Vanderbilt University Department of Economics Working Paper Series
Nashville, TN 37235-1819

Vanderbilt University Department of Economics Working Papers 15-00006

Tests for the validity of portfolio or group choice in financial and panel regressions

Atsushi Inoue
Vanderbilt University

Barbara Rossi
Universitat Pompeu Fabra

Abstract

In the capital asset pricing model (CAPM), estimating beta consistently is important to obtain a consistent estimate of the price of risk. However, it is often found that the estimate of beta is sensitive to the choice of portfolios used in the estimation. This paper provides a new test to evaluate whether the choice of portfolios in typical asset price regressions is valid, in the sense that the portfolios satisfy two conditions: (i) the way the portfolios are formed are exogenous; and (ii) the choice of the group of assets to include in the portfolios provides enough information to identify the parameters of interest. Thus, checking the validity of the portfolio choice is an important pre-requisite to ensure consistent estimates of the parameters of the model. We illustrate the performance of the test in small samples via Monte Carlo simulations. The proposed test is also applicable to group and pseudo panel data models.

Citation: Atsushi Inoue and Barbara Rossi, (2015) "Tests for the validity of portfolio or group choice in financial and panel regressions", *Vanderbilt University Department of Economics Working Papers,* VUECON-15-00006
Contact: Atsushi Inoue - atsushi.inoue@vanderbilt.edu, Barbara Rossi - barbara.rossi@upf.edu
Submitted: April 14, 2015 **Published:** April 15, 2015
URL: http://www.accessecon.com/Pubs/VUECON/VUECON-15-00006.pdf

1 Introduction

The empirical analysis of the capital asset pricing model (CAPM) has been a priority in finance. Fama and French (1992) proposed a two-step regression methodology to evaluate the CAPM model that is still widely used in finance. The procedure consists in grouping assets into portfolios, calculating their return above and beyond the risk-free return, and then regressing average excess portfolio returns on a series of factors. The model predicts, among other things, that the factors should be significant and allows to make inference on an important quantity of interest in empirical finance, the price of risk. The Fama and French (1992) regression approach goes beyond finance: for example, the carry-trade literature (e.g. Verdelhan, 2013) is an application of Fama-French regressions to exchange rates.

Portfolios are typically used to reduce noise in the estimated regressions. However, empirically, results turn out to be very sensitive to the portfolio choice (Kandel and Stambaugh, 1995); that is, it is possible that the empirical results are driven by the way the portfolios are chosen, and result in an especially poor or especially good performance of the theoretical model to be tested. Notwithstanding the importance of the issue in practice, no procedure is currently available to evaluate the appropriateness of the portfolio choice.[1] In this paper, we propose a new test to evaluate whether the choice of the portfolios are appropriate. For example, our proposed test can be used to determine whether portfolios should be sorted based on industry or size to identify beta and thus the price of risk.

We propose a new test to evaluate the validity and relevance of factors in these regressions. The test that we propose evaluates whether the portfolio chosen by the researcher satisfies two conditions. The first condition is that it is exogenous. Exogeneity requires that the way portfolios are formed should be uncorrelated with idiosyncratic shocks to individual assets. In other words, any noise at the individual asset level will vanish at the portfolio level. If this condition is not met, there will be systematic biases when parameters are estimated from portfolio regressions. Thus, lack of exogeneity leads to inconsistent parameter estimates and invalidates tests of significance. The second condition is that the choice of how to group assets into portfolios is valid so that parameters of asset pricing models are strongly identified.

[1] Cochrane (2005, p. 218) suggests to evaluate the invariance of the empirical results to the portfolio choice. However, if the empirical results turn out to depend on the portfolio choice, this approach does not shed light on which ones the researcher should trust.

Even when the exogeneity condition is satisfied, if the identification condition is not satisfied, parameters cannot be consistently estimated from portfolio regressions. It is only when both conditions are met that the parameter of interest (the effect that the factor has on the excess return) can be identified and valid inference on it as well as on the price of risk can be made. Our proposed test is a joint test for exogeneity and strong identification. When both conditions are met, the parameter of interest (the effect that the factor has on the excess return, that is the "beta" of the portfolio) can be identified.

We consider a framework in which the number of individual assets and the number of time periods are both large, a typical situation in the empirical CAPM literature. Even when the number of portfolios is small, the large time series dimension implies that the number of portfolio-time dummy variables is also large. Our analysis starts from the observation that the Fama-French regression estimator can be viewed as a two-stage least squares estimator in which these portfolio-time dummy variables are instruments, and are many. Thus, our methodology involves assessing whether the parameters are identifiable in an IV regression with many instruments. Because the number of instruments is large (Kunitomo, 1980; Morimune, 1983; Bekker, 1994), the two-stage least squares (2SLS) estimator is inconsistent but the jackknife instrumental variables estimator (JIVE) is consistent (Chao, Swanson, Hausman and Newey, 2012). Koopmans and Hood (1953) establish the condition for identifiability for instrumental variables estimators. Their condition involves evaluating the rank of the population coefficient matrix from regressing endogenous variables on exogenous variables. Because this matrix is infinite-dimensional, we transform this infinite dimensional problem into a finite dimensional one. We show that testing the identifiability of parameters of interest boils down to testing the rank of a finite-dimensional matrix in our framework. In a recent paper, Kleibergen and Paap (2006) propose a test of matrix rank using the singular value decomposition. They assume that the matrix is non-symmetric and thus their test cannot be used for our purpose (Kleibergen and Paap, 2006, p.103). Donald, Fortuna and Pipiras (2007) develop rank tests for symmetric matrices. Our implementation of rank testing is based on one of their tests.

Our proposed test is related to the existing tests of Anatolyev and Gospodinov (2011) and Chao, Hausman, Newey, Swanson and Woutersen (2014). Anatolyev and Gospodinov (2011) develop the Anderson-Rubin and the J overidentifying restriction tests when there are many instruments. Chao, Hausman, Newey, Swanson and Woutersen (2014) develop a test of overidentifying restrictions based on a jacknife version of Sargan's test statistic. These two tests are designed for testing instruments. Using the many instrument framework of Chao et al. (2012), we develop a test for the rank of a finite-dimensional matrix and we test the identification condition as well as the exogeneity condition.

While we motivate this problem using the Fama-French model, our test is applicable to a wide class of problems in which group averages (Angrist, 1988) and pseudo panels (Deaton, 1985) are used where group dummies are instruments that consist of ones and zeros.

The rest of the paper is organized as follows. Section 2 motivates our problem using the Fama-French regression. Section 3 presents assumptions and theoretical results. Section 4 shows Monte Carlo simulation results. Section 5 concludes the paper.

2 Motivation

The CAPM model suggests a particular relationship between the return of asset "i" at time "t" ($R_{i,t}$) and the excess market return at time "t" ($R_{M,t} - R_{f,t}$):[2]

$$R_{i,t} = R_{f,t} + [R_{M,t} - R_{f,t}]\,\beta_i + \varepsilon_{i,t},$$

where $i = 1, ..., N$, $t = 1, ..., T$, N is the total number of assets and T is the total sample size. Note that the excess market return is the difference between the market return, $R_{M,t}$, and the risk-free rate, $R_{f,t}$. Thus, the excess market return is a common factor that explains the variability of the individual asset returns. The excess market return may not be the only factor. According to the more general intertemporal capital asset pricing model (ICAPM), any state variable that predicts future investment opportunities serves as a factor. These additional factors can be selected via: (a) economic theory;[3] (b) statistical principal components; or (c) firm characteristics (e.g. size, value, momentum). Thus, the individual asset regression becomes:

$$R_{i,t} = R_{f,t} + \beta_i' f_t + \varepsilon_{i,t}, \tag{1}$$

where f_t is a $(K \times 1)$ vector of factors and β_i is a vector of $(K \times 1)$ individual assets' betas. However, estimates of beta for individual assets are very impre-

[2] See Fama and French (2004, p. 32), for example.

