Evaluating Latent and Observed Factors in Macroeconomics and Finance

Jushan Bai* Serena Ng †

August 2004

Abstract

Common factors play an important role in many disciplines of social science. In economics, the factors are the common shocks that underlie the co-movements of the large number of economic time series. The question of interest is whether some observable economic variables are in fact the underlying unobserved factors. We consider statistics to determine if the observed and the latent factors are exactly the same. We also provide simple to construct statistics that indicate the extent to which the two sets of factors differ. The key to the analysis is that the space spanned by the latent factors can be consistently estimated when the sample size is large in both the cross-section and the time series dimensions. The tests are used to assess how well the so-called Fama and French factors as well as several business cycle indicators approximate the factors in portfolio and individual stock returns. Data from a large panel of macroeconomic are also analyzed.

*Department of Economics, 269 Mercer St, NYU, New York, NY 10022 Email: Jushan.Bai@nyu.edu.

†Department of Economics, University of Michigan, Ann Arbor, MI 48109 Email: Serena.Ng@umich.edu
This paper was presented at Yale University, the 2003 IGIER-PIER Conference on Macroeconomics and Finance held at Bocconi University, and the 2003 ISI Meeting in Berlin. We thank the discussants Rob Engle, Ursula Gather, Domenico Giannone, the seminar participants, Sydney Ludvigson, two anonymous referees and the editor for many helpful suggestions and comments. The authors acknowledge financial support from the NSF (grants SBR-9896329, SES-0137084, SES-0136923).

1 Introduction

Many economic theories have found it useful to explain the behavior of the observed data by means of a small number of fundamental factors. For example, the arbitrage pricing theory (APT) of Ross (1976) is built upon the existence of a set of common factors underlying all asset returns. In the capital asset pricing theory (CAPM) of Sharpe (1964), Lintner (1965), and Merton (1973), the 'market return' is the common risk factor that has pervasive effects on all assets. In consumption based CCAPM models of Breeden (1989) and Lucas (1978), aggregate consumption is the source of systematic risk in one-good exchange economies. More recently, Lettau and Ludvigson (2001) suggest that the deviation between consumption and wealth is a fundamental factor. Because interest rate of different maturities are highly correlated, most models of interest rates have a factor structure. Indeed, Stambaugh (1988) derived the conditions under which an affine-yield model implies a latent-variable structure for bond returns. A fundamental characteristic of business cycles is the comovement of a large number of series, which is possible only if economic fluctuations are driven by common sources.

While the development of theory can proceed without a complete specification of how many and what the factors are, empirical testing does not have this luxury. To test the multifactor APT, one has to specify the empirical counterpart of the theoretical risk factors. To test the CAPM, one must specify what is the 'market', and to test the CCAPM, one must specifically define consumption. To test term structure models, the common practice is to identify instruments that are correlated with the underlying state variables, which amounts to finding proxies for the latent factors. To be able to asses the sources of business cycle fluctuations, one has to specify the candidates for the primitive shocks.

A small number of applications have proceeded by replacing the unobserved factors with statistically estimated ones. For example, Lehmann and Modest (1988) used factor analysis, while Connor and Korajzcyk (1998) adopted the method of principal components. The drawback is that the statistical factors do not have immediate economic interpretation. A more popular approach is to rely on intuition and theory as guide to come up with a list of observed variables as proxies of the unobserved theoretical factors. Variables such as the unemployment rate in excess of the natural rate and the deviation of output from its potential are both popular candidates for the state of the economy. In term structure analysis, the term spread and the short rate have all been used as proxies of the factors underlying bond yields. In the CAPM analysis, the equal-weighted and value-weighted market returns are often used in place of the theoretical market return. In the CCAPM analysis, non-durable consumption is frequently used as the systematic risk factor. By regressing asset returns on a large number of financial and macroeconomic variables and analyzing their explanatory power, Chen et al. (1986) found that the factors in APT are related to expected and unexpected inflation, interest rate risk, term structure risk, and industrial production. Perhaps the most well-known of observable risk factors are the three discussed in Fama and French (1993):- the market excess return, the small minus big factor, and the high minus low factor. Hereafter, we will simply refer to these as the Fama and French (FF) factors.

There is a certain appeal in associating the latent factors with observed variables as this facilitates economic interpretation. But as pointed out by Shanken (1992), estimation of beta pricing relations using proxy factors are meaningful only if the fundamental factors are spanned by the observed factors. Yet such a condition will be violated even if a pure measurement error is added to a perfect proxy. To date, there does not exist a formal test for the adequacy of the observed variables as proxies for the unobserved factors. The problem is not so much that the fundamental factors are unobserved because in principle, if we observe indicators of the factors, we can estimate the factors from the data. The problem is that latent factors estimated from a small number of indicators are imprecise, and in theory, consistent estimation of the latent factors cannot be achieved under the traditional assumption that T is large and N is fixed, or vice versa.

In this paper, we develop several procedures to compare the (individual or set of) observed variables with estimates of the unobserved factors. The point of departure is that we work with large dimensional panels. That is, we deal with datasets that have a large number of cross-section units (N) and a large number of time series observations (T). By allowing both N and T to tend to infinity, the space spanned by the common factors can be estimated consistently. Our analysis thus combines the statistical approach of Lehmann and Modest (1988) and Connor and Korajzcyk (1998), with the economic approach of using observed variables as proxies. We begin in Section 2 by considering estimation of the factors, stating the conditions under which the estimated factors can be treated as though they are known. Section 3 presents tests to compare the observed variables with the estimated factors, making precise how sampling variability of the factor estimates is reflected in the test statistics. Monte carlo simulations are reported In Section 4. We then use the procedures to compare observed variables with factors estimated from portfolio returns, individual stock returns, and a large set of macroeconomic time series. These are reported in Section 5. Proofs are given in the Appendix.

2 Preliminaries

Consider the factor representation for a panel of data x_{it}, $(i = 1, \dots N, t = 1, \dots T)$

$$x_{it} = \lambda_i' F_t + e_{it},$$

where F_t $(r \times 1)$ is the factor process, and λ_i $(r \times 1)$ is the factor loading for unit i. In classical factor analysis, the number of units, N, is fixed and the number of observations, T, tends to infinity. With macroeconomic and financial applications, this assumption is not fully satisfactory because data for a large number of cross-section units are often available over a long time span, and in some cases, N can be much larger than T (and vice versa). For example, daily data on returns for well over one hundred stocks are available since 1960, and a large number of price and interest rate series are available for over forty years on a monthly basis. Classical analysis also assumes e_{it} is iid over t and independent over i. This implies a diagonal variance covariance matrix for $e_t = (e_{1t}, ..., e_{Nt})$, an assumption that is rather restrictive. To overcome these two limitations, we work with high dimensional approximate factor models that allow both N and T to tend to infinity, and in which e_{it} may be serially and cross-sectionally correlated so that the covariance matrix of e_t does not have to be a diagonal matrix. In fact, all that is needed is that the largest eigenvalue of the covariance matrix of e_{it} is bounded as N tends to infinity.

