathbf{J}_\tau = \mathbf{U}_1\mathbf{H}_1, \tag{125}$$

since, according to definitions (10) and (24), we have

$$\left(\mathbf{I}_\tau - \mathbf{G}_1\mathbf{F}_1\right)\mathbf{J}_\tau = \mathbf{J}_\tau - \mathbf{J}_\tau\mathbf{F}_1'\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1}\mathbf{F}_1\mathbf{J}_\tau = \mathbf{H}_1.$$

It then follows from Assumption 1(ii) that, as $N \to \infty$,

$$\mathrm{tr}(\widehat{\mathbf{U}}_1'\widehat{\mathbf{U}}_1/N) = \mathrm{tr}(\mathbf{H}_1'(\mathbf{U}_1'\mathbf{U}_1/N)\mathbf{H}_1) \xrightarrow{p} \mathrm{tr}\left(\mathbf{H}_1'\left(v\mathbf{I}_\tau\right)\mathbf{H}_1\right) = v\mathrm{tr}\left(\mathbf{H}_1'\mathbf{H}_1\right) = v\mathrm{tr}\left(\mathbf{H}_1\right),$$

where we use the fact that $\mathbf{H}_1$ is symmetric and idempotent. Hence, as $N \to \infty$,

$$\widehat{v} = \frac{\mathrm{tr}(\widehat{\mathbf{U}}_1'\widehat{\mathbf{U}}_1/N)}{\mathrm{tr}\left(\mathbf{H}_1\right)} \xrightarrow{p} v,$$

and so the proof of the lemma is complete. ■

Proof of Lemma 6: First note that, using equations (6) and (7), we obtain

$$\mathbf{U}_1'\mathbf{1}_N = \sum\nolimits_{i=1}^N \mathbf{u}_{1,[i]}, \quad \mathbf{U}_2'\mathbf{1}_N = \sum\nolimits_{i=1}^N \mathbf{u}_{2,[i]}, \quad \mathbf{U}_1'\mathbf{U}_2 = \sum\nolimits_{i=1}^N \mathbf{u}_{1,[i]}\mathbf{u}_{2,[i]}'.$$

Invoking (6) and (7) again and using fact (F2) yields

$$\mathrm{vec}(\mathbf{U}_1'\mathbf{U}_1 - Nv\mathbf{I}_\tau) = \mathrm{vec}\left(\sum\nolimits_{i=1}^N \left(\mathbf{u}_{1,[i]}\mathbf{u}_{1,[i]}' - v\mathbf{I}_\tau\right)\right) = \sum\nolimits_{i=1}^N \left(\mathbf{u}_{1,[i]} \otimes \mathbf{u}_{1,[i]} - v\mathbf{i}_\tau\right),$$

and

$$\mathrm{vec}(\mathbf{U}_1'\mathbf{U}_2) = \mathrm{vec}\left(\sum\nolimits_{i=1}^N \mathbf{u}_{1,[i]}\mathbf{u}_{2,[i]}'\right) = \sum\nolimits_{i=1}^N \mathbf{u}_{2,[i]} \otimes \mathbf{u}_{1,[i]},$$

where the vector $\mathbf{i}_\tau$ is defined by (35). Finally, using equations (6) and (7) and fact (F2) again, we obtain

$$\mathrm{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_1) = \mathrm{vec}\left(\sum\nolimits_{i=1}^N (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)\,\mathbf{u}_{1,[i]}'\right) = \sum\nolimits_{i=1}^N \mathbf{u}_{1,[i]} \otimes (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta),$$

and

$$\mathrm{vec}((\mathbf{B} - \mathbf{1}_N\boldsymbol{\mu}_\beta')'\mathbf{U}_2) = \mathrm{vec}\left(\sum\nolimits_{i=1}^N (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)\,\mathbf{u}_{2,[i]}'\right) = \sum\nolimits_{i=1}^N \mathbf{u}_{2,[i]} \otimes (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta).$$

Collecting the above expressions and using the definition of $\mathbf{e}$ in (33), we obtain $\mathbf{e} = \sum_{i=1}^N \mathbf{e}_i/N$, where $\mathbf{e}_i$ is defined by (34), and so the proof of the lemma is complete. ■

Auxiliary Lemmas

Lemma 14 *The following identity holds:*

$$(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}' = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}} - N\widehat{v}\mathbf{M}'(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1')^{-1}\mathbf{M})^{-1}\widehat{\mathbf{X}}', \tag{126}$$

where $\widehat{\mathbf{X}}$, $\mathbf{F}_1$, $\widehat{v}$ *and* $\widetilde{\mathbf{X}}$ *are defined in (5), (3), (25) and (29), respectively, and* $\mathbf{M} = \begin{bmatrix} \mathbf{0}_K & \mathbf{I}_K \end{bmatrix}$.

Proof of Lemma 14: First, observe that

$$\begin{aligned}(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}' &= \begin{bmatrix} 1 & \frac{1}{N}\mathbf{1}_N'\widetilde{\mathbf{B}} \\ \frac{1}{N}\widetilde{\mathbf{B}}'\mathbf{1}_N & \frac{1}{N}\widetilde{\mathbf{B}}'\widetilde{\mathbf{B}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widetilde{\mathbf{B}}' \end{bmatrix} \\ &= \begin{bmatrix} 1 & \widehat{\boldsymbol{\mu}}_\beta' \\ \widehat{\boldsymbol{\mu}}_\beta & \widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + \boldsymbol{\Upsilon} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N' + \frac{1}{N}(\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N \end{bmatrix},\end{aligned}$$

where $\boldsymbol{\Upsilon} = (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}})/N$, since, according to equation (28), $\widetilde{\mathbf{B}} = \mathbf{1}_N\widehat{\boldsymbol{\mu}}_\beta' + \mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}})$ from which we also obtain that $\widetilde{\mathbf{B}}'\mathbf{1}_N/N = \widehat{\boldsymbol{\mu}}_\beta$ and $\widetilde{\mathbf{B}}'\widetilde{\mathbf{B}}/N = \widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + (\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}(\mathbf{I}_K - \widehat{\mathbf{C}})/N$, in light of the property $\mathbf{J}_N'\mathbf{1}_N = \mathbf{0}_N$ and the fact that $\mathbf{J}_N$ is an idempotent matrix. Using the formula for the inverse of a partitioned matrix (see Theorem 7.1 in Schott (1997)), we obtain

$$\begin{bmatrix} 1 & \widehat{\boldsymbol{\mu}}_\beta' \\ \widehat{\boldsymbol{\mu}}_\beta & \widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + \boldsymbol{\Upsilon} \end{bmatrix}^{-1} = \begin{bmatrix} 1 + \widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Upsilon}^{-1}\widehat{\boldsymbol{\mu}}_\beta & -\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Upsilon}^{-1} \\ -\boldsymbol{\Upsilon}^{-1}\widehat{\boldsymbol{\mu}}_\beta & \boldsymbol{\Upsilon}^{-1} \end{bmatrix}.$$

It follows from equations (26), (15), and (10) that $\mathbf{I}_K - \widehat{\mathbf{C}} = (\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N)^{-1}(\widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N - \widehat{v}\,(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1')^{-1})$ which then yields $\boldsymbol{\Upsilon} = (\mathbf{I}_K - \widehat{\mathbf{C}})'\boldsymbol{\Phi}$, where $\boldsymbol{\Phi} = \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N - \widehat{v}\,(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1')^{-1}$. Hence, upon using the definition of $\mathbf{J}_N$, we obtain that

$$\begin{aligned}(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}' &= \begin{bmatrix} 1 + \widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Upsilon}^{-1}\widehat{\boldsymbol{\mu}}_\beta & -\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Upsilon}^{-1} \\ -\boldsymbol{\Upsilon}^{-1}\widehat{\boldsymbol{\mu}}_\beta & \boldsymbol{\Upsilon}^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N' + \frac{1}{N}(\mathbf{I}_K - \widehat{\mathbf{C}})'\widehat{\mathbf{B}}'\mathbf{J}_N \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' - \frac{1}{N}\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Phi}^{-1}\widehat{\mathbf{B}}'\mathbf{J}_N \\ \frac{1}{N}\boldsymbol{\Phi}^{-1}\widehat{\mathbf{B}}'\mathbf{J}_N \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' + \frac{1}{N}\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Phi}^{-1}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N' - \frac{1}{N}\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Phi}^{-1}\widehat{\mathbf{B}}' \\ -\frac{1}{N}\boldsymbol{\Phi}^{-1}\widehat{\boldsymbol{\mu}}_\beta\mathbf{1}_N' + \frac{1}{N}\boldsymbol{\Phi}^{-1}\widehat{\mathbf{B}}' \end{bmatrix} \\ &= \begin{bmatrix} 1 + \widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Phi}^{-1}\widehat{\boldsymbol{\mu}}_\beta & -\widehat{\boldsymbol{\mu}}_\beta'\boldsymbol{\Phi}^{-1} \\ -\boldsymbol{\Phi}^{-1}\widehat{\boldsymbol{\mu}}_\beta & \boldsymbol{\Phi}^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widehat{\mathbf{B}}' \end{bmatrix} = \begin{bmatrix} 1 & \widehat{\boldsymbol{\mu}}_\beta' \\ \widehat{\boldsymbol{\mu}}_\beta & \widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + \boldsymbol{\Phi} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widehat{\mathbf{B}}' \end{bmatrix},\end{aligned}$$

where the last equality is obtained by invoking again the formula for the inverse of a partitioned matrix. Finally, noticing that

$$(\widehat{\mathbf{X}}'\widehat{\mathbf{X}} - N\widehat{v}\mathbf{M}'\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1}\mathbf{M})^{-1}\widehat{\mathbf{X}}' = \begin{bmatrix} 1 & \widehat{\boldsymbol{\mu}}_\beta' \\ \widehat{\boldsymbol{\mu}}_\beta & \widehat{\mathbf{B}}'\widehat{\mathbf{B}} - \widehat{v}\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N}\mathbf{1}_N' \\ \frac{1}{N}\widehat{\mathbf{B}}' \end{bmatrix}$$

and that $\widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + \boldsymbol{\Phi} = \widehat{\boldsymbol{\mu}}_\beta\widehat{\boldsymbol{\mu}}_\beta' + \widehat{\mathbf{B}}'\mathbf{J}_N\widehat{\mathbf{B}}/N - \widehat{v}\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1} = \widehat{\mathbf{B}}'\widehat{\mathbf{B}}/N - \widehat{v}\left(\mathbf{F}_1\mathbf{J}_\tau\mathbf{F}_1'\right)^{-1}$, which follows from the definition of $\mathbf{J}_N$, shows that the equality in (126) holds and completes the proof of the lemma. ■

The following lemma is used in various proofs.