[3] For example, Lettau and Ludvigson (2001) suggest that the consumption -to -wealth -to -income ratio ("cay") is a common factor.

cise, creating a measurement error problem. To address this issue, researchers estimate eq. (1) using portfolios returns, i.e. $R_{p,t} \equiv \sum_{i=1}^{N} \omega_{i,t} R_{i,t}$, where $\omega_{i,t}$ are weights (possibly time-varying). According to Fama and French (2004), estimates of β_p for diversified portfolios are more precise than estimates of β_i for individual assets, reducing the measurement error problem. Creating portfolios, however, reduces the range of β_i that can be estimated. Thus, researchers first sort securities based on their value of β and then form portfolios using quantiles of the distribution of β: the first portfolio is constructed using the assets with the lowest β_i; the second portfolio uses the assets with the smallest among the remaining β_i and so forth, until the last portfolio, which contains assets that have the highest β_i.

In the rest of the paper, we assume that β_i is the same for every return belonging to that portfolio: $\beta_i = \beta_g \; \forall i \in P_g$, where P_g is the set of i's that belong to portfolio g; in this case, ω_i is a dummy variable that equals one if asset i belongs to portfolio g. Thus, eq. (1) implies $\sum_{i=1, i \in P_g}^{N} \omega_i R_{i,t} = R_{f,t} + f_t' \left(\sum_{i=1, i \in P_g}^{N} \omega_{i,t} \beta_i \right) + \sum_{i=1, i \in P_g}^{N} \omega_i \varepsilon_{i,t}$, that is:

$$R_{p,t} = R_{f,t} + \beta_g' f_t + \varepsilon_{p,t} \tag{2}$$

where β_g is the beta of the portfolio: $\sum_{i=1, i \in P_g}^{N} \omega_i \beta_i = \beta_g$, $R_{p,t} = \sum_{i=1, i \in P_g}^{N} \omega_i R_{i,t}$, $\varepsilon_{p,t} = \sum_{i=1, i \in P_g}^{N} \omega_i \varepsilon_{i,t}$.

In this paper we are interested in testing whether the choice of which asset is included in the portfolio is valid, in the sense that the factors are exogenous and the grouping into portfolio provides relevant information to identify the parameter β_p. Note that the exogeneity of the factors is a necessary condition for the consistency of β_p in eq. (2), and it corresponds to the exogeneity of the factors in the individual asset return model, eq. (1), in the sense that if the factors are exogenous in the individual assets' return regressions then they are exogenous in the portfolio regressions.

3 Asymptotic Theory

Suppose that in portfolio g the i-th individual asset excess return ($r_{g,i,t} = R_{g,i,t} - R_{f,t}$) satisfies

$$r_{g,i,t} \;=\; \beta_g' f_t + \varepsilon_{g,i,t}, \tag{3}$$

at time t, for $g = 1, 2, ..., G$, $i = 1, ..., n$ and $t = 1, 2, ..., T$. For example, in the Fama-French three factor model, the $(K \times 1)$ vector of observed factors,

f_t, consists of (i) the excess return on a broad market portfolio, (ii) the difference between the return on a portfolio of small stocks and (iii) the return on a portfolio of large stocks, and the difference between the return on a portfolio of high-book-to-market stocks and the return on a portfolio of low-book-to-market equity (BE/ME) stocks in addition to the intercept term. For notational simplicity, we assume that the number of assets is the same in each portfolio at every time period and is denoted by n. [4]

We will show that an estimator of β in equation (3) can be interpreted as an instrumental variables estimator. By stacking (3) for $g = 1, 2, ..., G$, we can write (3) as:

$$\begin{bmatrix} r_{1,i,t} \\ r_{2,i,t} \\ \vdots \\ r_{G,i,t} \end{bmatrix} = \begin{bmatrix} \beta_1' \\ \beta_2' \\ \vdots \\ \beta_G \end{bmatrix} f_t + \begin{bmatrix} \varepsilon_{1,i,t} \\ \varepsilon_{2,i,t} \\ \vdots \\ \varepsilon_{G,i,t} \end{bmatrix}, \tag{4}$$

for $i = 1, ..., n$ and $t = 1, ..., T$.

The model can be more compactly written as

$$\begin{aligned} y_{i,t} &= \beta' f_t + u_{i,t}, && (5) \\ x_{i,t} &= f_t + v_{x,i,t}, && (6) \end{aligned}$$

for $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, where $y_{i,t} = [y_{1,i,t} \ \cdots \ y_{G,i,t}]'$, $\beta = [\beta_1 \ \beta_2 \ \cdots \ \beta_G]$ is a $(K \times G)$ matrix of parameters, $u_{i,t} = [\varepsilon_{1,i,t} \ \cdots \ \varepsilon_{G,i,t}]'$, and $v_{x,i,t}$ is a idiosyncratic measurement error (which is zero if the factor is directly observed). Here we assume that the number of signals $x_{i,t}$ equals n. In case there are G multiple of these n signals, one can think of $x_{i,t}$ as their average, i.e., $x_{i,t} = (1/G)\sum_{g=1}^{G} x_{g,i,t}$ where $x_{g,i,t}$ is a portfolio g version of $x_{i,t}$.

Taking time averages of (5) and (6) gives:

$$\begin{aligned} \bar{y}_{g,t} &= \beta_g' f_t + \bar{u}_{g,t}, \quad g = 1, ..., G, && (7) \\ \bar{x}_t &= f_t + \bar{v}_{x,t}, && (8) \end{aligned}$$

where $\bar{y}_{g,t} = (1/n)\sum_{i=1}^{n} y_{g,i,t}$, $\bar{x}_t = (1/n)\sum_{i=1}^{n} x_{i,t}$, $\bar{u}_{g,t} = (1/n)\sum_{i=1}^{n} \varepsilon_{i,g,t}$, and $\bar{v}_t = (1/n)\sum_{i=1}^{n} v_{x,i,t}$. If f_t is not observable, f_t in equation (7) is replaced by $\bar{x}_t$. If f_t is observable, $\bar{x}_t$ is numerically identical to f_t.

[4]In practice, this "balanced panel" assumption may not be satisfied. We can allow for unbalanced panels as long as the way unbalanced panel data are unbalanced is orthogonal to u_j.

It can be shown that the least squares estimator from regressing $\bar{y}_{g,t}$ on $\bar{x}_t$,

$$\left(\sum_{t=1}^{T} \bar{x}_t \bar{x}_t'\right)^{-1} \sum_{t=1}^{T} \bar{x}_t \bar{y}_{g,t}' \tag{9}$$

is an instrumental variables estimator of β_g. Define $nT \times K$ matrix X, $nT \times G$ matrix Y, and $nT \times T$ matrix Z by

$$X = \begin{bmatrix} x_{1,1}' \\ x_{1,2}' \\ \vdots \\ x_{1,T}' \\ x_{2,1}' \\ x_{2,2}' \\ \vdots \\ x_{2,T}' \\ \vdots \\ x_{n,1}' \\ x_{n,2}' \\ \vdots \\ x_{n,T}' \end{bmatrix}, \quad Y = \begin{bmatrix} y_{1,1,1} & y_{2,1,1} & \cdots & y_{G,1,1} \\ y_{1,1,2} & y_{2,1,2} & \cdots & y_{G,1,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,1,T} & y_{2,1,T} & \cdots & y_{G,1,T} \\ y_{1,2,1} & y_{2,2,1} & \cdots & y_{G,2,1} \\ y_{1,2,2} & y_{2,2,2} & \cdots & y_{G,2,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,2,T} & y_{2,2,T} & \cdots & y_{G,2,T} \\ \vdots & \vdots & \vdots & \vdots \\ y_{1,n,1} & y_{2,n,1} & \cdots & y_{G,n,1} \\ y_{1,n,2} & y_{2,n,2} & \cdots & y_{G,n,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,n,T} & y_{2,n,T} & \cdots & y_{G,n,T} \end{bmatrix}, \tag{10}$$