Suppose we observe G_t, an $(m \times 1)$ vector of economic variables. We are ultimately interested in the relationship between G_t and F_t, but we do not observe F_t. It would seem natural to proceed by regressing x_{it} on G_t, and then use some metric to assess the explanatory power of G_t[1]. The idea would be that if G_t is a good proxy for F_t, it should explain x_{it}. This, however, is not a satisfactory test because even if G_t equals F_t exactly, G_t might still only be weakly correlated with x_{it} if the variance of the idiosyncratic error e_{it} is large. In other words, a low explanatory power of G_t for x_{it} by itself may not be the proper criterion for deciding if G_t corresponds to the true factors.

In the CAPM analysis, several approaches have been used to check if inference is sensitive to the use of a proxy in place of the market portfolio. Stambaugh (1982) considered returns from a number of broadly defined markets and conclude that inference is not sensitive to the choice of proxies. This does not, however, suggest that the unobservability of the market portfolio has no implication for inference, an issue raised by Roll (1977). Another approach, used in Kandel and Stambaugh (1987) and Shanken (1987), is to obtain the cut-off correlation

[1] Such an approach was adopted in Chen et al. (1986), for example

between the proxy for market return and the true market return that would change the conclusion on the hypothesis being tested. They find that if the correlation between G_t and F_t is at or above .7, inference will remain intact. However, this begs the question of whether the correlation between the unobserved market return and the proxy variable is high. While these approaches provide far more cautions inference, the basic problem remains that we do not know how correlated are F_t and G_t. To be able to test F_t and G_t directly, we must first confront the problem that F_t is not observed.

We use the method of principal components to estimate the factors. Throughout, we use 'tilde' to denote the principal components estimates. Hatted variables are based on least squares regressions with the principal components estimates as regressors. Let X be the T by N matrix of observations such that the ith column is the time series of the ith cross section. Let $\widetilde{V}$ be a $r \times r$ diagonal matrix consisting of the r largest eigenvalues of XX'/NT. Let $\widetilde{F} = (\widetilde{F}_1, ..., \widetilde{F}_T)'$ be the principal component estimates of F under the normalization that $\frac{F'F}{T} = I_r$. Then $\widetilde{F}$ is comprised of the r eigenvectors (multiplied by $\sqrt{T}$) associated with the r largest eigenvalues of the matrix $XX'/(NT)$ in decreasing order. Let $\Lambda = (\lambda_1, ..., \lambda_N)'$ be the matrix of factor loadings. The principal components estimator of Λ is $\widetilde{\Lambda} = X'\widetilde{F}/T$. By definition $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i'\widetilde{F}_t$.

Denote the norm of a matrix A by $\|A\| = [tr(A'A)]^{1/2}$. The notation M stands for a finite positive constant, not depending on N and T. The following assumptions are needed for consistent estimation of the factors by the method of principal components.

Assumptions: *Assumption A: Common factors*

1. $E\|F_t\|^4 \le M$ and $\frac{1}{T}\sum_{t=1}^T F_t F_t' \xrightarrow{p} \Sigma_F$ for a $r \times r$ positive definite matrix Σ_F.

Assumption B: Heterogeneous factor loadings

The loading λ_i is either deterministic such that $\|\lambda_i\| \le M$ or it is stochastic such that $E\|\lambda_i\|^4 \le M$. In either case, $\Lambda'\Lambda/N \xrightarrow{p} \Sigma_\Lambda$ as $N \to \infty$ for some $r \times r$ positive definite non-random matrix Σ_Λ.

Assumption C: Time and cross-section dependence and heteroskedasticity

1. $E(e_{it}) = 0$, $E|e_{it}|^8 \le M$;

2. $E(e_{it}e_{js}) = \tau_{ij,ts}$, $|\tau_{ij,ts}| \le \tau_{ij}$ for all (t,s) and $|\tau_{ij,ts}| \le \gamma_{ts}$ for all (i,j) such that
$$\frac{1}{N}\sum_{i,j=1}^N \tau_{ij} \le M, \ \frac{1}{T}\sum_{t,s=1}^T \gamma_{ts} \le M, \text{ and } \frac{1}{NT}\sum_{i,j,t,s=1} |\tau_{ij,ts}| \le M$$

3. For every (t,s), $E|N^{-1/2}\sum_{i=1}^N \Big[e_{is}e_{it} - E(e_{is}e_{it})\Big]|^4 \le M$.

Assumption D: $\{\lambda_i\}, \{F_t\}$, and $\{e_{it}\}$ are three groups of mutually independent stochastic variables.

Assumptions A and B together imply r common factors. Assumption C allows for limited time series and cross section dependence in the idiosyncratic component. Heteroskedasticity in both the time and cross section dimensions is also allowed. Under stationarity in the time dimension, $\gamma_N(s,t) = \gamma_N(s-t)$, though the condition is not necessary. Given Assumption C1, the remaining assumptions in C are easily satisfied if the e_{it} are independent for all i and t. The allowance for weak cross-section correlation in the idiosyncratic components leads to the *approximate factor structure* of Chamberlain and Rothschild (1983). It is more general than a *strict factor model* which assumes e_{it} is uncorrelated across i. Assumption D is standard in factor analysis.

As is well known, the factor model is fundamentally unidentified because $\lambda_i' L L^{-1} F_t = \lambda_i' F_t$ for any invertible matrix L. In economics, exact identification of the factors, F_t, may not always be necessary. If the estimated F_t is used for forecasting as in Stock and Watson (2002), the distinction between F_t and LF_t is immaterial because they will give the same forecast. When stationarity or the cointegrating rank of F_t is of interest, knowing LF_t is sufficient, as F_t has the same cointegrating rank as LF_t. In these situations as well as addressing the question we are interested in, namely, determining if F_t is close to G_t, what is important is consistent estimation of the space spanned by the factors.