Lemma 15 *For each positive integer p, let $\ell_p \in \{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). Let $\boldsymbol{\mathcal{N}}$ be the $\tau^2 \times \tau$ matrix defined by $\boldsymbol{\mathcal{N}}(p,r) = \mathbf{1}_{[k_p=\ell_p=r]}$, for $p = 1,\ldots,\tau^2$ and $r = 1,\ldots,\tau$. Then, for any $\tau \times \tau$ matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\boldsymbol{\mathcal{N}}'\left(\mathbf{A}\otimes\mathbf{B}\right)\boldsymbol{\mathcal{N}} = \mathbf{A}\odot\mathbf{B}$, where $\otimes$ and $\odot$ denote the Kronecker and Hadamard products, respectively.*

<u>**Proof of Lemma 15**</u>: Define the $\tau^2 \times \tau^2$ matrix $\mathbf{C}$ by

$$\mathbf{C} = \mathbf{A}\otimes\mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1\tau}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2\tau}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\tau 1}\mathbf{B} & a_{\tau 2}\mathbf{B} & \cdots & a_{\tau\tau}\mathbf{B} \end{bmatrix}.$$

It follows that the (p,q) element of $\mathbf{C}$ is $c_{p,q} = a_{k_pk_q}b_{\ell_p\ell_q}$, for $p,q = 1,\ldots,\tau^2$. Then, the (p,s) element of the $\tau^2\times\tau$ matrix $\left(\mathbf{A}\otimes\mathbf{B}\right)\boldsymbol{\mathcal{N}} = \mathbf{C}\boldsymbol{\mathcal{N}}$ is $\sum_{q=1}^{\tau^2} c_{p,q}\mathbf{1}_{[k_q=\ell_q=s]} = a_{k_ps}b_{\ell_ps}$. Hence, the (r,s) element of the $\tau\times\tau$ matrix $\boldsymbol{\mathcal{N}}'\left(\mathbf{A}\otimes\mathbf{B}\right)\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{N}}'\mathbf{C}\boldsymbol{\mathcal{N}}$ is $\sum_{p=1}^{\tau^2}\mathbf{1}_{[k_p=\ell_p=r]}a_{k_ps}b_{\ell_ps} = a_{rs}b_{rs}$ which is the (r,s) element of $\mathbf{A}\odot\mathbf{B}$. ■

The following lemma is used in the proof of Theorem 7 in the Appendix.

Lemma 16 *The following expression holds*

$$\sqrt{N}(\widehat{v} - v) = \frac{1}{tr(\mathbf{H}_1)} vec\left(\mathbf{H}_1\right)'\, vec(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)),$$

where $\widehat{v}$ is defined in (25), the matrix $\mathbf{U}_1$ is defined in (6), and the matrix $\mathbf{H}_1$ is defined in (24).

<u>**Proof of Lemma 16**</u>: It follows from the proof of Lemma 3 that $\widehat{v} = \frac{1}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{tr}(\mathbf{H}_1'(\frac{1}{N}\mathbf{U}_1'\mathbf{U}_1)\mathbf{H}_1)$ and $\mathrm{tr}(\mathbf{H}_1'\mathbf{H}_1) = \mathrm{tr}(\mathbf{H}_1)$. Using of these two properties, we obtain

$$\begin{aligned}
\sqrt{N}\left(\widehat{v} - v\right) &= \frac{1}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{tr}(\mathbf{H}_1'\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)\mathbf{H}_1) \\
&= \frac{1}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{tr}(\mathbf{H}_1\mathbf{H}_1'\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)) \\
&= \frac{1}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{tr}(\mathbf{H}_1\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)) \\
&= \frac{1}{\mathrm{tr}(\mathbf{H}_1)}\mathrm{vec}(\mathbf{H}_1)'\mathrm{vec}(\sqrt{N}(\mathbf{U}_1'\mathbf{U}_1/N - v\mathbf{I}_\tau)),
\end{aligned}$$

where the second equality follows from fact (F1), the third equality follows from $\mathbf{H}_1$ being symmetric and idempotent, and the fourth equality follows from fact (F3). ■

Lemmas 17 and 18, stated and proved next, are used in the proofs of Lemmas 19, 20, and 21, that are, in turn, used in the proof of Lemma 12 in the Appendix.

Lemma 17 *For the matrix $\mathbf{H}_1$ defined in equation (24), we have*

$$\sqrt{tr\left(\mathbf{H}_1\otimes\mathbf{H}_1\right)} = \mathbf{1}_\tau'\left(\mathbf{H}_1\odot\mathbf{H}_1\right)\mathbf{1}_\tau = \mathbf{i}_\tau'\left(\mathbf{H}_1\otimes\mathbf{H}_1\right)\mathbf{i}_\tau = tr\left(\mathbf{H}_1\right)$$

and

$$\mathbf{1}_\tau'\left(\mathbf{H}_1\odot\mathbf{H}_1\right)\left(\mathbf{H}_1\odot\mathbf{H}_1\right)\mathbf{1}_\tau = tr\left(\mathbf{H}_1\odot\mathbf{H}_1\right).$$

Proof of Lemma 17: A number of implications follow from the fact that the matrix $\mathbf{H}_1$ is symmetric and idempotent. First, fact (F6) implies that $\text{tr}\,(\mathbf{H}_1 \otimes \mathbf{H}_1) = \text{tr}\,(\mathbf{H}_1)^2$. Second, it follows from fact (F7) that $\mathbf{1}_\tau' \,(\mathbf{H}_1 \odot \mathbf{H}_1)\, \mathbf{1}_\tau = \text{tr}\,(\mathbf{H}_1\mathbf{H}_1') = \text{tr}\,(\mathbf{H}_1)$. Third, since $\mathbf{i}_\tau = \text{vec}\,(\mathbf{I}_\tau)$, we have

$$\begin{aligned}\mathbf{i}_\tau' \,(\mathbf{H_1} \otimes \mathbf{H_1})\, \mathbf{i}_\tau &= \mathbf{i}_\tau' \,(\mathbf{H_1} \otimes \mathbf{H_1})\,(\mathbf{H_1} \otimes \mathbf{H_1})\, \mathbf{i}_\tau = \text{vec}\,(\mathbf{H}_1\mathbf{I}_\tau\mathbf{H}_1)'\,\text{vec}\,(\mathbf{H}_1\mathbf{I}_\tau\mathbf{H}_1)\\ &= \text{vec}\,(\mathbf{H}_1)'\,\text{vec}\,(\mathbf{H}_1) = \text{tr}\left(\mathbf{H}_1'\mathbf{H}_1\right) = \text{tr}(\mathbf{H}_1),\end{aligned}$$

where the first equality follows from fact (F5) and $\mathbf{H}_1$ being idempotent, the second equality follows from fact (F4), the third equality is due to $\mathbf{H}_1$ being idempotent, the fourth equality follows from fact (F3), and the fifth equality is due to $\mathbf{H}_1$ being symmetric and idempotent. Finally, note that $\mathbf{1}_\tau' \,(\mathbf{H}_1 \odot \mathbf{H}_1)\,(\mathbf{H}_1 \odot \mathbf{H}_1)\, \mathbf{1}_\tau = \mathbf{h}_{\mathtt{H}}'\mathbf{h}_{\mathtt{H}}$, where $\mathbf{h}_{\mathtt{H}} = (\mathbf{H}_1 \odot \mathbf{H}_1)\, \mathbf{1}_\tau$. Let h_{rs} denote the (r,s) element of the matrix $\mathbf{H}_1$. Then, the r-th element of the $\tau \times 1$ vector $\mathbf{h}_{\mathtt{H}}$ is $\sum_{s=1}^{\tau} h_{rs}^2 = \sum_{s=1}^{\tau} h_{rs}h_{sr} = h_{rr}$ since $\mathbf{H}_1$ is symmetric and idempotent. Hence, $\mathbf{h}_{\mathtt{H}}'\mathbf{h}_{\mathtt{H}} = \sum_{r=1}^{\tau} h_{rr}^2 = \text{tr}\,(\mathbf{H}_1 \odot \mathbf{H}_1)$ and so $\mathbf{1}_\tau' \,(\mathbf{H}_1 \odot \mathbf{H}_1)\,(\mathbf{H}_1 \odot \mathbf{H}_1)\, \mathbf{1}_\tau = \text{tr}\,(\mathbf{H}_1 \odot \mathbf{H}_1)$. ■

Lemma 18 *For a given $\tau^2 \times 1$ vector $\mathbf{x}$, let $\mathbf{X}$ be the $\tau \times \tau$ matrix such that $\mathbf{x} = vec\,(\mathbf{X})$. Then,*

$$\mathbf{x}'\boldsymbol{\mathcal{K}}^{[1]}\mathbf{x} = tr\,(\mathbf{X} \odot \mathbf{X})$$

and

$$\mathbf{x}'\boldsymbol{\mathcal{K}}^{[2]}\mathbf{x} = \frac{1}{2}\left(\mathbf{1}_\tau' \,(\mathbf{Y} \odot \mathbf{Y})\, \mathbf{1}_\tau - tr(\mathbf{Y} \odot \mathbf{Y})\right),$$

where the matrices $\boldsymbol{\mathcal{K}}^{[1]}$ and $\boldsymbol{\mathcal{K}}^{[2]}$ are defined in the statement of Proposition 8 and $\mathbf{Y} = \mathbf{X} + \mathbf{X}'$.