and $Z = I_T \otimes \ell_n$, respectively, where ℓ_n is the $(n \times 1)$ vector of ones. Then the time average regression estimator is a 2SLS estimator of β with Z as the instruments, i.e.,

$$\begin{aligned} \widehat{\beta}_{2SLS} &= [\widehat{\beta}_{1,2SLS}\ \widehat{\beta}_{2,2SLS}\ \cdots\ \widehat{\beta}_{G,2SLS}] \\ &= [X'Z(Z'Z)^{-1}Z'X]^{-1}X'Z(Z'Z)^{-1}Z'Y \\ &= \left[\sum_{i,j=1}^{n}\sum_{s,t=1}^{T} x_{i,s}P_{i,j,s,t}x_{j,t}'\right]^{-1} \sum_{i,j=1}^{n}\sum_{s,t=1}^{T} x_{i,s}P_{i,j,s,t}y_{j,t}', \end{aligned} \tag{11}$$

where $x_{i,s}'$ is the $(i-1)T+s$-th row of X, $y_{j,t}'$ is the $(j-1)T+t$-th row of Y, and $P_{i,j,s,t}$ is the $((i-1)T+s, (j-1)T+t)$-th element of the projection matrix $P = Z(Z'Z)^{-1}Z'$. Because the number of instruments equals the number of time periods, which is large relative to the total sample size in typical empirical applications, the two stage least squares estimator of $\beta = [\beta_1\ \beta_2\ \ \cdots\ \beta_G]$ is inconsistent (Bekker, 1994). In addition, volatilities may vary across different asset returns and portfolios. Thus, we consider a version of the jackknife instrumental variables (JIVE) estimator of Chao et al. (2012), which allows for heteroskedasticity as well as many weak instruments:

$$\begin{aligned} \widehat{\beta}_{JIVE} &= \left[\ \widehat{\beta}_{JIVE,1}\ \ \widehat{\beta}_{JIVE,2}\ \ \cdots\ \ \widehat{\beta}_{JIVE,G}\ \right] \\ &= \left[\sum_{i\neq j}\sum_{s,t=1}^{T} x_{i,s}P_{i,j,s,t}x'_{j,t}\right]^{-1}\sum_{i\neq j}\sum_{s,t=1}^{T} x_{i,s}P_{i,j,s,t}y'_{j,t}, \end{aligned} \tag{12}$$

where $\sum_{i\neq j}$ denotes $\sum_{i=1}^{n}\sum_{j=1,i\neq j}^{n}$ in the rest of the paper.

To understand the identifiability of β, we need to write the model in reduced form:

$$\begin{aligned} Y_{i,t} &\equiv \left[\begin{array}{c} x_{i,t} \\ y_{i,t} \end{array}\right] = \left[\begin{array}{c} \Pi_{x,T} \\ \Pi_{y,T} \end{array}\right] z_{i,t} + \left[\begin{array}{c} v_{x,i,t} \\ v_{y,i,t} \end{array}\right] \\ &= \left[\begin{array}{c} \Pi_{x,T} \\ \beta'\Pi_{x,T} \end{array}\right] z_{i,t} + \left[\begin{array}{c} v_{x,i,t} \\ v_{y,i,t} \end{array}\right] = \Pi_T z_{i,t} + v_{i,t}. \end{aligned} \tag{13}$$

where $z'_{i,t}$ is the $j = (i-1)T + t$-th row of Z, $\Pi_{x,T} = f' = [f_1\ f_2\ \cdots\ f_T]$ is $(K \times T)$, and $\Pi_{y,T} = \beta'\Pi_{x,T}$ is $(G \times T)$. [5] We now state Koopmans and Hood's (1953) necessary and sufficient condition for identifiability of β:

Proposition 1 (Koopmans-Hood Rank Condition). Suppose that (4) and (13) hold. For given N and T, β is identified if and only if $(K + G) \times T$ matrix Π_T has rank K, where K is the number of parameters in β in equation (13).

Remarks. When instruments are not exogenous, $\Pi_{y,T}$ cannot be written as a linear combination of the columns of $\Pi_{x,T}$ and thus the rank will be greater than K. The rank of Π_T will be less than K when the instruments are not relevant. The first case is a situation where the model is misspecified because $\Pi_{y,T}$ is not a linear combination of $\Pi_{x,T}$. As we will show in the Monte Carlo section, for example this may also happen when the researcher forgets to include a group-specific constant in the return regression. The second case is

[5] To obtain the reduced form (13), we start with (6): $x_{i,t} = f_t + v_{x,i,t}$ or $f_t = x_{i,t} - v_{x,i,t}$; substituting this in eq. (5), we have: $y_{i,t} = \beta'(x_{i,t} - v_{x,i,t}) + u_{i,t}$; rewriting eq. (6) as $x_{i,t} = \Pi_{x,T}z_{i,t} + v_{x,i,t}$ and substituting it in the preceeding equation for $y_{i,t}$: $y_{i,t} = \beta'\Pi_{x,T}$.

a situation of no identification because some of the factors are spurious and the rank of $\Pi_{x,T}$ is not full. For example, this is the case when $f_t = [f_{1t}, f_{2t}]'$ and $f_{2t} = 0\ \forall t$ or when the factors are linearly dependent.

We impose the following conditions:

Assumption.

(a) $Z = I_T \otimes \ell_n$.

(b) *Both n and T diverge to infinity while G is fixed.*

(c) $\|(1/T)\sum_{t=1}^T f_t f_t'\| \leq C$ *and* $\lambda_{\min}((1/T)\sum_{t=1}^T f_t f_t') \geq 1/C$ *almost surely.* $\sup_t E\|f_t\|^8 \leq C$.

(d) *Conditional on Z, $\{v_{i,t}\}$ are cross-sectionally independent and form a Martingale difference sequence in that $E(v_{i,s}v_{j,t}|Z) = 0$ for all i, j, s, t such that $i \neq j$ and $E(v_{it}|F_{i,t-1}, Z) = 0$ where $F_{i,t-1}$ is the sigma field generated by f_{s+1} and $v_{j,s}$ for all j and all $s < t$, and* $\sup_{i,t} E(\|v_{i,t}\|^8|Z) \leq C$.

(e)
$$\Sigma \ = \ \lim_{n\to\infty} \mathrm{Cov}\left(\frac{1}{\sqrt{nT}} \sum_{i\neq j} \sum_{s,t=1}^{T} \mathrm{vech}(Y_{i,s}P_{i,j,s,t}Y_{j,t}')\right) \tag{14}$$
is positive definite where $P_{i,j,s,t}$ is the $((i-1)T+s, (j-1)T+t)$-th element of the projection matrix $P = Z(Z'Z)^{-1}Z'$.

Remarks. An analog of assumption 1 of Chao et al. (2011) is satisfied under our assumptions (a) and (b). In their notation, we assume that $\Upsilon_i = \Pi Z_i$, $S_n = \sqrt{nT} I_{(G+K)\times(G+K)}$, $r_n = nT$, $K = T$. Because $z_{i,t}$ consists of zeros and ones and the elements of $z_{i,t}$ sum to one and Z has rank T, $P_{i,i,s,s} = 1/n$. Thus, their assumption that $P_{ii} < C < 1$ is satisfied. Assumption (b) implies $\|\sum_{i=1}^n \sum_{t=1}^T \Pi_T z_{i,t} z_{i,t}' \Pi_T'/(nT)\| \leq C$ and $\lambda_{\min}(\sum_{i=1}^n \sum_{t=1}^T \Pi_T z_{i,t} z_{i,t}' \Pi_T'/(nT)) \geq 1/C$ almost surely.