Lemma 1 *Let* $H = \widetilde{V}^{-1}(\widetilde{F}'F/T)(\Lambda'\Lambda/N)$. *Under Assumptions A-D, and as* $N, T \to \infty$,

i $\min[N,T]\Big(\frac{1}{T}\sum_{t=1}^T \|\widetilde{F}_t - HF_t\|^2\Big) = O_p(1)$;

ii If $\sqrt{N}/T \to 0$, $\sqrt{N}(\widetilde{F}_t - HF_t) \xrightarrow{d} V^{-1}QN(0,\Gamma_t)$, *where* $\widetilde{F}'F/T \xrightarrow{p} Q$, $\widetilde{V} \xrightarrow{p} V$, *and* $\Gamma_t = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \sum_{j=1}^N E(\lambda_i \lambda_j' e_{it} e_{jt})$.

Part (i), shown in Bai and Ng (2002), establishes that the squared difference between the estimated and the scaled true factors vanish as N and T tend to infinity. As shown in Bai (2003), $\sqrt{N}(\widetilde{F}_t - HF_t) = \widetilde{V}^{-1}(\widetilde{F}'F/T)\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i e_{it} + o_p(1)$, from which the sampling distribution of the factor estimates given in (ii) follows. These limiting distributions are asymptotically independent across t if e_{it} is serially uncorrelated. The tests to be developed are built upon Lemma 1 and the fact that Γ_t is consistently estimable.

3 Comparing the Estimated and the Observed Factors

We observe G_t, and want to know if its m elements are generated by (or is a linear combination of) the r latent factors, F_t. In general, r is an unknown parameter. Consider estimating r using one of the two panel information criterion:

$$\begin{aligned} \widehat{r} &= \underset{k=0,\dots kmax}{\text{argmax}}\; PCP(k) \quad \text{where} \quad PCP(k) = \widetilde{\sigma}^2(k) + \widetilde{\sigma}^2(kmax)k \cdot g(N,T), \\ \widehat{r} &= \underset{k=0,\dots kmax}{\text{argmax}}\; ICP(k) \quad \text{where} \quad ICP(k) = \log \widetilde{\sigma}^2(k) + k \cdot g(N,T), \end{aligned}$$

where $\widetilde{\sigma}^2(k) = \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \widetilde{e}_{it}^2$, $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i' \widetilde{F}_t$, with $\widetilde{\lambda}_i$ and $\widetilde{F}_t$ estimated by the method of principal components. In Bai and Ng (2002), we showed that $\text{prob}(\widehat{r} = r) \to 1$ as $N, T \to \infty$ if $g(N,T)$ is chosen such that $g(N,T) \to 0$ and $\min[N,T]g(N,T) \to \infty$. Because r can be consistently estimated, in what follows, we simply treat it as though it is known.

Obviously, if $m < r$, the m observed variables cannot span the space of the r latent factors. Testing for the adequacy of G as a set is meaningful only if $m \geq r$. Nonetheless, regardless of the dimension of G_t and F_t, it is still of interest to know if a particular G_t is in fact a fundamental factor. Section 3.1 therefore begins with testing the observed variables one by one. Section 3.2 considers testing the observed variables as a set.

3.1 Testing G_t One at a Time

Let G_{jt} be an element of the m vector G_t. The null hypothesis is that G_{jt} is an exact factor, or more precisely, that there exists a δ_j such that $G_{jt} = \delta_j' F_t$ for all t. Consider the regression $G_{jt} = \gamma_j' \widetilde{F}_t + error$. Let $\widehat{\gamma}_j$ be the least squares estimate of γ_j and let $\widehat{G}_{jt} = \widehat{\gamma}_j' \widetilde{F}_t$. Consider the t-statistic

$$\tau_t(j) = \frac{(\widehat{G}_{jt} - G_{jt})}{\left(\text{var}(\widehat{G}_{jt})\right)^{1/2}}. \tag{1}$$

Let Φ_α^τ be the α percentage point of the limiting distribution of $\tau_t(j)$. Then $\mathcal{A}(j) = \frac{1}{T}\sum_{t=1}^T 1(|\tau_t(j)| > \Phi_\alpha^\tau)$ is the frequency that $\tau_t(j)$ exceeds the α percent critical value in a sample of size T.

The $\mathcal{A}(j)$ statistic allows G_{jt} to deviate from $\widehat{G}_{jt}$ for a pre-specified number of time points as specified by α. A stronger test is to require G_{jt} not to deviate from $\widehat{G}_{jt}$ by more than sampling error at *every* t. To this end, we also consider the statistic $\mathcal{M}(j) = \max_{1 \leq t \leq T} |\tau_t(j)|$. The $\mathcal{M}(j)$ statistic is a test of how far is the $\widehat{G}_{jt}$ curve from G_{jt}. It is a stronger test than $A(j)$ which tests G_{jt} point by point.

Proposition 1 *(exact tests) Let G_t be a vector of m observed factors, and F_t be a vector of r latent factors. Let $\widehat{\tau}_t(j)$ be obtained with $var(\widehat{G}_{jt})$ replaced by its consistent estimate, $\widehat{var}(\widehat{G}_{jt})$. Consider the statistics:*

$$A(j) = \frac{1}{T}\sum_{t=1}^{T} 1(|\widehat{\tau}_t(j)| > \Phi_\alpha) \tag{2}$$

$$M(j) = \max_{1\le t\le T} |\widehat{\tau}_t(j)|. \tag{3}$$

Under the null hypothesis that $G_{jt} = \delta' F_t$ and as $N,T \to \infty$ with $\sqrt{N}/T \to 0$, then (i) $A(j) \xrightarrow{p} 2\alpha$. If, in addition, e_{it} is serially uncorrelated, then (ii) $P(M(j) \le x) \approx [2\Phi(x) - 1]^T$, where $\Phi(x)$ is the cdf of a standard normal random variable.