Proof of Lemma 18: Denote by x_{rs} and y_{rs} the (r,s) elements of the matrices $\mathbf{X}$ and $\mathbf{Y}$, respectively. It then follows that $y_{rs} = x_{rs} + x_{sr}$. For each positive integer p, let $\ell_p \in \{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). The p-th element of the $\tau^2 \times 1$ vector $\boldsymbol{\mathcal{K}}^{[1]}\mathbf{x}$ is $\sum_{q=1}^{\tau^2} \mathbf{1}_{[k_p=\ell_p=k_q=\ell_q]} x_{k_q\ell_q} = \mathbf{1}_{[k_p=\ell_p]} x_{k_p\ell_p}$, and so $\mathbf{x}'\boldsymbol{\mathcal{K}}^{[1]}\mathbf{x} = \sum_{p=1}^{\tau^2} x_{k_p\ell_p}\mathbf{1}_{[k_p=\ell_p]} x_{k_p\ell_p} = \sum_{r=1}^{\tau} x_{rr}^2 = \text{tr}\,(\mathbf{X} \odot \mathbf{X})$. Moreover, note that the p-th element of the $\tau^2 \times 1$ vector $\boldsymbol{\mathcal{K}}^{[2]}\mathbf{x}$ is equal to $\sum_{q=1}^{\tau^2}\left(\mathbf{1}_{[k_p=k_q\neq\ell_p=\ell_q]} + \mathbf{1}_{[k_p=\ell_q\neq\ell_p=k_q]}\right) x_{k_q\ell_q} = \mathbf{1}_{[k_p\neq\ell_p]}\left(x_{k_p\ell_p} + x_{\ell_p k_p}\right)$. Hence,

$$\begin{aligned}\mathbf{x}'\boldsymbol{\mathcal{K}}^{[2]}\mathbf{x} &= \sum_{p=1}^{\tau^2} x_{k_p\ell_p}\mathbf{1}_{[k_p\neq\ell_p]}\left(x_{k_p\ell_p} + x_{\ell_p k_p}\right)\\ &= \frac{1}{2}\sum_{r=1}^{\tau}\sum_{s=1}^{\tau} \mathbf{1}_{[r\neq s]}\left[x_{rs}\,(x_{rs} + x_{sr}) + x_{sr}\,(x_{sr} + x_{rs})\right]\\ &= \frac{1}{2}\sum_{r=1}^{\tau}\sum_{s=1}^{\tau} \mathbf{1}_{[r\neq s]}\,(x_{rs} + x_{sr})^2 = \frac{1}{2}\sum_{r=1}^{\tau}\sum_{s=1}^{\tau} \mathbf{1}_{[r\neq s]} y_{rs}^2\\ &= \frac{1}{2}\left(\sum_{r=1}^{\tau}\sum_{s=1}^{\tau} y_{rs}^2 - \sum_{r=1}^{\tau} y_{rr}^2\right) = \frac{1}{2}\left(\mathbf{1}_\tau' \,(\mathbf{Y} \odot \mathbf{Y})\, \mathbf{1}_\tau - \text{tr}\,(\mathbf{Y} \odot \mathbf{Y})\right),\end{aligned}$$

completing the proof of the lemma. ■

The following three Lemmas 19, 20 and 21, used in the proof of Lemma 12 in the Appendix, provide important properties of the matrices

$$\boldsymbol{\mathcal{D}}^{[i]} = (\mathbf{H}_1 \otimes \mathbf{H}_1)'\, \boldsymbol{\mathcal{K}}^{[i]}\, (\mathbf{H}_1 \otimes \mathbf{H}_1)\,,\ i = 1,2,3, \tag{127}$$

where the matrices $\boldsymbol{\mathcal{K}}^{[i]}$, $i = 1,2,3$, are defined in Proposition 8.

Lemma 19 *Consider the matrices $\boldsymbol{\mathcal{D}}^{[i]}$, $i = 1,2,3$, defined in equation (127). Then,*

$$\begin{aligned}&tr(\boldsymbol{\mathcal{D}}^{[1]}) = tr\,(\mathbf{H}_1 \odot \mathbf{H}_1)\,,\\ &tr(\boldsymbol{\mathcal{D}}^{[2]}) = tr\,(\mathbf{H}_1)\,(tr\,(\mathbf{H}_1) + 1) - 2tr(\mathbf{H}_1 \odot \mathbf{H}_1),\\ &tr(\boldsymbol{\mathcal{D}}^{[3]}) = tr\,(\mathbf{H}_1) - tr\,(\mathbf{H}_1 \odot \mathbf{H}_1)\,.\end{aligned}$$

Proof of Lemma 19: Recall that the matrix $\mathbf{H}_1$ is symmetric and idempotent. In addition, consider the matrix $\boldsymbol{\mathcal{N}}$ defined in Lemma 15 and for each positive integer p, let $\ell_p \in \{1,\ldots,\tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). Then, note that the (p,q) element of the $\tau^2 \times \tau^2$ matrix $\boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'$ is $\sum_{r=1}^{\tau} \mathbf{1}_{[k_p=\ell_p=r]}\mathbf{1}_{[k_q=\ell_q=r]} = \mathbf{1}_{[k_p=\ell_p=k_q=\ell_q]}$ which implies that $\boldsymbol{\mathcal{K}}^{[1]} = \boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'$. Then, we obtain

$$\begin{aligned}\text{tr}(\boldsymbol{\mathcal{D}}^{[1]}) &= \text{tr}((\mathbf{H_1} \otimes \mathbf{H_1})'\, \boldsymbol{\mathcal{K}}^{[1]}\, (\mathbf{H_1} \otimes \mathbf{H_1})) = \text{tr}((\mathbf{H_1} \otimes \mathbf{H_1})'\, \boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'\, (\mathbf{H_1} \otimes \mathbf{H_1}))\\ &= \text{tr}\left(\boldsymbol{\mathcal{N}}'\,(\mathbf{H_1} \otimes \mathbf{H_1})\,(\mathbf{H_1} \otimes \mathbf{H_1})'\,\boldsymbol{\mathcal{N}}\right) = \text{tr}\left(\boldsymbol{\mathcal{N}}'\,(\mathbf{H_1} \otimes \mathbf{H_1})\,\boldsymbol{\mathcal{N}}\right) = \text{tr}\,(\mathbf{H_1} \odot \mathbf{H_1})\,,\end{aligned}$$

where the third equality follows from fact (F1), the fourth equality follows from fact (F5) and $\mathbf{H}_1$ being idempotent, and the fifth equality follows from Lemma 15. Note that $\mathrm{tr}(\boldsymbol{\mathcal{D}}^{[2]}) = \mathrm{tr}((\mathbf{H_1} \otimes \mathbf{H_1})' \boldsymbol{\mathcal{K}}^{[2]} (\mathbf{H_1} \otimes \mathbf{H_1})) = \mathrm{tr}\left(\boldsymbol{\mathcal{K}}^{[2]} (\mathbf{H_1} \otimes \mathbf{H_1})\right)$. Furthermore, let h_{rs} denote the (r, s) element of the matrix $\mathbf{H}_1$. The (p, q) element of the $\tau^2 \times \tau^2$ matrix $\mathbf{H_1} \otimes \mathbf{H_1}$ is $h_{k_p k_q} h_{\ell_p \ell_q}$ and, therefore the (q, q) element of the $\tau^2 \times \tau^2$ matrix $\boldsymbol{\mathcal{K}}^{[2]} (\mathbf{H_1} \otimes \mathbf{H_1})$ is

$$\sum\nolimits_{p=1}^{\tau^2} \left(\mathbf{1}_{[k_q = k_p \neq \ell_q = \ell_p]} + \mathbf{1}_{[k_q = \ell_p \neq \ell_q = k_p]}\right) h_{k_p k_q} h_{\ell_p \ell_q} = \mathbf{1}_{[k_q \neq \ell_q]} \left(h_{k_q k_q} h_{\ell_q \ell_q} + h_{\ell_q k_q} h_{k_q \ell_q}\right).$$

Hence,

$$\begin{aligned}
\mathrm{tr}(\boldsymbol{\mathcal{K}}^{[2]} (\mathbf{H_1} \otimes \mathbf{H_1})) &= \sum\nolimits_{k_q=1}^{\tau} \sum\nolimits_{\ell_q=1}^{\tau} \mathbf{1}_{[k_q \neq \ell_q]} \left(h_{k_q k_q} h_{\ell_q \ell_q} + h_{\ell_q k_q} h_{k_q \ell_q}\right) \\
&= \sum\nolimits_{r=1}^{\tau} \sum\nolimits_{s=1}^{\tau} (h_{rr} h_{ss} + h_{sr} h_{rs}) - \sum\nolimits_{r=1}^{\tau} (h_{rr} h_{rr} + h_{rr} h_{rr}) \\
&= \sum\nolimits_{r=1}^{\tau} \sum\nolimits_{s=1}^{\tau} h_{rr} h_{ss} + \sum\nolimits_{r=1}^{\tau} \sum\nolimits_{s=1}^{\tau} h_{rs}^2 - 2 \sum\nolimits_{r=1}^{\tau} h_{rr}^2 \\
&= \mathrm{tr}\left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) + \mathbf{1}_\tau' \left(\mathbf{H}_1 \odot \mathbf{H}_1\right) \mathbf{1}_\tau - 2\mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right) \\
&= \mathrm{tr}\left(\mathbf{H}_1\right)^2 + \mathrm{tr}\left(\mathbf{H}_1\right) - 2\mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right),
\end{aligned}$$