Their assumption 2 is also satisfied under our assumptions (b) and (c). Their assumption 3 is satisfied under our assumption (d). Their assumption 4 is trivially satisfied because we assume that the reduced form is linear in the instruments. Their assumption 5 is imposed in our assumption (e).

Given our proposition, testing identifiability of β boils down to testing whether or not the rank of Π_T is K. Rather than test the rank of Π_T whose dimension $(K \times T)$ diverges to infinity, we will test the rank of

$$H = \lim_{n,T\to\infty} \frac{1}{nT} \sum_{i\neq j} \sum_{s,t=1}^{T} \Pi_T z_{i,s} z'_{j,t} \Pi'_T, \tag{15}$$

whose dimension $((G+K)\times(G+K))$ is fixed relative to the sample size. Let

$$\hat{H} = \frac{1}{nT} \sum_{i\neq j} \sum_{s,t=1}^{T} (Y_{i,s} P_{i,j,s,t} Y'_{j,t}); \tag{16}$$

then, H is effectively the probability limit of $\hat{H}$, and its rank is the same as the rank of the limit of Π_T.

Theorem 1 (Asymptotic Distribution of Concentration Matrix).

$$\sqrt{T}\left\{ \frac{1}{nT} \sum_{i\neq j} \sum_{s,t=1}^{T} \left[vech(Y_{i,s} P_{i,j,s,t} Y'_{j,t}) - vech(H) \right] \right\} \xrightarrow{d} N(0,\Sigma), \tag{17}$$

where Σ is given in assumption (e).

Remarks. Donald, Fortuna and Pipiras (2007) show that, for the asymptotic covariance matrix of a symmetric matrix estimator to be positive definite, the symmetric matrix estimator must be indefinite (Proposition 2.1 of Donald, Fortuna and Pipiras, 2007, p.1219). To be precise, their proof requires that the symmetric matrix estimator needs to be indefinite with positive probability (it does not have to be indefinite with probability one as their statement might imply). We show that this condition is indeed satisfied in our problem.

Because

$$\frac{1}{nT} \sum_{i\neq j} \sum_{s,t=1}^{T} Y_{i,s} P_{i,j,s,t} Y'_{j,t} \xrightarrow{p} H, \tag{18}$$

and H is positive semi-definite with rank K, the largest eigenvalue of $\sum_{i\neq j} Y_i P_{ij} Y'_j$ is positive with probability approaching one under the null hypothesis that β is identifiable.

For $(G+K)\times 1$ vector a,

$$\begin{aligned}
a'\frac{1}{nT}\sum_{i\neq j}\sum_{s,t=1}^{T} Y_{i,s}P_{i,j,s,t}Y'_{j,t}a &= \frac{1}{nT}\sum_{i\neq j}\sum_{s,t=1}^{T} b'_{i,s}b_{j,t} \\
&= \frac{1}{nT}\left(\sum_{i=1}^{n}\sum_{s=1}^{T} b'_{i,s}\sum_{j=1}^{n}\sum_{t=1}^{T} b_{j,t} - \sum_{i=1}^{n} b'_i b_i\right) \\
&= \frac{1}{T}\left[(n-1)\bar{b}'\bar{b} - \sum_{k=1}^{G+K}\widehat{Var}(b_k)\right] \qquad (19)
\end{aligned}$$

where $b_{i,t} = (Z'Z)^{-\frac{1}{2}}z_{i,t}Y'_{i,t}a$, $\bar{b} = (1/n)\sum_{i=1}^{n}\sum_{t=1}^{T} b_{i,t}$, and $\widehat{Var}(b_k)$ is the sample variance of the k-th element of $\sum_{t=1}^{T} b_{i,t}$. For given n and T, the sum of the variances of $\sum_{t=1}^{T} b_{i,t}$ can be larger than $n\bar{b}'\bar{b}$ with positive probability, provided that the support of v_i is large enough, Thus, the smallest eigenvalue is negative with positive probability. Therefore, the necessary condition of Donald et al. (2007) is satisfied provided that the support of v_i is large enough given Z.

Theorem 1 provides a basis for testing the rank of H. Therefore we will extend Donald et al.'s (2007) approach to symmetric matrices. Following Kleibergen and Paap (2006) we may also normalize $\hat{H}$ and consider $C\hat{H}C'$ where C is a $(G+K)\times(G+K)$ non-singular matrix. For example, C may be the square root of the covariance matrix of the endogenous variables.

Corollary 1 (Asymptotic Distribution of Scaled Concentration Matrix) In addition to Assumptions (a)–(e), suppose that

(f) $D^{+}_{G+K}(C\otimes C)D^{+}_{G+K}$ has rank $(G+K)(G+K+1)/2$.

Then

$$\sqrt{T}vech(C\hat{H}C' - CHC') \xrightarrow{d} N(0,\Omega), \qquad (20)$$

where

$$\Omega = D^{+}_{G+K}(C\otimes C)D^{+}_{G+K}\Sigma D^{+\prime}_{G+K}(C'\otimes C')D^{+\prime}_{G+K},\ D^{+}_{G+K} = (D'_{G+K}D_{G+K})^{-1}D'_{G+K}$$

and D_{G+K} is the $(G+K)^2\times(G+K)(G+K+1)/2$ duplication matrix (Magnus and Neudecker, 1999, p.49).

Given the asymptotic normality of $\widehat{H}$, we employ Donald et al.'s (2007) implementation of Kleibergen and Paap's (2006) singular value decomposition (SVD) rank test. For completeness, we provide the definition of the test statistic.

To test that the instruments are orthogonal, we consider the null hypothesis that $\widehat{H}$ is rank K against the alternative hypothesis that $\widehat{H}$ is rank $(K+1)$. To test that the instruments are not relevant, we test the null hypothesis that $\widehat{H}$ is rank $(K-1)$ against the alternative hypothesis that $\widehat{H}$ is rank K. Let K_0 be the rank under the null hypothesis. That is, $K_0 = K$ under the first null hypothesis and $K_0 = K-1$ under the second null hypothesis.

For the notational simplicity, we assume that $\widehat{H}$ is used. Because $\widehat{H}$ is symmetric, its singular value decomposition is identical to the Schur decomposition of $\widehat{H}$:

$$\widehat{H} = UDU' = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} U'_{11} & U'_{21} \\ U'_{12} & U'_{22} \end{bmatrix}, \tag{21}$$

where U is a unitary matrix, D is a diagonal matrix whose diagonal elements are the eigenvalues of $\widehat{H}$ in non-increasing order, U_{11} and D_{11} are $(K_0 \times K_0)$, U_{12} and U'_{21} are $(K_0 \times (K+G-K_0))$, and U_{22} and D_{22} are $((K+G-K_0) \times (K+G-K_0))$. Define

$$\begin{aligned} A_\perp &= \begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix} U_{22}(U_{22}U'_{22})^{-\frac{1}{2}} \\ &= \begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix}, && (22) \\ B_\perp &= A'_\perp, && (23) \\ \widehat{\Lambda} &= (U_{22}U'_{22})^{-\frac{1}{2}} U_{22} D_{22} U'_{22} (U_{22}U'_{22})^{-\frac{1}{2}}, && (24) \end{aligned}$$

Then under the null hypothesis that the rank of H is K_0, it follows from Proposition 4.1 of Donald et al. (2007) that

$$T \cdot vech(\widehat{\Lambda})' \widehat{\Omega}^{-1} vech(\widehat{\Lambda}) \xrightarrow{d} \chi^2_{(K+G-K_0)(K+G-K_0+1)/2}, \tag{25}$$

where

$$\widehat{\Omega} = D^{+}_{K+G-K_0}(B_\perp \otimes A'_\perp) D_{G+K-1} \widehat{\Sigma} D'_{G+K-K_0}(B'_\perp \otimes A_\perp) D^{+\prime}_{K+G-K_0},$$

$D^+_{K+G-K_0} = (D'_{K+G-K_0} D_{K+G-K_0})^{-1} D'_{K+G-K_0}$ and D_{K+G-K_0} is the $((K+G-K_0)^2 \times (K+G-K_0)(K+G-K_0+1)/2)$ duplication matrix.