Under the null hypothesis that $G_{jt} = \delta_j' F_t$ and that δ_j is time invariant, the rate of convergence of $\widehat{G}_{jt}$ to G_{jt} is $\sqrt{N}$ and the distribution is asymptotically normal. To see this, rewrite $G_{jt} = \gamma_j' H F_t$ with $\gamma_j' = \delta_j' H^{-1}$. Adding and subtracting, $\sqrt{N}(\widehat{G}_{jt} - G_{jt}) = \sqrt{N}(\widehat{\gamma}_j' - \gamma_j')HF_t + \widehat{\gamma}_j'\sqrt{N}(\widetilde{F}_t - HF_t)$. The first term is shown to be $o_p(1)$ in the appendix under the null hypothesis. So the second term dictates the limiting distribution of $\sqrt{N}(\widehat{G}_{jt} - G_{jt})$, but this is asymptotically normal by Lemma 1 (ii). Thus $\tau_t(j)$, the standardized $\widehat{G}_{jt} - G_{jt}$, has a standard normal limiting distribution. Part (i) of the proposition permits e_{it} to be serially correlated, since the law of large numbers holds even if $\tau_t(j)$ is time dependent. In contrast, part (ii) of the proposition requires e_{it} to be serially uncorrelated, so that $\tau_t(j)$ is asymptotically uncorrelated and thus independent by normality.[2] The stated property of the $M(j)$ statistic then follows from consideration of the maximum of iid normal variates.

Let $Avar(\widehat{G}_{jt})$ denote the asymptotic variance of $\sqrt{N}(\widehat{G}_{jt} - G_{jt})$. It follows that

$$Avar(\widehat{G}_{jt}) = \text{plim } \widehat{\gamma}_j' Avar(\widetilde{F}_t)\widehat{\gamma}_j = \text{plim } \widehat{\gamma}_j'\widetilde{V}^{-1}(\frac{\widetilde{F}'F}{T})\Gamma_t(\frac{\widetilde{F}'F}{T})\widetilde{V}^{-1}\widehat{\gamma}_j,$$

with $\Gamma_t = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} E(\lambda_i\lambda_j' e_{it}e_{jt})$. Let $\text{var}(\widehat{G}_{jt})$ denote the asymptotic variance divided by N. That is,

$$\text{var}(\widehat{G}_{jt}) = \frac{1}{N} Avar(\widehat{G}_{jt}).$$

An estimate of it can be obtained by first substituting $\widetilde{F}$ for F, and noting that $\widetilde{F}'\widetilde{F}/T$ is an r-dimensional identity matrix by construction. Thus,

$$\widehat{\text{var}}(\widehat{G}_{jt}) = \frac{1}{N}\widehat{\gamma}_j'\widetilde{V}^{-1}\widetilde{\Gamma}_t\widetilde{V}^{-1}\widehat{\gamma}_j,$$

[2]Extreme value distribution can also be used to approximate the distribution of $M(j)$ after proper centering and scaling when e_{it} is serially correlated.

where $\widetilde{\Gamma}_t$ is a consistent estimate of $H'^{-1}\Gamma_t H^{-1}$. For each fixed t, unless e_{it} is cross-sectionally uncorrelated, Γ_t (more precisely $H'^{-1}\Gamma_t H^{-1}$) is not consistently estimable. The problem is akin to the problem in time series that summing the T autocovariances will not yield a consistent estimate of the spectrum. With cross-section data, the data have no natural ordering, and the time series solution of truncation is neither intuitive nor possible without a precise definition of distance between observations. However, if Γ_t does not depend on t, that is, if e_{it} is stationary so that $E(e_{it}^2) = \sigma_i^2$ for all t, the following estimator is consistent for $H'^{-1}\Gamma_t H^{-1}$.

$$\widetilde{\Gamma}_t = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\widetilde{\lambda}_i\widetilde{\lambda}_j'\frac{1}{T}\sum_{t=1}^{T}\widetilde{e}_{it}\widetilde{e}_{jt} \quad \forall\, t, \tag{4}$$

where $\frac{n}{\min[N,T]} \to 0$. This covariance estimator is robust to cross-correlation, and accordingly, we refer to it as CS-HAC. It makes use of repeated observations over time, so that under covariance stationarity, the time series observations can be used to estimate the cross-section correlation matrix. However, because of error from estimating of the factors, we can only use $n < N$ observations when estimating Γ_t. A formal analysis of the CS-HAC is given in Bai and Ng (2004).

When e_{it} is cross-sectionally uncorrelated, an alternative estimator is

$$\widetilde{\Gamma}_t = \frac{1}{N}\sum_{i=1}^{N}\widetilde{e}_{it}^2\widetilde{\lambda}_i\widetilde{\lambda}_i'. \tag{5}$$

The above estimator does not require time-series homoskedasticity of e_{it}. In the absence of cross-section correlation, time series heteroskedasticity of e_{it} is permitted. If, in addition, $E(e_{it}^2) = \sigma_e^2$ for all i and all t, an appropriate estimator is

$$\widetilde{\Gamma} = \widehat{\sigma}_e^2\frac{\widetilde{\Lambda}'\widetilde{\Lambda}}{N}, \tag{6}$$

where $\widehat{\sigma}_e^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{e}_{it}^2$, and $\widetilde{\Lambda}$ is the $N \times r$ matrix of estimated factor loadings. In summary, equation (4) allows cross-section correlation but assumes time-series stationarity of e_{it}. Equation (5) allows time-series heteroskedasticity, but assumes no cross-section correlation. Equation (6) assumes no cross-section correlation and constant variance for all i and all t.

Once an appropriate $\widetilde{\Gamma}_t$ is chosen to construct $\widehat{\text{var}}(\widehat{G}_{jt})$, testing if G_{jt} is an exact factor is then quite simple. For example, if $\alpha = 0.025$, the fraction of $\tau_t(j)$ that exceeds 1.96 in absolute value should be close to 5% for large N and T. Thus, $A(j)$ should be close to .05.

Furthermore, $M(j)$ should not exceed the maximum of a vector of $N(0,1)$ random variables (in absolute values) of length T. This maximum value increases with T, but can be easily tabulated by simulations or theoretical computations. The 1, 5 and 10 % critical values are as follows:

Critical Values for the $M(j)$ test

	T			
	50	100	200	400
.01	3.775	3.935	4.109	4.219
.05	3.283	3.467	3.656	3.830
.10	3.076	3.278	3.475	3.632

Requiring that G_t be an exact linear combination of the latent factors is rather strong. An observed series might match the variations of the latent factors very closely, and yet is not an exact factor in a mathematical sense. Measurement error and time aggregation, for example, could be responsible for deviations between the observed variables and the latent factors, as discussed in Breeden et al. (1989). In such cases, it would be useful to gauge how far are the proxies from the true factors.