where the last equality follows from Lemma 17. Therefore, $\mathrm{tr}(\boldsymbol{\mathcal{D}}^{[2]}) = \mathrm{tr}\left(\mathbf{H}_1\right)\left(\mathrm{tr}\left(\mathbf{H}_1\right) + 1\right) - 2\mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)$. Note that the (p, q) element of the matrix $\boldsymbol{\mathcal{K}}^{[1]} + \boldsymbol{\mathcal{K}}^{[3]}$ is $\mathbf{1}_{[k_p = \ell_p = k_q = \ell_q]} + \mathbf{1}_{[k_p = \ell_p \neq k_q = \ell_q]} = \mathbf{1}_{[k_p = \ell_p]} \mathbf{1}_{[k_q = \ell_q]}$ which is the (p, q) element of $\mathbf{i}_\tau' \mathbf{i}_\tau$. Hence, $\boldsymbol{\mathcal{K}}^{[1]} + \boldsymbol{\mathcal{K}}^{[3]} = \mathbf{i}_\tau' \mathbf{i}_\tau$ and so $\boldsymbol{\mathcal{K}}^{[3]} = \mathbf{i}_\tau \mathbf{i}_\tau' - \boldsymbol{\mathcal{K}}^{[1]}$. It then follows that

$$\begin{aligned}
\mathrm{tr}(\boldsymbol{\mathcal{D}}^{[3]}) &= \mathrm{tr}\left(\left(\mathbf{H}_1' \otimes \mathbf{H}_1'\right) \mathbf{i}_\tau \mathbf{i}_\tau' \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right)\right) - \mathrm{tr}(\boldsymbol{\mathcal{D}}^{[1]}) \\
&= \mathbf{i}_\tau' \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \mathbf{i}_\tau - \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right) = \mathrm{tr}\left(\mathbf{H}_1\right) - \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right),
\end{aligned}$$

where the last equality follows from Lemma 17. ■

Lemma 20 *Consider the matrices* $\boldsymbol{\mathcal{D}}^{[i]}$, $i = 1, 2, 3$, *defined in equation (127). Then,*

$$\begin{aligned}
\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[1]} \mathbf{i}_\tau &= tr\left(\mathbf{H}_1 \odot \mathbf{H}_1\right), \\
\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[2]} \mathbf{i}_\tau &= 2\left(tr\left(\mathbf{H}_1\right) - tr\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right), \\
\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[3]} \mathbf{i}_\tau &= tr\left(\mathbf{H}_1\right)^2 - tr\left(\mathbf{H}_1 \odot \mathbf{H}_1\right).
\end{aligned}$$

<u>Proof of Lemma 20</u>: Define $\mathbf{h}_\mathsf{K} = \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \mathbf{i}_\tau$ and note that $\mathbf{h}_\mathsf{K} = \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \mathrm{vec}(\mathbf{I}_\tau) = \mathrm{vec}(\mathbf{H}_1 \mathbf{I}_\tau \mathbf{H}_1) = \mathrm{vec}(\mathbf{H}_1)$ due to fact (F4) and $\mathbf{H}_1$ being symmetric and idempotent. It follows that $\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[1]} \mathbf{i}_\tau = \mathbf{h}_\mathsf{K}' \boldsymbol{\mathcal{K}}^{[1]} \mathbf{h}_\mathsf{K}$ and $\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[2]} \mathbf{i}_\tau = \mathbf{h}_\mathsf{K}' \boldsymbol{\mathcal{K}}^{[2]} \mathbf{h}_\mathsf{K}$. Lemma 18 then yields $\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[1]} \mathbf{i}_\tau = \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)$ and

$$\begin{aligned}
\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[2]} \mathbf{i}_\tau &= \frac{1}{2} \left(\mathbf{1}_\tau' \left((2\mathbf{H}_1) \odot (2\mathbf{H}_1)\right) \mathbf{1}_\tau - \mathrm{tr}\left((2\mathbf{H}_1) \odot (2\mathbf{H}_1)\right)\right) \\
&= 2\left(\mathbf{1}_\tau' \left(\mathbf{H}_1 \odot \mathbf{H}_1\right) \mathbf{1}_\tau - \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right) \\
&= 2\left(\mathrm{tr}\left(\mathbf{H}_1\right) - \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right),
\end{aligned}$$

where we use the symmetry of $\mathbf{H}_1$ and the last equality is obtained by invoking Lemma 17. It follows from the proof of Lemma 19 that $\boldsymbol{\mathcal{K}}^{[3]} = \mathbf{i}_\tau \mathbf{i}_\tau' - \boldsymbol{\mathcal{K}}^{[1]}$ which, in turn, yields

$$\mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[3]} \mathbf{i}_\tau = \left[\mathbf{i}_\tau' \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \mathbf{i}_\tau\right]^2 - \mathbf{i}_\tau' \boldsymbol{\mathcal{D}}^{[1]} \mathbf{i}_\tau = \left[\mathrm{tr}\left(\mathbf{H}_1\right)\right]^2 - \mathrm{tr}\left(\mathbf{H}_1 \odot \mathbf{H}_1\right),$$

where the last equality follows from Lemma 17. ■

Lemma 21 *Consider the matrices* $\boldsymbol{\mathcal{D}}^{[i]}$, $i = 1, 2, 3$, *defined in equation (127). Then,*

$$\begin{aligned}
tr(\boldsymbol{\mathcal{N}}' \boldsymbol{\mathcal{D}}^{[1]} \boldsymbol{\mathcal{N}}) &= tr\left(\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right), \\
tr(\boldsymbol{\mathcal{N}}' \boldsymbol{\mathcal{D}}^{[2]} \boldsymbol{\mathcal{N}}) &= 2\left(tr\left(\mathbf{H}_1 \odot \mathbf{H}_1\right) - tr\left(\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right)\right), \\
tr(\boldsymbol{\mathcal{N}}' \boldsymbol{\mathcal{D}}^{[3]} \boldsymbol{\mathcal{N}}) &= tr(\mathbf{H}_1 \odot \mathbf{H}_1) - tr\left(\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right).
\end{aligned}$$

<u>Proof of Lemma 21</u>: Using the fact $\boldsymbol{\mathcal{K}}^{[1]} = \boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'$ (see the proof of Lemma 19), we obtain

$$\mathrm{tr}(\boldsymbol{\mathcal{N}}' \boldsymbol{\mathcal{D}}^{[1]} \boldsymbol{\mathcal{N}}) = \mathrm{tr}\left(\boldsymbol{\mathcal{N}}' \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}' \left(\mathbf{H}_1 \otimes \mathbf{H}_1\right) \boldsymbol{\mathcal{N}}\right) = \mathrm{tr}\left(\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\left(\mathbf{H}_1 \odot \mathbf{H}_1\right)\right),$$

since, according to Lemma 15, $\boldsymbol{\mathcal{N}}'(\mathbf{H}_1 \otimes \mathbf{H}_1)\boldsymbol{\mathcal{N}} = \mathbf{H}_1 \odot \mathbf{H}_1$.

Note that $\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[2]}\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{G}}'\boldsymbol{\mathcal{K}}^{[2]}\boldsymbol{\mathcal{G}}$, where $\boldsymbol{\mathcal{G}} = (\mathbf{H}_1 \otimes \mathbf{H}_1)\boldsymbol{\mathcal{N}}$. Denote the $\boldsymbol{g}_r$ the r-th column of $\tau^2 \times \tau$ matrix $\boldsymbol{\mathcal{G}}$, for $r = 1, \ldots, \tau$. For each positive integer p, let $\ell_p \in \{1, \ldots, \tau\}$ and k_p be the unique positive integers such that $p = (k_p - 1)\tau + \ell_p$ (see (37)). Furthermore, let h_{rs} denote the (r, s) element of the matrix $\mathbf{H}_1$. Then, the (p, q) element of $\mathbf{H}_1 \otimes \mathbf{H}_1$ is $h_{k_p k_q} h_{\ell_p \ell_q}$. Hence, since the (q, r) element of $\boldsymbol{\mathcal{N}}$ is $\mathbf{1}_{[k_q = \ell_q = r]}$, the p-th element of the vector $\boldsymbol{g}_r$ is $\sum_{q=1}^{\tau^2} h_{k_p k_q} h_{\ell_p \ell_q} \mathbf{1}_{[k_q = \ell_q = r]} = h_{k_p r} h_{\ell_p r}$. Denoting by $\mathbf{h}_r$ the r-th column of $\mathbf{H}_1$, we have $\boldsymbol{g}_r = \mathbf{h}_r \otimes \mathbf{h}_r = \mathrm{vec}(\mathbf{h}_r \mathbf{h}_r')$. Therefore, the (r, r) element of the matrix $\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[2]}\boldsymbol{\mathcal{N}}$ is equal to $\boldsymbol{g}_r'\boldsymbol{\mathcal{K}}^{[2]}\boldsymbol{g}_r$. Lemma 18 then implies

$$\begin{aligned}
\boldsymbol{g}_r'\boldsymbol{\mathcal{K}}^{[2]}\boldsymbol{g}_r &= \frac{1}{2}\left(\mathbf{1}_\tau'\left((2\mathbf{h}_r\mathbf{h}_r') \odot (2\mathbf{h}_r\mathbf{h}_r')\right)\mathbf{1}_\tau - \mathrm{tr}\left((2\mathbf{h}_r\mathbf{h}_r') \odot (2\mathbf{h}_r\mathbf{h}_r')\right)\right) \\
&= 2\left(\mathbf{1}_\tau'\left((\mathbf{h}_r\mathbf{h}_r') \odot (\mathbf{h}_r\mathbf{h}_r')\right)\mathbf{1}_\tau - \mathrm{tr}\left((\mathbf{h}_r\mathbf{h}_r') \odot (\mathbf{h}_r\mathbf{h}_r')\right)\right).
\end{aligned}$$

Note that $\mathbf{H}_1 \odot \mathbf{H}_1 = \begin{bmatrix} \mathbf{h}_1 \odot \mathbf{h}_1 & \cdots & \mathbf{h}_\tau \odot \mathbf{h}_\tau \end{bmatrix}$ and so, due to the symmetry of $\mathbf{H}_1$, we have

$$\begin{aligned}
(\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1) &= (\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1)' \\
&= \sum\nolimits_{r=1}^{\tau}(\mathbf{h}_r \odot \mathbf{h}_r)(\mathbf{h}_r \odot \mathbf{h}_r)' \\
&= \sum\nolimits_{r=1}^{\tau}(\mathbf{h}_r\mathbf{h}_r') \odot (\mathbf{h}_r\mathbf{h}_r'),
\end{aligned}$$

where the last equality follows from fact (F8). Collecting the above results, we obtain

$$\begin{aligned}
\mathrm{tr}(\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[2]}\boldsymbol{\mathcal{N}}) &= \sum\nolimits_{r=1}^{\tau} \boldsymbol{g}_r'\boldsymbol{\mathcal{K}}^{[2]}\boldsymbol{g}_r \\
&= 2\left(\mathbf{1}_\tau'(\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1)\mathbf{1}_\tau - \mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1))\right) \\
&= 2\left(\mathrm{tr}(\mathbf{H}_1 \odot \mathbf{H}_1) - \mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1))\right),
\end{aligned}$$

where the last equality follows from Lemma 17.