We suggest the following testing procedure:

Step 1. Test if $K_0 = K$. If this null is rejected, the model is misspecified. If this null is not rejected, proceed to step 2.

Step 2. Test if $K_0 = K - 1$. If this null hypothesis is rejected, the model is correctly specified and the parameter is identified (thus we fail to reject the null hypothesis that the model is correct and the parameter is identified). If the second smallest eigenvalue is also zero, then the parameter is not identified.

To make our testing procedure operational, one needs a consistent estimator of Σ, $\widehat{\Sigma}$. We need to find the asymptotic covariance matrix of the *vech* of

$$\frac{1}{nT} \sum_{i \neq j} \sum_{t=1}^{T} Y_{i,t} P_{i,j,t,t} Y'_{j,t}. \tag{26}$$

For the factor model, we propose the following estimator of the asymptotic covariance matrix:

$$\begin{aligned} \widehat{\Sigma} &= \frac{1}{n^2 T} \sum_{t=1}^{T} \left[\sum_{i<j} \text{vech}(Y_{i,t} Y_{j,t} + Y_{j,t} Y'_{i,t}) \text{vech}(Y_{i,t} Y_{j,t} + Y_{j,t} Y'_{i,t})' \right. \\ & \left. -\hat{\mu}_t \hat{\mu}'_t \right], \end{aligned} \tag{27}$$

where $\hat{\mu}_t = (1/n) \sum_{i \neq j} (\text{vech}(Y_{i,t} Y_{j,t} + Y_{j,t} Y_{i,t}))$.

Proposition 2 (Consistency of the Asymptotic Covariance Matrix Estimator) Suppose that Assumptions (a)–(e) hold. Then

$$\widehat{\Sigma} \xrightarrow{p} \Sigma. \tag{28}$$

4 Monte Carlo Experiments

We consider data generating processes (DGPs) calibrated on an asset pricing model with parameters estimated from U.S. data. We consider the traditional

CAPM model with parameters calibrated using gross returns on the three-month Treasury-bill and the 25 Fama and French size and book-to-market portfolios from 1952:2 to 2000:4, for a total of 195 *time series* observations. The data are the same as in Gospodinov et al. (2013).

We consider three data generating processes (DGPs). In the first DGP, the parameter β is the following CAPM model is identified:

$$\begin{aligned} r_{igt} &= \beta f_t + \sigma \varepsilon_{igt}, && (29) \\ f_{igt} &= f_t + \sigma \eta_{igt}, && (30) \\ f_t &= \mu_f + \sigma_f \zeta_t, && (31) \end{aligned}$$

for $i = 1, ..., n$, $g = 1, ..., G$, $t = 1, ..., T$, where $G \in \{1, 5\}$ is the number of Fama-French portfolios, $n \in \{100, 200\}$ is the number of assets in each group, and $T \in \{100, 200\}$ is the number of time periods. Because the identifiability of β does not depend on its value, the value of β is set to 1 without of generality. The scalar common factor f_t is the simulated market return at time t (generated from a normal with mean zero and variance σ_f^2, where σ_f^2 is set to the variance of the market return in the data, $\sigma_f^2 = 0.0066$).[6] The calibrated value of σ is based on Ang, Liu and Schwarz (2008): $\sigma^2 = 0.1225$. ε_{igt}, η_{igt} and ζ_t are independent iid standard normal random variables.

To see why β is identifiable in this model rewrite the model using the notation in the previous section as:

$$\begin{aligned} x &= (\ell_G \otimes f) \otimes \ell_{N_{gt}} + v_x \\ &= Z\Pi_x + v_x, && (32) \\ y &= (\ell_G \otimes f) \otimes \ell_{N_{gt}} \beta + \varepsilon \\ &= Z\Pi_x \beta + v_y, && (33) \end{aligned}$$

where $Z = I_{GT} \otimes \ell_{N_{gt}}$, $\Pi_x = \ell_G \otimes f$, $\Pi_y = \Pi_x \beta$, $v_x = \sigma\eta$, $v_y = \varepsilon - v_x \beta$. Then note that

$$\Pi = [\Pi_x \; \Pi_y] = [\ell_G \otimes f \;\; (\ell_G \otimes f)\beta] \qquad (34)$$

actually has rank 1 with probability one. Thus, β is identifiable with probability one in this model.

[6] The mean of the market return for the sample period we consider is 0.0191 and it is indeed close to zero, so we set $\mu_f = 0$.

We calculate the test statistic for testing that Π has rank 0 (under-identification) and the test statistic for testing that Π has rank 1 (correct specification). Then we calculate frequencies in which each of the two null hypotheses is rejected at the 5% significance level over 1,000 Monte Carlo simulations. Table 1 shows that the rank 0 test always rejects the null of under-identification while the rank 1 tends to have good size across different sample sizes and portfolio sizes.

Next, we consider an alternative hypothesis under which the model is misspecified. Specifically, we replace equation (29) by

$$r_{igt} \;=\; a_g + \beta f_t + \sigma \varepsilon_{igt}, \tag{35}$$

where a_g is a portfolio-specific mean that is 0.025 times the sample means of the portfolio returns:

$$[1.0131\ 1.0394; 1.0394\ 1.0460\ 1.0492\ 1.0286\ 1.0369\ 1.0423\ 1.0441\ 1.0474]'.$$

The tests always have power one when the sample mean is used as a_g. To make the Monte Carlo experiment more meaningful, we use 0.025 times the sample portfolio means. To see the effect of G, n and T on the power of the tests, we multiply the sample mean by 0.05. The portfolio specific mean could be interpreted as measurement error that does not disappear asymptotically when constructing portfolio averages. Thus, this is a case in which the researcher has chosen a characteristic to sort portfolios that results in large measurement errors and renders the parameter estimate inconsistent. As a result the population first-stage regression coefficient matrix takes the form of

$$\Pi \;=\; [\Pi_x \;\; \Pi_y] \;=\; [\ell_G \otimes f \;\; (\ell_G \otimes f)\beta + a \otimes \ell_T] \tag{36}$$

The rank of this matrix is 2 with probability one, meaning that the model is misspecified.

Table 2 shows the empirical rejection frequencies when testing the null hypothesis that Π has rank 0 and when testing the null hypothesis that Π has rank 1. Again the rank 0 test always rejects the null hypothesis of under-identification. The rank 1 test rejects the null hypothesis of correct specification with probability much higher than the nominal size. The rejection frequency of the rank 1 test is increasing in sample sizes and approaches one as expected.

Finally, we consider a DGP in which there is no factor structure:

$$\begin{aligned} r_{igt} &= \sigma\varepsilon_{igt}, \\ f_{igt} &= \sigma\eta_{igt}, \end{aligned}$$

Using the notation in the previous section, the model can be written as

$$\begin{aligned} x &= Z\Pi_x + v_x, \\ y &= x\beta + \varepsilon = Z\underbrace{\Pi_x\beta}_{=\Pi_y} + v_y, \end{aligned}$$

where $\Pi_x = 0$. Because rank(Π_T) =rank$\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = 0$, the parameter β is not identified. Because the rank 0 test is designed to test the null of under-identification, not rejecting this null means that the parameter is likely to be under-identified. Table 3 shows that the rank 0 test is undersized and suggests that one would conclude that the parameter is not identified. It is interesting to note that the rejection frequencies of the rank 1 test are very small in this case. This is because the rank 1 test is designed to test the null hypothesis that the rank is one against the alternative hypothesis that the rank is higher than one.