Proposition 2 *(approximate tests) Suppose* $G_{jt} = \delta_j' F_t + \varepsilon_{jt}$, *with* $\varepsilon_{jt} \sim (0, \sigma_\varepsilon^2(j))$. *Let* $\widehat{G}_{jt} = \widehat{\gamma}_j' \widetilde{F}_t$, *where* $\widehat{\gamma}_j$ *is obtained by least squares from a regression of* G_{jt} *on* $\widetilde{F}_t$. *Let* $\widehat{\varepsilon}_{jt} = G_{jt} - \widehat{G}_{jt}$. *Then as* $N, T \to \infty$, *(i)* $\widehat{\sigma}_\varepsilon^2(j) = \frac{1}{T}\sum_{t=1}^T \widehat{\varepsilon}_{jt}^2 \xrightarrow{p} \sigma_\varepsilon^2(j)$, *and (ii)*

$$\frac{(\widehat{\varepsilon}_{jt} - \varepsilon_{jt})}{s_{jt}} \xrightarrow{d} N(0,1),$$

where $s_{jt}^2 = \frac{1}{T} F_t' Avar(\bar{\delta}_j) F_t + \frac{1}{N} Avar(\widehat{G}_{jt})$, *where* $Avar(\bar{\delta}_j)$ *is the asymptotic variance from a hypothetical regression of* G_{jt} *on* F_t.

We generically refer to ε_{jt} as a measurement error, even though it might be due to systematic differences between F_t and G_{jt}. If F_t was observed, $\bar{\delta}_j$ would have been the resulting estimate. The error in fitting G_{jt} would then be confined to the first term of s_{jt}^2, as is standard of regression analysis. But because we regress G_t on $\widetilde{F}_t$, $\widehat{\varepsilon}_{jt}$ now consists of the error from estimation of F_t. Notably, the error from having to estimate F_t is decreasing in N. The convergence rate of $\widehat{\varepsilon}_{jt}$ to ε_{jt} is $\min[\sqrt{N}, \sqrt{T}]$ and follows from Theorem 2 of Bai and Ng (2004). As can be seen from the expression of s_{jt}, no restriction is placed on the relative rate of increase between N and T.

When ε_t is homoskedastic, $Avar(\bar{\delta}_j) = \sigma_\varepsilon^2(\frac{F'F}{T})^{-1}$ and thus $s_{jt}^2 = \frac{1}{T}F_t'(\frac{F'F}{T})^{-1}F_t\sigma_\varepsilon^2 + \frac{1}{N}Avar(\widehat{G}_{jt})$. A consistent estimator of s_{jt}^2 when ε_{jt} is conditionally homoskedastic is thus

$$\widehat{s}_{jt}^2 = \frac{1}{T}\widetilde{F}_t'\widetilde{F}_t\widehat{\sigma}_\varepsilon^2(j) + \frac{1}{N}\widehat{Avar}(\widehat{G}_{jt}),$$

where $\widehat{Avar}(\widehat{G}_{jt})$ is given earlier, and again noting that $\frac{\widetilde{F}'\widetilde{F}}{T} = I_r$. When ε_{jt} is heteroskedastic, the first term of $\widehat{s}_{jt}^2$ is replaced by $\frac{1}{T}\widetilde{F}_t'(\frac{1}{T}\sum_{s=1}^T \widetilde{F}_s\widetilde{F}_s'\widehat{\varepsilon}_{js}^2)\widetilde{F}_t$ following White (1980).

Explicit consideration of measurement error allows us to construct, for each t, a confidence interval for ε_{jt}. For example, at the 95% level, the confidence interval is

$$(\varepsilon_{jt}^-, \varepsilon_{jt}^+) = (\widehat{\varepsilon}_{jt} - 1.96\widehat{s}_{jt}, \quad \widehat{\varepsilon}_{jt} + 1.96\widehat{s}_{jt}). \tag{7}$$

If G_{jt} is an exact factor, zero should lie in the confidence interval for each t. It could be of economic interest to know at which time points G_{jt} deviates from $\widehat{G}_{jt}$.

Instead of comparing G_{jt} with $\widehat{G}_{jt}$ at each t, two overall statistics can also be constructed.

$$NS(j) = \frac{\widehat{\text{var}}(\widehat{\varepsilon}(j))}{\widehat{\text{var}}(\widehat{G}(j))} \tag{8}$$

$$R^2(j) = \frac{\widehat{\text{var}}(\widehat{G}(j))}{\widehat{\text{var}}(G(j))}. \tag{9}$$

The $NS(j)$ statistic is simply the noise-to-signal ratio. If G_{jt} is an exact factor, the population value of $NS(j)$ is zero. A large $NS(j)$ thus indicates important departures of G_{jt} from the latent factors. A limitation of this statistic is that it leaves open the question of what is small and what is large. For this reason, we also consider $R^2(j)$, which should be unity if G_{jt} is an exact factor, and zero if the observed variable is irrelevant. The practitioner can draw inference as to how good is G_{jt} as a proxy factor by picking cut off points for $NS(j)$ and $R^2(j)$, in the same way we select the size of the test. These statistics are useful because instead of asking how large is the measurement error in the proxy variable that would overturn hypothesis, we now have consistent estimates for the size of the measurement error.

3.2 Testing G_t as a set

Suppose there are r latent and m observed factors. Whether r exceeds m or vice versa, it is useful to gauge the general coherency between F_t and G_t. To this end, we consider the canonical correlations between F_t and G_t. Let S_{FF} and S_{GG} be the sample variance-covariance matrix of F and G respectively. The sample squared-canonical correlations, denoted by $\widehat{\rho}_k^2, k = 1, \dots \min[m, r]$, are the largest eigenvalues of the $r \times r$ matrix

$S_{FF}^{-1}S_{GF}S_{GG}^{-1}S_{FG}$. It is well known that if F and G are observed and are normally distributed, $z_k = \frac{\sqrt{T}(\widehat{\rho}_k^2-\rho_k^2)}{2\rho_k(1-\rho_k^2)} \xrightarrow{d} N(0,1)$ for $k = 1, \ldots \min[m, r]$, see Anderson (1984). Muirhead and Waternaux (1980) provide results for non-normal distributions. For elliptical distributions, they showed that $z_k = \frac{1}{(1+\kappa/3)} \frac{\sqrt{T}(\widehat{\rho}_k^2-\rho_k^2)}{2\rho_k(1-\rho_k^2)} \xrightarrow{d} N(0,1)$, where κ is the excess kurtosis of the distribution.[3] Their results cover the multivariate normal, some contaminated normal (mixture normal), and the multivariate t, which are all elliptical distributions. Properties of canonical correlations for non-elliptical distributions can be obtained using results in Yuan and Bentler (2000). Our analysis is complicated by the fact that F is not observed but estimated. Nevertheless, the following holds.