Since $\boldsymbol{\mathcal{K}}^{[3]} = \mathbf{i}_\tau\mathbf{i}_\tau' - \boldsymbol{\mathcal{K}}^{[1]}$ (see the proof of Lemma 19) and $\boldsymbol{\mathcal{K}}^{[1]} = \boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'$, we obtain

$$\begin{aligned}
\mathrm{tr}(\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[3]}\boldsymbol{\mathcal{N}}) &= \mathrm{tr}\left(\boldsymbol{\mathcal{N}}'(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathbf{i}_\tau\mathbf{i}_\tau'(\mathbf{H}_1 \otimes \mathbf{H}_1)\boldsymbol{\mathcal{N}}\right) - \mathrm{tr}(\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[1]}\boldsymbol{\mathcal{N}}) \\
&= \mathrm{tr}\left(\mathbf{i}_\tau'(\mathbf{H}_1 \otimes \mathbf{H}_1)\boldsymbol{\mathcal{N}}\boldsymbol{\mathcal{N}}'(\mathbf{H}_1 \otimes \mathbf{H}_1)\mathbf{i}_\tau\right) - \mathrm{tr}(\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[1]}\boldsymbol{\mathcal{N}}) \\
&= \mathbf{i}_\tau'\boldsymbol{\mathcal{D}}^{[1]}\mathbf{i}_\tau - \mathrm{tr}(\boldsymbol{\mathcal{N}}'\boldsymbol{\mathcal{D}}^{[1]}\boldsymbol{\mathcal{N}}) \\
&= \mathrm{tr}(\mathbf{H}_1 \odot \mathbf{H}_1) - \mathrm{tr}((\mathbf{H}_1 \odot \mathbf{H}_1)(\mathbf{H}_1 \odot \mathbf{H}_1)),
\end{aligned}$$

where the last equality follows from Lemma 20. ■

Table 8: **Empirical bias in the estimation of γ with Student-t distributed shocks: the role of the EIV correction.** This table presents simulation results on the bias (in annualized basis points) in the estimation of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \cdots \ \lambda_K\ ]'$ is the vector of ex-post prices of risk, and K is the number of factors. We estimate $\boldsymbol{\gamma}$ with and without the EIV correction. The shocks u_{it} are assumed to follow a Student-t distribution with 6 degrees of freedom. The number of individual stocks, N, is equal to 2,000 and the number of clusters, M_N, is set equal to 50. The pairwise correlation of shocks, assumed to be constant within each cluster, is set equal to 0.10. The simulation is calibrated to two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. The results are based on 10,000 Monte Carlo repetitions.

Estimation Period	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-00	01-05
Testing Period	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-00	01-05	06-10
CAPM												
Without EIV correction												
α	587.5	332.4	397.5	-25.2	-65.6	221.3	138.7	177.4	274.5	573.1	21.1	72.2
λ_1	-670.7	-379.5	-456.5	29.2	75.8	-254.5	-158.4	-200.6	-313.9	-655.3	-25.2	-83.3
With EIV correction												
α	-4.2	-4.5	-3.9	1.0	1.8	-0.5	-1.1	-1.3	-3.8	-7.5	-1.0	-0.5
λ_1	6.6	6.3	2.9	-0.8	-1.2	-0.7	1.6	4.0	4.6	9.2	0.1	-0.2
Fama-French model												
Without EIV correction												
α	570.5	424.0	697.3	148.0	102.0	243.6	436.5	127.3	580.5	644.7	541.5	114.1
λ_1	-836.1	-462.9	-468.1	193.2	0.6	-28.0	-157.4	-377.2	-348.3	-743.1	-143.1	-28.7
λ_2	484.8	-173.6	-56.2	-491.3	92.7	-467.7	-46.6	515.4	-168.4	211.1	-435.5	-169.0
λ_3	-13.8	177.8	-633.0	-316.3	-358.4	-70.8	-698.4	18.6	-493.5	-133.9	-596.1	-50.9
With EIV correction												
α	-21.5	-41.3	-27.6	0.3	-0.9	-1.2	-8.4	-1.0	-16.1	-20.4	-11.9	-1.4
λ_1	55.3	51.7	12.7	-26.9	0.8	-1.8	2.9	12.7	6.2	23.0	5.8	-1.1
λ_2	-65.9	41.5	0.9	39.0	-3.7	5.7	-0.6	-22.8	6.6	-6.8	5.8	4.1
λ_3	-7.6	-49.9	38.6	23.4	4.3	1.0	16.9	-3.1	22.5	5.6	11.4	0.1

Table 9: **Empirical rejection frequencies with Student-t distributed shocks.** This table presents simulation results on the finite-sample performance of the EIV-corrected estimator $\widetilde{\boldsymbol{\gamma}}$ of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \ \cdots \ \ \lambda_K\ ]'$ is the vector of ex-post prices of risk, and K is the number of factors. Reported are empirical p-values (in percentages) for the χ^2 statistic associated with the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ and the t statistics associated with the simple hypotheses $\alpha = 0$ as well as $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. The shocks u_{it} are assumed to follow a Student-t distribution with 6 degrees of freedom. We consider two possible values for the number of clusters, M_N, namely 50 and 100. The pairwise correlation of shocks, assumed to be constant within each cluster, is denoted by ρ and takes three values, namely 0, 0.10, and 0.20. Three nominal levels of significance are considered: 1%, 5%, and 10%. The simulation is calibrated to two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. The results are based on 10,000 Monte Carlo repetitions.

ρ	0						0.10						0.20					
M_N	50			100			50			100			50			100		
Nominal Size	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10
								CAPM										
Joint χ^2	0.9	4.9	9.9	1.1	5.1	9.8	1.1	5.3	10.4	1.0	5.1	10.4	1.1	5.1	10.0	1.1	5.3	10.2
$t(\alpha)$	1.0	5.0	10.1	1.1	5.1	10.1	1.0	5.4	10.2	1.1	5.1	10.3	1.2	5.1	9.9	1.2	5.3	10.4
$t(\lambda_1)$	1.1	5.2	10.2	1.1	4.9	9.5	1.1	5.2	10.5	1.1	5.2	9.7	1.1	5.0	10.1	1.0	5.0	10.4
								Fama-French model										
Joint χ^2	1.0	4.8	9.9	1.1	4.9	9.9	1.2	5.2	10.1	1.1	5.0	10.2	1.3	5.3	10.4	1.1	5.2	10.2
$t(\alpha)$	1.0	4.7	9.7	1.0	5.2	10.2	1.2	5.1	10.3	0.9	4.8	9.7	1.1	5.0	10.5	1.1	5.1	9.8
$t(\lambda_1)$	0.9	4.7	9.5	1.0	4.7	9.5	1.0	5.1	10.3	1.0	4.8	10.1	1.0	5.1	10.1	1.0	4.7	10.3
$t(\lambda_2)$	0.9	4.9	9.9	1.0	5.1	10.1	1.0	4.9	9.9	1.0	5.0	9.8	1.0	4.9	9.8	1.0	4.8	9.6
$t(\lambda_3)$	1.1	5.2	10.6	1.1	5.4	10.1	1.2	5.2	9.9	1.0	5.4	10.0	1.3	5.7	10.5	1.0	4.8	9.5

Table 10: **Test statistics and p-values for six-year testing periods from 1951 to 2010 using 50 clusters based on size.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 10 non-overlapping six-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
					CAPM					
Joint χ^2	139.52	273.71	146.45	14.30	77.21	337.33	18.96	94.93	1009.88	159.75
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	5.82	16.29	1.13	-0.37	-0.37	17.06	3.58	8.86	3.59	7.04
	[0.000]	[0.000]	[0.257]	[0.710]	[0.715]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\lambda_1)$	-8.20	-14.35	1.51	1.35	3.03	-12.49	-2.49	-9.74	8.91	0.79
	[0.000]	[0.000]	[0.132]	[0.178]	[0.002]	[0.000]	[0.013]	[0.000]	[0.000]	[0.428]
					Fama-French model					
Joint χ^2	93.22	286.65	52.31	96.08	21.11	157.10	39.01	92.66	310.90	142.55
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	6.88	14.37	3.26	-0.81	-0.31	11.38	0.19	6.50	2.23	0.35
	[0.000]	[0.000]	[0.001]	[0.420]	[0.753]	[0.000]	[0.845]	[0.000]	[0.026]	[0.724]
$t(\lambda_1)$	-8.59	-13.81	-3.17	0.84	0.51	-10.69	-0.65	-7.27	2.37	2.04
	[0.000]	[0.000]	[0.002]	[0.403]	[0.609]	[0.000]	[0.518]	[0.000]	[0.018]	[0.042]
$t(\lambda_2)$	2.07	3.16	0.05	4.81	-2.95	0.93	1.35	7.11	2.48	1.91
	[0.038]	[0.002]	[0.961]	[0.000]	[0.003]	[0.353]	[0.176]	[0.000]	[0.013]	[0.056]
$t(\lambda_3)$	1.14	-3.51	-0.68	-1.41	-0.92	-0.77	0.34	-1.72	-6.38	1.86
	[0.253]	[0.000]	[0.497]	[0.160]	[0.357]	[0.439]	[0.732]	[0.086]	[0.000]	[0.064]