5 Conclusion

This paper proposed a new test to evaluate the portfolio choice in widely-used finance regressions based on portfolio returns. The new test allows researchers to tackle the empirical problem that empirical results are sensitive to the portfolio choice. This paper provides techniques to evaluate whether the portfolio choice is appropriate by testing the validity of the choice of how to group assets into portfolios and the exogeneity of the portfolio choice. Only under these condition the parameters of interest(such as β and the price of risk of a portfolio) can be correctly estimated. Monte Carlo simulations based on the CAPM model demonstrate the good size and power properties of our test.

While we motivate this problem from the portfolio choice perspective, our test can be used to test for identification in group and pseudo and panel data models as well.

Appendix

First, we present a lemma that is an extension of Lemma A2 of Chao et al. (2012) to our dependence structure of the error terms.

Lemma. Suppose that Assumptions (a)-(e) hold. Then

$$\begin{aligned} S_{n,T} &= \frac{1}{\sqrt{nT}} \sum_{i \neq j} \sum_{s,t=1}^{T} vech(\Pi_T z_{i,s} P_{i,j,s,t} v_{j,t} + v_{i,s} P_{i,j,s,t} z'_{j,t} \Pi'_T + v_{i,s} P_{i,j,s,t} v'_{j,t}) \\ &\xrightarrow{d} N(0, \Sigma). \end{aligned} \tag{37}$$

Proof of Lemma. For notational simplicity, and suppose that f_t and $v_{i,t}$ are scalar and that we focus on the $(1,1)$ element of $S_{n,T}$. The desired result follows from applying the Cramér-Wold device to this scalar result.

Because $P = I_T \otimes (1/n)\ell_n \ell'_n$ is block diagonal and the $(1,1)$ element of $\Pi_T z_{i,s}$ is f_s, we can write the (1,1) element of $S_{n,T}$ as

$$S_{n,T} = \frac{1}{n\sqrt{nT}} \sum_{i \neq j} \sum_{t=1}^{T} (f_t v_{i,t} + f_t v_{j,t} + v_{i,t} v_{j,t}), \tag{38}$$

with some notational abuse. Let

$$s_t = \frac{1}{n\sqrt{nT}} \sum_{i \neq j} (f_t v_{i,t} + f_t v_{j,t} + v_{i,t} v_{j,t}). \tag{39}$$

Then we can rewrite $S_{n,T}$ as

$$S_{n,T} = \sum_{t=1}^{T} s_t. \tag{40}$$

Let $\mathcal{F}_t$ denote the sigma field generated by $\{f_{s+1},\ v_{1,s},\ v_{2,s}, \ldots, v_{n,s}\}_{s=1}^{t}$. Because $v_{i,s}$ and $v_{j,t}$ are independent for all i, j, s, t such that $i \neq j$ or $s \neq t$, $\{s_t\}$ is a martingale difference sequence relative to $\{\mathcal{F}_t\}$. If we show

$$\sum_{t=1}^{T} E\left[s_t^2 | \mathcal{F}_{t-1}, Z\right] \ \to \ \Sigma \ \ a.s., \tag{41}$$

$$\sum_{t=1}^{T} E[s_t^2 I(|s_t| \geq \epsilon) | Z] \ \to \ 0 \ \ a.s., \tag{42}$$

for each $\epsilon > 0$, the martingale central limit theorem (Theorem 35.12 of Billinsgley, 1995, p.476) gives

$$P(S_{n,T} \leq x|Z) \ \rightarrow \ \frac{1}{\sqrt{\Sigma}}\Phi\left(\frac{x}{\sqrt{\Sigma}}\right) \ \ a.s. \tag{43}$$

for all $x \in \Re$ conditional on Z. Note that, for $\epsilon > 0$:

$$\sup_{n,T} E[P(S_{n,T} \leq x|Z)^{1+\epsilon}] < \infty \tag{44}$$

because probabilities are bounded between zero and one. Given that (44) holds, it follows from the dominated convergence theorem (Corollary in Billingsley, 1995, p.338) and (43) that[7]

$$P(S_{n,T} \leq x) \ \rightarrow \ \frac{1}{\sqrt{\Sigma}}\Phi\left(\frac{x}{\sqrt{\Sigma}}\right), \tag{45}$$

completing the proof for the case in which $v_{i,t}$ is scalar.

It remains to show (41) and (42). Note that

$$\begin{aligned}
& E\left(s_t^2|Z\right) \\
= \ & \frac{1}{n^3 T}\sum_{i\neq j}\sum_{k\neq \ell}[E(f_t^2 v_{i,t}v_{k,t}|Z) + E(f_t^2 v_{i,t}v_{\ell,t}|Z) + E(f_t^2 v_{j,t}v_{k,t}|Z) + E(f_t^2 v_{j,t}v_{\ell,t}|Z) \\
& + E(v_{i,t}v_{j,t}v_{k,t}v_{\ell,t}|Z) + 2E(f_t^2 v_{i,t}v_{j,t}) + 2E(f_t v_{i,t}^2 v_{j,t}|Z) + 2E(f_t v_{i,t} v_{j,t}^2|Z)] \\
= \ & \frac{4(n-1)^2}{n^3 T}\sum_{i=1}^{n} E(f_t^2 v_{i,t}^2|Z) + \frac{1}{n^3 T}\sum_{i\neq j}\sum_{k\neq \ell} E(v_{i,t}v_{j,t}v_{k,t}v_{\ell,t}|Z)]. \qquad (46)
\end{aligned}$$

The first term on the right hand side of (46) can be written as

$$\frac{4(n-1)^2}{n^3 T}\sum_{i=1}^{n} E(f_t^2 v_{i,t}^2|Z) \ = \ \frac{4(n-1)^2}{n^3 T}\sigma_f^2\sum_{i=1}^{n}\sigma_{i,t}^2. \tag{47}$$

The summation over $i \neq j$ and $k \neq \ell$ in the second term on the right hand

[7]Note that the unconditional probability is the expected value of the conditional probability.

side of (46) can be split into summations over three sets of i, j, k, ℓ:

$$
\begin{aligned}
I_0 &= \{(i,j,k,\ell) : \ i,j,k,\ell \text{ are distint from each other}\}, && (48)\\
I_1 &= \{(i,j,k,\ell) : \text{ only one pair are identical among } i \neq j, k \neq \ell, \\
&\quad \text{i.e., } (i = k \text{ and } j \neq \ell) \text{ or } (i \neq k \text{ and } j = \ell) \text{ or } (i = \ell \text{ and } j \neq k) \text{ or } (i \neq \ell \text{ and } j = k)\}, && (49)\\
I_2 &= \{(i,j,k,\ell) : \text{ two pairs are identical among } i \neq j, k \neq \ell, \\
&\quad \text{i.e., } (i = k \text{ and } j = \ell) \text{ or } (i = \ell \text{ and } j = k)\}. && (50)
\end{aligned}
$$

The expectations in the second term on the right hand side of (46) are zero when i, j, k, ℓ are in I_0 and I_1 and thus can be written as

$$
\begin{aligned}
\frac{1}{n^3 T} \sum_{i \neq j} \sum_{k \neq \ell} E(v_{i,t} v_{j,t} v_{k,t} v_{\ell,t} | Z) &= \frac{1}{n^3 T} \sum_{i,j,k,\ell \in I_2} E(v_{i,t} v_{j,t} v_{k,t} v_{\ell,t} | Z) \\
&= \frac{2}{n^3 T} \sum_{i \neq j} \sigma_{i,t}^2 \sigma_{j,t}^2 \qquad (51)
\end{aligned}
$$

where $\sigma_{i,t}^2 = E(v_{i,t}^2 | Z)$. Because s_t is a martingale difference sequence relative to $\mathcal{F}_t$ conditional on Z, it follows from (46), (47), (51) and Assumption (e) that

$$
\begin{aligned}
\sum_{t=1}^{T} E\left(s_t^2 | \mathcal{F}_t, Z\right) &= \frac{4\sigma_f^2 (n-1)^2}{n^3 T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sigma_{i,t}^2 + \frac{2}{n^3 T} \sum_{i \neq j} \sum_{t=1}^{T} \sigma_{i,t}^2 \sigma_{j,t}^2 \\
&= \frac{4\sigma_f^2}{nT} \sum_{t=1}^{T} \sum_{i=1}^{n} \sigma_{i,t}^2 + O(n^{-1}) \\
&\to \Sigma \ \ a.s. \qquad (52)
\end{aligned}
$$

Thus the law of iterated expectations completes the proof of (41).