Proposition 3 *Let $\widetilde{\rho}_1^2 \ldots, \widetilde{\rho}_p^2$ be the largest $p = \min[m, r]$ sample squared canonical correlations between $\widetilde{F}$ and G, where $\widetilde{F}_t$ is the principal components estimate of F_t. Let $N, T \to \infty$ with $\sqrt{T}/N \to 0$.*

(i) Suppose that $(F_t', G_t')'$ are iid normally distributed,

$$\widetilde{z}_k = \frac{\sqrt{T}(\widetilde{\rho}_k^2 - \rho_k^2)}{2\widetilde{\rho}_k(1-\widetilde{\rho}_k^2)} \xrightarrow{d} N(0,1), \qquad k = 1, \ldots, \min[m, r]. \tag{10}$$

(ii) Suppose that $(F_t', G_t')'$ are iid and elliptically distributed. Then

$$\widetilde{z}_k = \frac{1}{(1+\kappa/3)} \frac{\sqrt{T}(\widetilde{\rho}_k^2 - \rho_k^2)}{2\widetilde{\rho}_k(1-\widetilde{\rho}_k^2)} \xrightarrow{d} N(0,1), \qquad k = 1, \ldots, \min[m, r]. \tag{11}$$

where κ is the excess kurtosis.

The results hold for a class of non-elliptical distributions characterized by Yuan and Bentler (2000). The iid assumption on $(F_t', G_t')'$ can also be relaxed. For example, suppose F_t and G_t can be described by finite order VARs such that $A(L)F_t = \xi_t$ and $B(L)G_t = \eta_t$, where $(\xi_t', \eta_t')'$ are iid. Then canonical correlations can be performed using the residuals after fitting the VARs.[4]. We expect the same limiting distributions will be obtained when using $\widetilde{F}_t$ in place of F_t in the VAR.

Proposition 3 establishes that $\widetilde{z}_k$ is asymptotically the same as z_k, so that having to estimate F has no effect on the sampling distribution of the canonical correlations. This allows us to construct $(1-\alpha)$ percent confidence intervals for the population canonical correlations as follows. For $k = 1, \ldots \min[m, r]$,

$$\left(\rho_k^{2-}, \rho_k^{2+}\right) = \left(\widetilde{\rho}_k^2 - 2\Phi_\alpha \frac{\widetilde{\rho}_i(1-\widetilde{\rho}_i^2)}{\sqrt{T}}, \qquad \widetilde{\rho}_k^2 + 2\Phi_\alpha \frac{\widetilde{\rho}_i(1-\widetilde{\rho}_i^2)}{\sqrt{T}}\right). \tag{12}$$

[3] A random vector Y is said to have an elliptical distribution if its density is of the form $c|\Omega^{-1/2}|g(y'\Omega^{-1}y)$ for some constant c, positive-definite matrix Ω, and nonnegative function g.

[4] Johansen (1988) used the same idea, albeit in a non-stationary context.

If every element of G_t is an exact factor, all the non-zero population canonical correlations should be unity. The confidence interval for the smallest non-zero canonical correlation is thus a bound for the weakest correlation between F_t and G_t.

The only non-zero canonical correlation between a single series, say, G_{jt} and $\widetilde{F}_t$ is $\widehat{\rho}_1^2$. But this is simply the coefficient of determination from a projection of G_{jt} onto $\widetilde{F}_t$, and thus coincide with $R^2(j)$ as defined in (9). The formula in (12) can therefore be used to obtain a confidence interval for $R^2(j)$ also.

The results provided in this section suggest that tests of the factor pricing theory using consistently estimated factors should be more precise than using proxy variables. In two pass regression tests, one first regresses excess returns on the proxy factors to obtain the 'betas'. In the second step, the estimated betas are used as regressors to obtain estimates of the risk premia. Although the betas with respect to G_t can be consistently estimated, $\frac{\text{cov}(X_{it},G_t)}{\text{var}(G_t)}$ will, in general, be different from $\frac{\text{cov}(X_{it},F_t)}{\text{var}(F_t)}$, the true betas with respect to F_t. Instead of asking how large must measurement errors be to overturn conclusions drawn from the proxy factors as in Shanken (1992), we can estimate the size of the measurement error. Furthermore, the true betas can be consistently estimated up to a matrix transformation using $\widetilde{F}_t$ in the first pass regression. With additional identifying assumptions, the estimates can, in principle, be rotated so that the risk premia estimated in the second step can be given interesting economic interpretation.[5] Formulas in the preceding section can then be used to provide standard errors for the risk premia.

Factor pricing theory also implies that if X_{it} is excess returns and F_t is the vector of factor returns, then in the time series regression $X_{it} = \alpha_i + \beta_i F_t + \varepsilon_{it}$, the restriction $\alpha_i = 0$ should hold for all i. If $\widetilde{F}_t$ is used in place of F_t in the regression, we can use the foregoing results to adjust the variance of $\widehat{\alpha}_i(\widetilde{F})$ to reflect the error due to $\widetilde{F}_t$. Note, however, that this error falls with N. But if we use G_t to proxy for F_t and $G_t = \delta' F_t + \varepsilon_t$, the measurement error has a non vanishing effect on $\widehat{\alpha}_i(G)$ even when N and T are large. In addition, the estimated α_i will not be consistent in the presence of measurement errors. Since inference may be invalid, use of consistently estimated F_t is preferred over use of proxy variables, with the proviso that inference takes into the account that the factors are being estimated. The present analysis shows how this correction can be made for the tests considered. A more general analysis of using estimated factors in forecast and regression analysis is given in Bai and Ng (2004).

[5]Identification of factors using a priori restrictions is also referred to as confirmatory factor analysis. See Anderson (1984).

4 Simulations

We use simulations to assess the finite sample properties of the tests. Throughout, we assume $F_{kt} \sim N(0,1), k = 1, \dots r$, and $e_{it} \sim N(0, \sigma_e^2(i))$, where e_{it} is uncorrelated with e_{jt} for $i \neq j, i, j = 1, \dots N$. When $\sigma_e^2(i) = \sigma_e^2$ for all i, we have the case of homogeneous data. The factor loadings are standard normal, i.e. $\lambda_{ij} \sim N(0,1), j = 1, \dots r, i = 1, \dots N$. The data are generated as $x_{it} = \lambda_i' F_t + e_{it}$. In the experiments, we assume that there are $r = 2$ factors and that this is known.[6] The data are standardized to have mean zero and unit variance prior to estimation of the factors by the method of principal components.