Table 11: **Test statistics and p-values for six-year testing periods from 1951 to 2010 using 30 clusters based on size.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 10 non-overlapping six-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
					CAPM					
Joint χ^2	107.77	247.35	116.16	10.86	53.28	265.13	15.18	80.02	722.40	110.07
	[0.000]	[0.000]	[0.000]	[0.004]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	5.68	15.65	1.06	-0.35	-0.35	15.52	3.31	7.97	3.28	5.93
	[0.000]	[0.000]	[0.291]	[0.726]	[0.730]	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]
$t(\lambda_1)$	-7.80	-14.00	1.40	1.25	2.81	-11.42	-2.30	-8.95	8.13	0.71
	[0.000]	[0.000]	[0.160]	[0.211]	[0.005]	[0.000]	[0.021]	[0.000]	[0.000]	[0.481]
					Fama-French model					
Joint χ^2	85.78	268.89	54.51	89.73	20.24	133.28	34.92	84.66	258.81	111.01
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	6.73	13.91	3.12	-0.78	-0.30	10.72	0.19	6.29	2.15	0.33
	[0.000]	[0.000]	[0.002]	[0.434]	[0.764]	[0.000]	[0.852]	[0.000]	[0.031]	[0.740]
$t(\lambda_1)$	-8.31	-13.49	-3.08	0.81	0.49	-10.35	-0.62	-7.14	2.25	1.85
	[0.000]	[0.000]	[0.002]	[0.416]	[0.625]	[0.000]	[0.538]	[0.000]	[0.025]	[0.064]
$t(\lambda_2)$	2.04	3.13	0.05	4.66	-2.84	0.90	1.30	6.89	2.48	1.90
	[0.042]	[0.002]	[0.961]	[0.000]	[0.005]	[0.370]	[0.192]	[0.000]	[0.013]	[0.057]
$t(\lambda_3)$	1.13	-3.38	-0.66	-1.37	-0.90	-0.75	0.33	-1.72	-6.02	1.77
	[0.260]	[0.001]	[0.512]	[0.172]	[0.369]	[0.453]	[0.741]	[0.086]	[0.000]	[0.077]

Table 12: **Test statistics and p-values for three-year testing periods from 1951 to 2010 using 50 clusters based on size.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 20 non-overlapping three-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
					CAPM					
Joint χ^2	86.58	27.40	203.73	87.58	35.09	181.43	14.73	37.75	155.15	9.62
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.008]
$t(\alpha)$	3.63	3.46	13.53	9.32	-0.39	2.95	0.62	-0.79	3.17	2.32
	[0.000]	[0.001]	[0.000]	[0.000]	[0.699]	[0.003]	[0.536]	[0.432]	[0.002]	[0.020]
$t(\lambda_1)$	-5.72	-4.51	-11.16	-9.08	1.64	0.38	-1.65	2.85	1.03	-1.23
	[0.000]	[0.000]	[0.000]	[0.000]	[0.100]	[0.704]	[0.100]	[0.004]	[0.302]	[0.219]
					Fama-French model					
Joint χ^2	17.76	41.74	164.22	120.74	8.33	108.00	5.45	184.64	25.38	65.52
	[0.000]	[0.000]	[0.000]	[0.000]	[0.016]	[0.000]	[0.066]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	2.45	6.32	11.08	9.36	-0.34	5.60	-2.16	-0.70	2.36	4.45
	[0.014]	[0.000]	[0.000]	[0.000]	[0.737]	[0.000]	[0.031]	[0.482]	[0.018]	[0.000]
$t(\lambda_1)$	-3.33	-6.21	-10.57	-9.00	-0.12	-4.73	2.27	-0.38	-1.55	-3.84
	[0.001]	[0.000]	[0.000]	[0.000]	[0.903]	[0.000]	[0.023]	[0.701]	[0.122]	[0.000]
$t(\lambda_2)$	0.84	-0.39	0.85	3.86	0.96	-2.74	-1.41	11.18	-0.41	-3.27
	[0.402]	[0.694]	[0.394]	[0.000]	[0.339]	[0.006]	[0.158]	[0.000]	[0.680]	[0.001]
$t(\lambda_3)$	-0.80	0.43	-2.59	-3.07	-1.34	0.25	0.52	-0.17	-4.59	-1.10
	[0.425]	[0.669]	[0.010]	[0.002]	[0.180]	[0.800]	[0.604]	[0.868]	[0.000]	[0.273]

Continued on next page

Table 12 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
					CAPM					
Joint χ^2	248.02	205.66	30.46	25.84	75.87	109.39	512.82	271.10	4.61	352.51
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.100]	[0.000]
$t(\alpha)$	11.30	13.75	5.12	-2.15	8.48	6.83	-5.68	11.15	2.15	2.70
	[0.000]	[0.000]	[0.000]	[0.032]	[0.000]	[0.000]	[0.000]	[0.000]	[0.032]	[0.007]
$t(\lambda_1)$	-5.48	-14.28	-5.52	3.55	-6.47	-9.40	13.94	-2.50	-1.72	4.68
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.013]	[0.085]	[0.000]
					Fama-French model					
Joint χ^2	62.94	110.58	26.92	40.49	67.41	84.52	254.44	50.65	10.31	186.63
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.006]	[0.000]
$t(\alpha)$	7.12	4.79	0.37	-0.90	6.79	4.54	-1.20	2.08	-1.09	1.51
	[0.000]	[0.000]	[0.709]	[0.370]	[0.000]	[0.000]	[0.232]	[0.037]	[0.275]	[0.130]
$t(\lambda_1)$	-6.51	-6.06	-1.06	1.91	-5.38	-6.70	4.38	-1.45	-1.29	1.56
	[0.000]	[0.000]	[0.287]	[0.057]	[0.000]	[0.000]	[0.000]	[0.147]	[0.197]	[0.119]
$t(\lambda_2)$	-1.04	2.42	0.84	-0.24	2.53	7.41	4.34	0.62	2.93	-1.33
	[0.297]	[0.015]	[0.402]	[0.811]	[0.011]	[0.000]	[0.000]	[0.535]	[0.003]	[0.183]
$t(\lambda_3)$	-0.31	1.28	0.39	-2.49	-2.16	-0.81	-10.28	1.22	0.62	1.51
	[0.755]	[0.200]	[0.697]	[0.013]	[0.031]	[0.415]	[0.000]	[0.223]	[0.535]	[0.132]

Table 13: **Test statistics and p-values for three-year testing periods from 1951 to 2010 using 30 clusters based on size.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \dots, K$. We consider 20 non-overlapping three-year testing periods from 1951 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
					CAPM					
Joint χ^2	68.71	21.00	176.96	81.93	27.12	154.24	11.04	27.07	107.65	7.45
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.004]	[0.000]	[0.000]	[0.024]
$t(\alpha)$	3.52	3.21	12.88	8.99	-0.37	2.83	0.58	-0.73	3.01	2.18
	[0.000]	[0.001]	[0.000]	[0.000]	[0.708]	[0.005]	[0.561]	[0.466]	[0.003]	[0.029]
$t(\lambda_1)$	-5.43	-4.11	-10.73	-8.74	1.58	0.36	-1.52	2.60	0.96	-1.13
	[0.000]	[0.000]	[0.000]	[0.000]	[0.114]	[0.720]	[0.129]	[0.009]	[0.335]	[0.259]
					Fama-French model					
Joint χ^2	17.22	39.02	153.48	110.41	7.82	104.54	5.22	178.94	23.73	57.71
	[0.000]	[0.000]	[0.000]	[0.000]	[0.020]	[0.000]	[0.073]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	2.38	6.11	10.83	9.07	-0.32	5.48	-2.10	-0.67	2.32	4.08
	[0.017]	[0.000]	[0.000]	[0.000]	[0.748]	[0.000]	[0.036]	[0.504]	[0.020]	[0.000]
$t(\lambda_1)$	-3.25	-5.99	-10.41	-8.68	-0.12	-4.65	2.21	-0.37	-1.53	-3.57
	[0.001]	[0.000]	[0.000]	[0.000]	[0.907]	[0.000]	[0.027]	[0.713]	[0.127]	[0.000]
$t(\lambda_2)$	0.82	-0.39	0.83	3.75	0.97	-2.69	-1.33	10.92	-0.38	-3.14
	[0.413]	[0.699]	[0.405]	[0.000]	[0.334]	[0.007]	[0.183]	[0.000]	[0.704]	[0.002]
$t(\lambda_3)$	-0.81	0.40	-2.55	-2.97	-1.31	0.24	0.49	-0.16	-4.45	-1.03
	[0.417]	[0.689]	[0.011]	[0.003]	[0.189]	[0.808]	[0.624]	[0.871]	[0.000]	[0.305]

Continued on next page

Table 13 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
					CAPM					
Joint χ^2	176.25	166.48	26.70	19.47	58.62	89.91	384.73	178.86	3.30	258.29
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.192]	[0.000]
$t(\alpha)$	10.07	12.16	4.79	-1.96	7.59	6.24	-5.14	9.40	1.80	2.42
	[0.000]	[0.000]	[0.000]	[0.050]	[0.000]	[0.000]	[0.000]	[0.000]	[0.071]	[0.016]
$t(\lambda_1)$	-4.90	-12.87	-5.17	3.25	-5.98	-8.67	12.75	-2.20	-1.52	4.23
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.028]	[0.128]	[0.000]
					Fama-French model					
Joint χ^2	51.93	92.22	23.80	35.65	55.79	81.83	240.95	36.97	9.82	148.32
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.007]	[0.000]
$t(\alpha)$	6.50	4.67	0.36	-0.86	6.48	4.47	-1.05	1.96	-1.03	1.41
	[0.000]	[0.000]	[0.719]	[0.389]	[0.000]	[0.000]	[0.296]	[0.050]	[0.305]	[0.157]
$t(\lambda_1)$	-6.01	-5.94	-1.03	1.85	-5.29	-6.63	3.87	-1.33	-1.18	1.49
	[0.000]	[0.000]	[0.304]	[0.064]	[0.000]	[0.000]	[0.000]	[0.183]	[0.238]	[0.137]
$t(\lambda_2)$	-1.00	2.33	0.81	-0.23	2.47	7.31	4.22	0.55	2.86	-1.30
	[0.318]	[0.020]	[0.419]	[0.821]	[0.013]	[0.000]	[0.000]	[0.581]	[0.004]	[0.194]
$t(\lambda_3)$	-0.29	1.24	0.38	-2.40	-2.11	-0.79	-9.73	1.17	0.58	1.43
	[0.771]	[0.215]	[0.707]	[0.017]	[0.035]	[0.428]	[0.000]	[0.242]	[0.559]	[0.153]