To prove (42), write $\sum_{t=1}^{T} E(s_t^4|Z)$ as

$$\begin{aligned}
\sum_{t=1}^{T} E(s_t^4|Z) &= \frac{1}{n^2T^2}\sum_{j_1,j_2,j_3,j_4=1}^{n} E(f_t^4 v_{j_1}v_{j_2}v_{j_3}v_{j_4}) \\
&\quad + \frac{4}{n^3T^2}\sum_{j_1,j_2,j_3,j_4,j_5,j_4\neq j_5} E(f_t^3 v_{j_1}v_{j_2}v_{j_3}v_{j_4}v_{j_5}) \\
&\quad + \frac{6}{n^4T^2}\sum_{j_1,j_2,j_3,j_4,j_5,j_6,j_3\neq j_4,j_5\neq j_6} E(f_t^2 v_{j_1}v_{j_2}v_{j_3}v_{j_4}v_{j_5}v_{j_6}) \\
&\quad + \frac{4}{n^5T^2}\sum_{j_1,j_2,j_3,j_4,j_5,j_6,j_7,j_2\neq j_3,j_4\neq j_5,j_6\neq j_7} E(f_t v_{j_1}v_{j_2}v_{j_3}v_{j_4}v_{j_5}v_{j_6}v_{j_7}) \\
&\quad + \frac{4}{n^6T^2}\sum_{j_1\neq j_2,j_3\neq j_4,j_5\neq j_6,j_7\neq j_8} E(v_{j_1}v_{j_2}v_{j_3}v_{j_4}v_{j_5}v_{j_6}v_{j_7}v_{j_8}) \\
&= E_1 + E_2 + E_3 + E_4 + E_5. \qquad (53)
\end{aligned}$$

It follows from the assumption of mutual independence that

$$\begin{aligned}
E_1 &= O\left(\frac{n}{n^2T^2}\right) + O\left(\frac{n^2}{n^2T^2}\right), \qquad (54)\\
E_2 &= O\left(\frac{n^2}{n^3T^2}\right), \qquad (55)\\
E_3 &= O\left(\frac{n^2}{n^4T^2}\right) + O\left(\frac{n^3}{n^4T^2}\right), \qquad (56)\\
E_4 &= O\left(\frac{n^2}{n^5T^2}\right) + O\left(\frac{n^3}{n^5T^2}\right), \qquad (57)\\
E_5 &= O\left(\frac{n^2}{n^6T^2}\right) + O\left(\frac{n^3}{n^6T^2}\right) + O\left(\frac{n^4}{n^6T^2}\right), \qquad (58)
\end{aligned}$$

almost surely. Thus,

$$\sum_{t=1}^{T} E(s_t^4|Z) = O(T^{-1}) \qquad (59)$$

almost surely from which (42) follows.

Proof of Proposition 1. Suppose that the rank of Π_T is K. Pre-multiplying each side of the reduced-form equation (13) by $G\times(G+K)$ matrix $c[-\beta'\ I_G]$

yields

$$c(y_{i,t} - \beta' x_{i,t}) \;=\; c[1 \;\; -\beta']\Pi_T z_{i,t} + c[-\beta' \;\; I_G]v_j, \tag{60}$$

which implies that $c[-\beta' \;\; I_G]\Pi_T = 0$ must hold. Because Π_T has rank K, $c[1 \;\; -\beta']$ is identified up to scale. Setting $c = 1$, β is identified.

Suppose that the model is correctly specified and that the parameter β is identified. Then the former implies $\Pi_{y,T} = \beta'\Pi_{x,T}$ while the latter implies that the rank of $\Pi_{x,T}$ is K. Therefore the rank of Π_T is K.

Proof of Theorem 1. Because

$$\begin{aligned}\frac{1}{\sqrt{nT}}\sum_{i\neq j}\sum_{s,t=1}^{T} Y_{i,s}P_{i,j,s,t}Y'_{j,t} \;&=\; \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T}\Pi_T z_{i,t}z'_{j,t}\Pi'_T + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T}\Pi_T z_{i,t}v'_{j,t}\\ &+\frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}z'_{j,t}\Pi'_T + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}v'_{j,t}\end{aligned} \tag{61}$$

we have

$$\begin{aligned}&\frac{1}{\sqrt{nT}}\sum_{i\neq j}\sum_{s,t=1}^{T} Y_{i,s}P_{i,j,s,t}Y'_{j,t} - \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T}\Pi_T z_{i,t}z'_{j,t}\Pi'_T\\ =\;& \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T}\Pi_T z_{i,t}v'_{j,t} + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}z'_{j,t}\Pi'_T + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}v'_{j,t}\\ =\;& \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} f_t v'_{j,t} + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}f'_t + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}v'_{j,t}\\ =\;& \frac{1}{\sqrt{nT}}\sum_{t=1}^{T} f_t\sum_{j=1}^{n} v'_{j,t} + \frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\sum_{i=1}^{n} v_{i,t}f'_t + \frac{1}{n\sqrt{nT}}\sum_{i\neq j}\sum_{t=1}^{T} v_{i,t}v'_{j,t}\\ =\;& I + II + III.\end{aligned} \tag{62}$$

The variance-covariance matrix of I is

$$\frac{1}{nT}\sum_{s,t=1}^{T}\sum_{i,j=1}^{n} E[\mathrm{vech}(f_s v'_{i,s})\mathrm{vech}(f_t v'_{j,t})'] \;=\; \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n} E[\mathrm{vech}(f_t v'_{i,t})\mathrm{vech}(f_t v'_{i,t})']. \tag{63}$$

Similarly, the covariance matrix between I and II and the variance covariance matrix of II are given by

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}E[\text{vech}(f_t v_{i,t}')\text{vech}(v_{i,t}f_t')'], \tag{64}$$

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{j=1}^{n}E[\text{vech}(v_{j,t}f_t')\text{vech}(v_{j,t}f_t')'], \tag{65}$$

respectively. I and III are uncorrelated because

$$\begin{aligned}\frac{1}{nT}\sum_{s,t=1}^{T}\sum_{i=1}^{n}\sum_{j\neq k}E[\text{vech}(f_s v_{i,s}')\text{vech}(v_{j,t}v_{k,t}')'] &= \frac{1}{nT}\sum_{t=1}^{T}\sum_{j\neq k}E[\text{vech}(f_t v_{j,t}')\text{vech}(v_{j,t}v_{k,t}')'] \\ &\quad +\frac{1}{nT}\sum_{t=1}^{T}\sum_{j\neq k}E[\text{vech}(f_t v_{k,t}')\text{vech}(v_{j,t}v_{k,t}')'] \\ &= 0 \end{aligned} \tag{66}$$

Analogously, II and III are uncorrelated. III is $O_p(n^{-1})$. Therefore it follows from our lemma that

$$vech(I+II+III) \overset{d}{\to} N(0,\Sigma). \tag{67}$$

Thus, it follows from (62)—(67) that

$$\frac{1}{\sqrt{NT}}\sum_{i\neq j}vech(v_i P_{ij} v_j') \overset{d}{\to} N(0,\Sigma). \tag{68}$$

Proof of Corollary 1. The corollary follows from Theorem 1 because

$$\begin{aligned} vech(C\hat{H}C') &= D_n^+ vec(C\hat{H}C') = D_n^+ vec(C\hat{H}C') \\ &= D_n^+(C\otimes C)vec(\hat{H}) = D_n^+(C\times C)D_n^+ vech(\hat{H}). \end{aligned} \tag{69}$$

Proof of Proposition 2.