The observed factors are generated as $G_{jt} = \delta_j' F_t + \varepsilon_{jt}$, where δ_j is a $r \times 1$ vector of weights, and $\varepsilon_{jt} \sim \sigma_\varepsilon(j) N(0, \text{var}(\delta_j' F_t))$. We test $m = 7$ observed variables parameterized as follows[7]:

Parameters for G_{jt}

j	1	2	3	4	5	6	7
δ_{j1}	1	1	1	1	1	1	0
δ_{j2}	1	0	1	0	1	0	0
σ_ε	0	0	.2	.2	2	2	1

The first two factors, G_{1t} and G_{2t} are exact factors since $\sigma_\varepsilon = 0$. Factors three to six are linear combinations of the two latent factors but are contaminated by errors. The variance of this error is small relative to the variations of the factors for G_{3t} and G_{4t}, but is large for G_{5t} and G_{6t}. Finally, G_{7t} is an irrelevant factor as it is simply a $N(0,1)$ random variable unrelated to F_t. Prior to testing, the G_ts are also standardized to have mean zero and unit variance. We conduct 1000 replications using Matlab 6.5.

We report results with α set to 0.025. Results with $\widehat{\text{var}}(\widehat{G}_t)$ defined as in (6), (5) and (4) are given in Table 1a-c, respectively. According to theory, $A(j)$ should be 2α if G_{jt} is a true factor, and unity if the factor is irrelevant. Furthermore, $M(j)$ should exceed the critical value at percentage point α with probability 2α. Columns four and five report the properties of these tests averaged over 1000 replications. Indeed, for G_{1t} and G_{2t}, the rejection rates are close to the nominal size of 5%, even for small samples. For the irrelevant factor G_{7t}, the tests reject the null hypothesis with high probabilities, showing the tests have power.[8] The

[6] In finite samples, the criteria developed in Bai and Ng (2002) for selecting the number of factors are excellent even under heteroskedasticity, mild weak and cross section correlation.

[7] Results for fat-tailed and cross-correlated errors are available in working version of this paper, available at `http://www.econ.lsa.umich.edu/~ngse/papers/observe.pdf`.

[8] This notwithstanding, there may exist other tests that have higher power. Admissibility of the test is beyond the scope of the present analysis.

power of the $M(j)$ test is especially impressive. Even when heteroskedasticity in the errors has to be accounted for, the rejection rate is 100 percent, for all sample sizes considered. This means that even with $(N, T) = (50, 50)$, the test can very precisely determine if an observed variable is an exact factor. The $NS(j)$ test reinforces the conclusion that G_{1t} and G_{2t} are exact factors, and that G_{7t} has no coherence with the latent factors. Table 1 also reports the average estimate of $R^2(j)$. The two exact factors have estimates of $R^2(j)$ well above 0.95, while uninformative factors have $R^2(j)$ well below 0.05, with tight confidence intervals.[9]

Because we know the data generating process, we can assess whether or not the confidence intervals around ε_{jt} give correct inference. In the column labelled $CI_\varepsilon(j)$ in Table 1 , we report the probability that the true ε_{jt} lies inside the two-tailed 95% confidence interval defined by (7). Evidently, the coverage is excellent for ε_{1t}, ε_{2t}, and ε_{7t}. The result for ε_{7t} might seem surprising at first, but this is in fact showing that the measurement error can be precisely estimated even when F_t and G_{jt} are totally unrelated.

In theory, $\tau(j)$ and $M(j)$ should always reject the null hypothesis when G_3, G_4, G_5 and G_6 are being tested since none of these are exact factors. Table 1 shows that this is the case when N and T both exceed 100. For smaller N and/or T, the power of the tests depend on how large are the measurement errors. For G_{5t} and G_{6t}, which have a high noise-to-signal ratio of 4, $M(j)$ still rejects with probability one when N and T are small, while the $A(j)$ has a respectable rejection rate of 0.85. However, when the signal-to-noise ratio is only .04 as in G_{3t} and G_{4t}, both the $A(j)$ and the $M(j)$ under-reject the null hypothesis, and the problem is much more severe for $A(j)$.

Notice that when the noise-to-signal ratio is .04, $R^2(j)$ remains at around .95. This would be judged high in a typical regression analysis, and yet we would reject the null hypothesis that G_{jt} is an exact factor. It is for cases such as these that having a sense of how big is the measurement error is useful. In our experience, an $NS(j)$ above .5, and/or a $R^2(j)$ below 0.95 is symptomatic of non-negligible measurement errors. By these guides, G_{3t} and G_{4t} are strong proxies for the latent factors.

It is of interest to remark that whether the measurement error has large or small variance, the confidence intervals constructed according to (7) bracket the true error quite precisely. This is useful in empirical work since we can learn if the discrepancy between G_{jt} and the latent factors are systematic or occasional.

[9] Since $R^2(j)$ is bounded between zero and one, the estimated lower bound should be interpreted in this light.

Table 2 reports results for testing four sets of observed factors using canonical correlations. These are formed from G_{1t} to G_{6t} as defined above, plus four $N(0,1)$ random variables unrelated to F_t, labelled G_{7t} to G_{10t}. The four sets of factors are defined as follows:

Set 1: G_{3t}, G_{4t}, and G_{7t} Set 2: G_{5t}, G_{6t}, and G_{7t}
Set 3: G_{1t}, G_{2t}, and G_{7t} Set 4: G_{7t}, G_{8t}, G_{9t}, and G_{10t}.

Because min$[m, r]$ is 2 in each of the four case, the largest two canonical correlations should be unity if every G_t is an exact factor. We use (12) to construct confidence intervals for ρ_2^2, i.e. the smaller of the two non-zero canonical correlations. Table 2 shows that Set 3 has maximal correlation with $\widetilde{F}_t$ as should be the case since G_{1t} and G_{2t} are exact factors, and a weight of zero on G_{7t} would indeed maximize the correlation between G_t and $\widetilde{F}_t$. When Set 4 is being tested, zero is in the confidence interval as should be the case since this is a set of irrelevant factors. For Set 2 which has factors contaminated by large measurement errors, the test also correctly detects a very small canonical correlation. When measurement errors are small but non-zero, the sample correlations are non-zero but also not unity. The practitioner again has to take a stand on whether a set of factors is useful. The values of $\widehat{\rho}_2^2$ for Set 1 are around .9, below our cut-off point of .95. We would thus be more concerned with accepting Set 1 as a valid set than accepting its elements (i.e. G_{3t} and G_{4t}) as individually valid factors.