Table 14: **Estimation results for six-year testing periods from 1969 to 2010 using stocks with book-to-market data.** This table presents the estimates of $\boldsymbol{\gamma} = [\ \alpha\ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1\ \ \cdots\ \ \lambda_K\ ]'$ is the vector of ex-post factor risk premia, for two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. We report the estimates with and without the EIV correction as well as the realized factor average over the testing period $\bar{\mathbf{f}}_2$. The linear factor model implies that $\alpha = 0$ and $\boldsymbol{\lambda} = \bar{\mathbf{f}}_2$. We consider seven non-overlapping six-year testing periods from 1969 to 2010. For each testing period, the beta estimation period consists of the preceding five years. All numbers are reported in annualized percentages.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
CAPM										
Without EIV correction										
α	–	–	–	-3.82	6.48	18.31	5.86	9.45	9.40	7.19
λ_1	–	–	–	-5.38	12.32	-7.12	3.35	3.31	7.00	2.41
With EIV correction										
α	–	–	–	0.24	-0.55	20.82	4.26	7.92	5.41	6.82
λ_1	–	–	–	-8.82	18.71	-9.40	5.05	4.93	11.69	2.91
Realized factor average										
λ_1	–	–	–	-10.19	12.69	5.80	8.29	14.80	0.62	2.62
Fama-French model										
Without EIV correction										
α	–	–	–	-5.06	5.54	16.89	5.71	9.16	7.98	5.17
λ_1	–	–	–	-3.88	9.04	-8.30	3.05	2.70	4.55	2.92
λ_2	–	–	–	-3.43	9.26	0.19	0.89	1.92	7.26	2.96
λ_3	–	–	–	1.88	-0.29	5.35	-1.20	-0.62	-2.63	0.79
With EIV correction										
α	–	–	–	-2.47	-2.12	18.41	-0.37	7.89	2.83	-0.32
λ_1	–	–	–	-6.72	15.01	-12.28	7.45	3.64	5.31	5.20
λ_2	–	–	–	-3.68	10.89	1.85	3.92	2.69	14.59	8.14
λ_3	–	–	–	3.59	2.16	10.53	5.56	-0.61	-1.39	5.15
Realized factor average										
λ_1	–	–	–	-10.19	12.69	5.80	8.29	14.80	0.62	2.62
λ_2	–	–	–	-9.66	14.11	1.57	-1.64	-4.73	10.66	3.15
λ_3	–	–	–	6.29	2.03	12.46	0.44	3.16	7.63	1.39

Table 15: **Test statistics and p-values for six-year testing periods from 1969 to 2010 using 50 clusters based on book-to-market.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider seven non-overlapping six-year testing periods from 1969 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
					CAPM					
Joint χ^2	–	–	–	12.96	74.41	261.27	17.27	65.33	855.76	147.66
	–	–	–	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	–	–	–	0.13	-0.23	14.82	3.19	6.97	3.62	6.71
	–	–	–	[0.898]	[0.815]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
$t(\lambda_1)$	–	–	–	0.78	2.65	-11.38	-2.19	-8.04	5.79	0.20
	–	–	–	[0.434]	[0.008]	[0.000]	[0.029]	[0.000]	[0.000]	[0.839]
					Fama-French model					
Joint χ^2	–	–	–	76.25	19.29	135.62	31.54	74.01	239.24	117.53
	–	–	–	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	–	–	–	-0.75	-0.60	10.96	-0.06	5.33	1.18	-0.15
	–	–	–	[0.456]	[0.550]	[0.000]	[0.955]	[0.000]	[0.237]	[0.881]
$t(\lambda_1)$	–	–	–	0.97	0.70	-10.42	-0.18	-6.31	1.84	1.85
	–	–	–	[0.330]	[0.482]	[0.000]	[0.855]	[0.000]	[0.066]	[0.064]
$t(\lambda_2)$	–	–	–	4.27	-3.39	0.26	1.37	6.82	2.22	1.83
	–	–	–	[0.000]	[0.001]	[0.796]	[0.172]	[0.000]	[0.026]	[0.068]
$t(\lambda_3)$	–	–	–	-1.34	0.07	-0.99	0.52	-2.38	-6.17	1.87
	–	–	–	[0.179]	[0.942]	[0.321]	[0.605]	[0.017]	[0.000]	[0.061]

Table 16: **Test statistics and p-values for six-year testing periods from 1969 to 2010 using 30 clusters based on book-to-market.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider seven non-overlapping six-year testing periods from 1969 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	52-56	58-62	64-68	70-74	76-80	82-86	88-92	94-98	00-04
Testing Period	51-56	57-62	63-68	69-74	75-80	81-86	87-92	93-98	99-04	05-10
					CAPM					
Joint χ^2	–	–	–	10.52	53.63	190.99	14.65	52.38	628.56	105.09
	–	–	–	[0.005]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	–	–	–	0.12	-0.22	13.02	3.06	6.29	3.31	5.67
	–	–	–	[0.905]	[0.825]	[0.000]	[0.002]	[0.000]	[0.001]	[0.000]
$t(\lambda_1)$	–	–	–	0.73	2.48	-10.21	-2.08	-7.22	5.27	0.17
	–	–	–	[0.464]	[0.013]	[0.000]	[0.037]	[0.000]	[0.000]	[0.862]
					Fama-French model					
Joint χ^2	–	–	–	72.74	18.21	109.63	28.91	72.41	206.62	94.89
	–	–	–	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$t(\alpha)$	–	–	–	-0.72	-0.58	10.00	-0.05	5.06	1.17	-0.15
	–	–	–	[0.469]	[0.564]	[0.000]	[0.957]	[0.000]	[0.241]	[0.883]
$t(\lambda_1)$	–	–	–	0.94	0.67	-9.76	-0.18	-6.15	1.78	1.70
	–	–	–	[0.346]	[0.502]	[0.000]	[0.860]	[0.000]	[0.075]	[0.089]
$t(\lambda_2)$	–	–	–	4.09	-3.29	0.25	1.33	6.60	2.14	1.87
	–	–	–	[0.000]	[0.001]	[0.801]	[0.182]	[0.000]	[0.032]	[0.062]
$t(\lambda_3)$	–	–	–	-1.31	0.07	-0.94	0.50	-2.30	-5.80	1.80
	–	–	–	[0.189]	[0.943]	[0.348]	[0.616]	[0.022]	[0.000]	[0.072]

Table 17: **Estimation results for three-year testing periods from 1966 to 2010 using stocks with book-to-market data.** This table presents the estimates of $\boldsymbol{\gamma} = [\ \alpha \ \ \boldsymbol{\lambda}'\ ]'$, where $\boldsymbol{\lambda} = [\ \lambda_1 \ \ \cdots \ \ \lambda_K\ ]'$ is the vector of ex-post factor risk premia, for two commonly used linear asset pricing models: the single-factor CAPM and the three-factor Fama-French model. We report the estimates with and without the EIV correction as well as the realized factor average over the testing period $\bar{\mathbf{f}}_2$. The linear factor model implies that $\alpha = 0$ and $\boldsymbol{\lambda} = \bar{\mathbf{f}}_2$. We consider 15 non-overlapping three-year testing periods from 1966 to 2010. For each testing period, the beta estimation period consists of the preceding five years. All numbers are reported in annualized percentages.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
CAPM										
Without EIV correction										
α	–	–	–	–	–	7.91	-1.01	-5.25	15.38	5.10
λ_1	–	–	–	–	–	5.29	-4.33	-8.85	10.64	8.19
With EIV correction										
α	–	–	–	–	–	4.80	2.77	-1.59	9.05	0.99
λ_1	–	–	–	–	–	8.11	-7.39	-11.84	16.37	11.82
Realized factor average										
λ_1	–	–	–	–	–	5.93	-3.10	-17.27	13.67	11.71
Fama-French model										
Without EIV correction										
α	–	–	–	–	–	10.93	-3.00	-5.96	13.11	6.46
λ_1	–	–	–	–	–	-1.83	-0.10	-11.04	6.28	6.02
λ_2	–	–	–	–	–	7.88	-5.52	-0.19	11.85	4.79
λ_3	–	–	–	–	–	2.04	0.88	5.70	0.89	-4.05
With EIV correction										
α	–	–	–	–	–	20.17	-9.94	-1.53	7.46	6.64
λ_1	–	–	–	–	–	-15.38	8.82	-18.18	9.55	5.15
λ_2	–	–	–	–	–	15.02	-8.89	2.68	14.46	7.74
λ_3	–	–	–	–	–	4.13	2.48	11.69	3.48	-8.16
Realized factor average										
λ_1	–	–	–	–	–	5.93	-3.10	-17.27	13.67	11.71
λ_2	–	–	–	–	–	20.30	-6.32	-13.01	15.30	12.92
λ_3	–	–	–	–	–	3.08	0.40	12.18	11.37	-7.31