To simplify the notation, we assume that f_t and $v_{i,t}$ scalar. Because $P_{i,j,t,t} =$

$1/n$, $\widehat{\Sigma}$ can be written as

$$\widehat{\Sigma} = \frac{2}{n^2 T} \sum_{t=1}^{T} \left[\sum_{i \neq j} (f_t + v_{i,t})^2 (f_t + v_{j,t})^2 - \left(\frac{1}{n} \sum_{i \neq j} (f_t + v_{i,t})(f_t + v_{j,t}) \right)^2 \right]. \tag{70}$$

Repeating the martingale argument in the proof of the lemma, it can be shown that

$$\begin{aligned}
\widehat{\Sigma} &= \frac{2}{n^2 T} \sum_{t=1}^{T} \sum_{i \neq j} f_t^2 (v_{i,t}^2 + v_{j,t}^2) + o_p(1) \\
&= \frac{2(n-1)}{n^2 T} \sum_{t=1}^{T} \sum_{i=1}^{n} f_t^2 v_{i,t}^2 + \frac{2(n-1)}{n^2 T} \sum_{t=1}^{T} \sum_{j=1}^{n} f_t^2 v_{j,t}^2 + o_p(1) \\
&= 2\sigma_f^2 \frac{1}{n} \sum_{i=1}^{n} \sigma_{i,t}^2 + 2\sigma_f^2 \frac{1}{n} \sum_{j=1}^{n} \sigma_{j,t}^2 + o_p(1) \\
&= 4\sigma_f^2 \frac{1}{n} \sum_{i=1}^{n} \sigma_{i,t}^2 + o_p(1).
\end{aligned} \tag{71}$$

Thus, $\widehat{\Sigma}$ is consistent for Σ.

References

1. Anatolyev, Stanislav, and Nikolay Gospodinov (2011), "Specification Testing in Models with Many Instruments," *Econometric Theory,* 27, 427–441.

2. Ang, Andrew, Jun Liu, and Krista Schwarz (2009), "Using Individual Stocks or Portfolios in Tests of Factor Models", *SSRN Working Paper*, http://ssrn.com/abstract=1106463.

3. Angrist, Joshua D. (1988), "Grouped Data Estimation and Testing in Simple Labor Supply Models," *Journal of Econometrics,* 47, 243–266.

4. Bekker, Paul A. (1994), "Alternative Approximations to the Distributions of Instrumental Variables Estimators," *Econometrica,* 62, 657–681.

5. Billingsley, Patrick (1995), *Probability and Measures,* Third Edition, John Wiley & Sons: New York, NY.

6. Chao, John C., Norman R. Swanson, Jerry A. Hausman, Whitney K. Newey and Tiemen Woutersen (2012), "Asymptotic Distribution of JIVE in a Heteroskedastic IV Regression with Many Instruments," *Econometric Theory,* 28, 42–86.

7. Chao, John C., Jerry A. Hausman, Whitney K. Newey, Norman R. Swanson and Tiemen Woutersen (2014), "Testing Overidentifying Restrictions with Many Instruments and Heteroskedasticity," *Journal of Econometrics,* 178, 15–21.

8. Cochrane, John H. (2005), *Asset Pricing*, Princeton University Press.

9. Deaton, Angus (1985), "Panel Data from a Time Series of Cross-Sections," *Journal of Econometrics,* 30, 109–126.

10. Donald, Stephen G., Natércia Fortuna and Vladas Pipiras (2007), "On Rank Estimation in Symmetric Matrices: The Case of Indefinite Matrix Estimators," *Econometric Theory,* 23, 1217–1232.

11. Fama, Eugene F. and Kenneth R. French (1992), "The Cross-Section of Expected Stock Returns," *Journal of Finance* 47:2, 427–65.

12. Fama, Eugene F. and Kenneth R. French (2004), "The Capital Asset Pricing Model: Theory and Evidence", *Journal of Economic Perspectives* 18(3), 25–46.

13. Gospodinov, Nikolay, Raymond Kan and Cesare Robotti (2013), "Chi-squared Tests for Evaluation and Comparison of Asset Pricing Models", *Journal of Econometrics* 173, 108–125.

14. Kandel, S. and R.F. Stambaugh (1995), "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Finance* 50, 157-184.

15. Kleibergen, Frank, and Richard Paap, (2006), "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics,* 133, 97–126.

16. Kleibergen, Frank, and Zhaoguo Zhan (2013), "Unexplained Factors and Their Effects on Second Pass R-squared's and t-tests", *mimeo*, Brown and Tsinghua University.

17. Koopmans, Tjalling C., and William C. Hood, (1953), "The Estimation of Simultaneous Linear Economic Relationships," in W.C. Hood and T.C. Koopmans, eds., *Studies in Econometric Methods,* Yale University Press: New Haven, CT, Chapter 6.

18. Kunitomo, Naoto (1980), "Asymptotic Expansions of Distributions of Estimators in a Linear Functional Relationship and Simultaneous Equations," *Journal of the American Statistical Associations,* 75, 693–700.

19. Lettau, Martin, and Sydney Ludvigson (2001), "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance* 56, 815–849.

20. Magnus, Jan R., and Heinz Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics,* Revised Edition, John Wiley & Sons: Chichester, England.

21. Morimune, Kimio (1983), "Approximate Distributions of the k-class Estimators When the Degree of Overidentifiability is Large Compared with the Sample Size," *Econometrica,* 51, 821–841.

22. Verdelhan, Adrien (2012), "The Share of Systematic Variation in Bilateral Exchange Rates", *SSRN Working Paper*, http://ssrn.com/abstract=1930516.

Table 1. Empirical Rejection Frequencies
When the Parameter is Identified

G	5	5	5	5	10	10	10	10
n	100	100	200	200	100	100	200	200
T	100	200	100	200	100	200	100	200
Null Hypothesis								
Under-identified	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Correctly Specified	0.008	0.005	0.005	0.006	0.006	0.002	0.006	0.005

Table 2. Empirical Rejection Frequencies
When the Portfolio-Specific Mean is Neglected

G	5	5	5	5	10	10	10	10
n	100	100	200	200	100	100	200	200
T	100	200	100	200	100	200	100	200
Null Hypothesis								
Under-identified	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Correctly specified	0.464	0.750	0.945	0.998	0.379	0.623	0.889	0.998

Table 3. Empirical Rejection Frequencies
When There is No Factor Structure

G	5	5	5	5	10	10	10	10
n	100	100	200	200	100	100	200	200
T	100	200	100	200	100	200	100	200
Null Hypothesis								
Under-identified	0.059	0.056	0.063	0.042	0.077	0.073	0.075	0.053
Correctly specified	0.004	0.003	0.004	0.001	0.005	0.001	0.001	0.001

Notes to the tables: The tables reports empirical rejection frequencies of our tests in data generating processes calibrated on the CAPM models at the 5% significance level.