Continued on next page

Table 17 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
CAPM										
Without EIV correction										
α	16.73	20.84	8.95	-0.38	9.96	12.01	-1.15	20.08	3.46	4.73
λ_1	-4.99	-12.55	-0.32	9.83	1.77	1.40	15.75	-2.63	1.10	4.37
With EIV correction										
α	18.75	26.73	9.11	-4.90	9.14	10.97	-10.70	20.75	3.28	1.24
λ_1	-6.78	-18.23	-0.49	14.49	2.65	2.60	26.71	-3.57	1.33	7.38
Realized factor average										
λ_1	2.38	9.21	10.00	6.58	10.27	19.34	-3.66	4.89	4.89	0.35
Fama-French model										
Without EIV correction										
α	14.19	18.24	8.72	0.74	9.34	11.83	-1.34	16.32	2.10	4.40
λ_1	-7.16	-10.19	0.55	8.63	1.22	0.98	10.78	-1.65	1.21	3.96
λ_2	3.06	-3.10	-1.04	0.78	2.06	1.20	10.65	2.40	2.33	3.08
λ_3	8.14	7.79	1.21	-3.58	0.25	-0.73	-7.74	4.10	0.76	0.65
With EIV correction										
α	15.19	17.79	3.41	-2.75	8.49	11.29	-7.79	4.60	-2.18	0.46
λ_1	-11.83	-10.72	4.71	12.65	1.40	1.13	12.13	1.53	2.48	6.35
λ_2	6.04	-3.23	0.83	0.29	3.16	2.10	20.28	11.70	7.00	5.22
λ_3	15.16	13.36	8.46	-4.60	0.75	-1.22	-9.40	11.59	4.43	3.91
Realized factor average										
λ_1	2.38	9.21	10.00	6.58	10.27	19.34	-3.66	4.89	4.89	0.35
λ_2	8.68	-5.53	-4.60	1.32	-0.52	-8.93	11.37	9.94	-2.95	9.24
λ_3	16.13	8.80	1.59	-0.72	5.54	0.78	7.15	8.10	2.73	0.06

Table 18: **Test statistics and p-values for three-year testing periods from 1966 to 2010 using 50 clusters based on book-to-market.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 15 non-overlapping three-year testing periods from 1966 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
					CAPM					
Joint χ^2	–	–	–	–	–	150.15	10.76	41.18	166.67	1.64
	–	–	–	–	–	[0.000]	[0.005]	[0.000]	[0.000]	[0.440]
$t\,(\alpha)$	–	–	–	–	–	1.94	1.05	-0.71	2.89	0.40
	–	–	–	–	–	[0.053]	[0.296]	[0.478]	[0.004]	[0.688]
$t\,(\lambda_1)$	–	–	–	–	–	0.91	-1.80	2.68	0.90	0.05
	–	–	–	–	–	[0.362]	[0.072]	[0.007]	[0.367]	[0.963]
					Fama-French model					
Joint χ^2	–	–	–	–	–	82.00	5.73	163.79	14.27	58.56
	–	–	–	–	–	[0.000]	[0.057]	[0.000]	[0.001]	[0.000]
$t\,(\alpha)$	–	–	–	–	–	4.69	-1.93	-0.49	1.69	2.53
	–	–	–	–	–	[0.000]	[0.053]	[0.622]	[0.091]	[0.011]
$t\,(\lambda_1)$	–	–	–	–	–	-4.08	2.13	-0.30	-1.01	-2.41
	–	–	–	–	–	[0.000]	[0.034]	[0.763]	[0.312]	[0.016]
$t\,(\lambda_2)$	–	–	–	–	–	-2.08	-1.55	10.48	-0.68	-3.41
	–	–	–	–	–	[0.037]	[0.120]	[0.000]	[0.497]	[0.001]
$t\,(\lambda_3)$	–	–	–	–	–	0.39	0.76	-0.16	-3.27	-0.42
	–	–	–	–	–	[0.693]	[0.448]	[0.876]	[0.001]	[0.673]

Continued on next page

Table 18 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
					CAPM					
Joint χ^2	209.66	165.54	27.14	25.86	42.16	82.33	429.66	264.68	5.84	282.12
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.054]	[0.000]
$t(\alpha)$	9.68	12.30	5.03	-2.26	6.28	6.72	-4.79	10.59	2.42	0.74
	[0.000]	[0.000]	[0.000]	[0.024]	[0.000]	[0.000]	[0.000]	[0.000]	[0.016]	[0.458]
$t(\lambda_1)$	-5.07	-12.85	-5.20	3.44	-4.85	-8.55	10.97	-2.88	-1.92	4.58
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.004]	[0.054]	[0.000]
					Fama-French model					
Joint χ^2	54.03	88.21	27.66	38.63	34.86	80.55	236.58	38.98	11.11	117.18
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.004]	[0.000]
$t(\alpha)$	6.60	4.29	0.39	-0.85	4.92	4.49	-2.32	1.08	-0.82	0.25
	[0.000]	[0.000]	[0.694]	[0.397]	[0.000]	[0.000]	[0.020]	[0.280]	[0.415]	[0.803]
$t(\lambda_1)$	-5.82	-5.24	-0.91	1.82	-4.27	-6.18	4.27	-1.16	-1.35	2.40
	[0.000]	[0.000]	[0.362]	[0.068]	[0.000]	[0.000]	[0.000]	[0.245]	[0.177]	[0.017]
$t(\lambda_2)$	-1.82	1.46	1.02	-0.82	2.47	6.69	3.57	0.45	2.95	-1.74
	[0.069]	[0.144]	[0.306]	[0.414]	[0.013]	[0.000]	[0.000]	[0.649]	[0.003]	[0.082]
$t(\lambda_3)$	-0.40	1.46	0.57	-2.61	-2.50	-0.78	-8.96	1.16	0.66	1.32
	[0.687]	[0.145]	[0.569]	[0.009]	[0.012]	[0.436]	[0.000]	[0.247]	[0.506]	[0.187]

Table 19: **Test statistics and p-values for three-year testing periods from 1966 to 2010 using 30 clusters based on book-to-market.** This table presents test statistics and p-values for various implications of a linear asset pricing model. Two commonly used models are tested: the single-factor CAPM and the three-factor Fama-French model. We report χ^2 statistics for the joint hypothesis $\alpha = 0$ and $\boldsymbol{\lambda} = \overline{\mathbf{f}}_2$ as well as t statistics for the simple hypotheses $\alpha = 0$ and $\lambda_k = \overline{f}_{k,2}$, for $k = 1, \ldots, K$. We consider 15 non-overlapping three-year testing periods from 1966 to 2010. For each testing period, the beta estimation period consists of the preceding five years. We report the corresponding p-values in square brackets below the test statistics.

Estimation Period	46-50	49-53	52-56	55-59	58-62	61-65	64-68	67-71	70-74	73-77
Testing Period	51-53	54-56	57-59	60-62	63-65	66-68	69-71	72-74	75-77	78-80
					CAPM					
Joint χ^2	–	–	–	–	–	127.55	8.14	31.03	118.18	1.14
	–	–	–	–	–	[0.000]	[0.017]	[0.000]	[0.000]	[0.566]
$t(\alpha)$	–	–	–	–	–	1.86	0.97	-0.66	2.74	0.37
	–	–	–	–	–	[0.062]	[0.334]	[0.511]	[0.006]	[0.714]
$t(\lambda_1)$	–	–	–	–	–	0.86	-1.64	2.46	0.85	0.04
	–	–	–	–	–	[0.388]	[0.102]	[0.014]	[0.396]	[0.967]
					Fama-French model					
Joint χ^2	–	–	–	–	–	76.99	5.66	152.86	13.25	50.19
	–	–	–	–	–	[0.000]	[0.059]	[0.000]	[0.001]	[0.000]
$t(\alpha)$	–	–	–	–	–	4.68	-1.92	-0.47	1.60	2.43
	–	–	–	–	–	[0.000]	[0.055]	[0.640]	[0.111]	[0.015]
$t(\lambda_1)$	–	–	–	–	–	-4.09	2.13	-0.29	-0.95	-2.32
	–	–	–	–	–	[0.000]	[0.033]	[0.771]	[0.342]	[0.020]
$t(\lambda_2)$	–	–	–	–	–	-2.07	-1.51	10.14	-0.66	-3.23
	–	–	–	–	–	[0.039]	[0.130]	[0.000]	[0.512]	[0.001]
$t(\lambda_3)$	–	–	–	–	–	0.37	0.73	-0.15	-3.17	-0.40
	–	–	–	–	–	[0.714]	[0.466]	[0.883]	[0.002]	[0.690]

Continued on next page

Table 19 – continued from previous page

Estimation Period	76-80	79-83	82-86	85-89	88-92	91-95	94-98	97-01	00-04	03-07
Testing Period	81-83	84-86	87-89	90-92	93-95	96-98	99-01	02-04	05-08	08-10
					CAPM					
Joint χ^2	149.65	131.46	23.73	19.98	35.15	66.44	306.55	178.57	4.10	211.81
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.129]	[0.000]
$t(\alpha)$	8.57	10.87	4.73	-2.09	5.78	6.14	-4.36	9.09	2.02	0.72
	[0.000]	[0.000]	[0.000]	[0.037]	[0.000]	[0.000]	[0.000]	[0.000]	[0.043]	[0.470]
$t(\lambda_1)$	-4.51	-11.45	-4.83	3.16	-4.44	-7.81	9.91	-2.49	-1.64	4.43
	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.013]	[0.101]	[0.000]
					Fama-French model					
Joint χ^2	44.12	73.95	24.18	33.92	31.75	76.50	202.82	29.37	9.92	98.97
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.007]	[0.000]
$t(\alpha)$	6.10	4.14	0.37	-0.83	4.65	4.53	-2.26	1.02	-0.79	0.24
	[0.000]	[0.000]	[0.711]	[0.409]	[0.000]	[0.000]	[0.024]	[0.306]	[0.431]	[0.812]
$t(\lambda_1)$	-5.48	-5.06	-0.85	1.77	-4.06	-6.24	4.20	-1.11	-1.23	2.33
	[0.000]	[0.000]	[0.395]	[0.077]	[0.000]	[0.000]	[0.000]	[0.269]	[0.218]	[0.020]
$t(\lambda_2)$	-1.74	1.37	0.99	-0.78	2.33	6.42	3.57	0.42	2.92	-1.69
	[0.082]	[0.170]	[0.324]	[0.435]	[0.020]	[0.000]	[0.000]	[0.677]	[0.004]	[0.091]
$t(\lambda_3)$	-0.38	1.39	0.55	-2.49	-2.37	-0.77	-8.33	1.08	0.63	1.26
	[0.707]	[0.165]	[0.585]	[0.013]	[0.018]	[0.439]	[0.000]	[0.279]	[0.531]	[0.208]