Finally, to illustrate the performance of proposed tests, we first consider the case of two latent factors (F_{1t}, F_{2t}) and two observed factors ($G_{1t} = F_{1t}$, $G_{2t} = F_{2t}$). Thus by design, the observed variables are exact factors. Figure 1 displays the true G_t along with the confidence intervals computed as

$$([\widehat{G}_{jt} - 1.96\widehat{\text{var}}(\widehat{G}_{jt})^{1/2},\, \widehat{G}_{jt} + 1.96\widehat{\text{var}}(\widehat{G}_{jt})^{1/2}]$$

for $t = 1, ..., T$ and $j = 1, 2$. The top left panel is for the first factor with $N = 50$, and the top right panel is for the second factor. The bottom panel plots the confidence intervals when $N = 100$. Clearly, the true G_{jt} is inside the confidence intervals and these become narrower for larger N.

Next, we assume $G_{jt} = F_{jt} + \varepsilon_{jt}$ with $\varepsilon_{jt} \sim N(0,1)$ for $j = 1, 2$. The measurement errors ε_{jt} are estimated for $t = 1, ..., T$, $j = 1, 2$, and the confidence intervals are constructed according to Proposition 2:

$$[\widehat{\varepsilon}_{jt} - 1.96\,\widehat{s}_{jt},\, \widehat{\varepsilon}_{jt} + 1.96\,\widehat{s}_{jt}].$$

Since in simulations, the true error processes ε_{jt} are known, they are also plotted in Figure 2 along with the confidence intervals. It is clear that the confidence intervals cover the true process. However, in contrast to Figure 1 with exact factors, the confidence intervals do not become narrower as N increases. This is because δ is estimated with T observations. In results not reported, we verify that the confidence bands are narrower as T increases.

5 Empirical Applications

In this section, we take our tests to the data. Factors estimated from portfolios, stock returns, and a large set of economic variables will be tested against various G_{jt}s. The base factors are the three factors considered in Fama and French (1993), denoted 'Market', 'SMB', and 'HML'.[10] In addition to the FF factors, we also include variables considered in Chen et al. (1986). These are the first lag of innovations to annual consumption growth 'DC', inflation 'DP', the growth rate of industrial production 'DIP', a term premia 'TERM', and a risk premia 'RISK'.[11] For annual data, the innovations are the residuals from estimation of autoregressions with two lags. For monthly data, the innovations are the residuals from an autoregression with six lags. In each case, we analyze the data from 1960-1996. We also split the sample at various points to look for changes in relations between the observed and the latent factors over the forty years. The data are standardized to be mean zero with unit variances prior to estimation by the method of principal components. The G_ts are likewise standardized prior to implementation of the tests. In view of the properties of the data, we only report results for heteroskedastic errors with $\widehat{\text{var}}(\widehat{G}_t)$ defined as in (5).

5.1 Portfolios

In this application, x_{it} are monthly or annual observations on 100 portfolios available from Kenneth French's web site.[12] These are the intersections of 10 portfolios formed on size (market equity) and 10 portfolios formed on the ratio of book to market equity. A total of 89 portfolios are continuously available for the full sample. Depending on the sample period,

[10] Small minus big is the difference between the average return of three small portfolios and three big portfolios. High minus low is the average return on two value and two growth portfolios. See Fama and French (1993). $Market = R_m - R_f$ is value weighted return on all NYSE, AMEX, and NASDAQ minus the one month treasury bill rate from Ibbotson Associates.

[11] The data are taken from citibase. 'DC' is the growth rate of PCEND, 'DIP' is the growth rate of IP, 'DP' is the growth rate of PUNEW. The risk premium,'RISK', is the BAA rated bond rate (FYBAAC) minus the 10 year government bond rate (FYGT10). The term premia, 'TERM' is 10 year government bond FYGT10 minus the three month treasury bill rate FYGM3.

[12] Web site `mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_100_port_sz.html`.

the PCP and ICP select between 4 and 6 factors.[13] We set $r = 6$ in all subsequent tests. The results are reported in Table 3.

For annual data, the $A(j)$ test rejects the null hypothesis of exact factors in more than 5% of the sample. The critical value for the $M(j)$ test is 3.28. It cannot reject the null hypothesis that SMB is an exact factor at the 5% level, and HML and Market at around the 10% level. However, the evidence does not support the macroeconomic variables as exact factors. The $NS(j)$ and $R^2(j)$ also indicate the presence of non-trivial measurement errors. The canonical correlations suggest only three well defined relations between F_t and G_t. The remaining relations are extremely weak. The canonical correlations between the three FF factors alone, and $\widetilde{F}_t$ are .970, .962, and .950, respectively. Little is gained by including 'DC', 'DIP', 'DP', 'TERM', and 'RISK'. This suggests that the FF factors underlie the three non-zero canonical correlations in the eight variable set. Of the macroeconomic variables considered, 'RISK' has the highest explanatory power for returns.

Results for testing monthly data are given in Table 4. Several features are noteworthy. First, the FF factors continue to be strong proxies for systematic risks. Statistically, we cannot reject the null hypothesis that the SMB is an exact factor in the 1988-1996 subsample. Of the three FF factors, 'Market' has the highest $R^2(j)$. Second, over the entire sample, RISK has the highest $R^2(j)$ and the lowest $NS(j)$, but its relation with the latent factors is far weaker than the FF factors. Third, it could be argued that the macroeconomic variables have unstable relations with the unobserved factors over time, and that we are jointly testing the joint hypothesis that G_{jt} is a linear combination of the fundamental factors *and* the relation is stable. We therefore consider various subsamples. The change from fixed to flexible exchange rate regime and the 1982 recession are plausible times that the relation between the observed and fundamental factors changed.[14] The statistics suggest that the relations have become stronger in the more recent subsamples. However, even after allowing for parameter instability, the evidence remains that there is lack of coherence between the observed and the macroeconomic variables. In contrast, the relation between the FF factors and the latent factors appears to be quite stable, displaying little variation in both $R^2(j)$ and $NS(j)$ over time. Fourth, three of the six sample canonical correlations (since $\min[m, r] = 6$) between $\widetilde{F}_t$ and G_t are practically zero, from which we can conclude that the

[13] We consider $g_1(N,T) = \log\left(\frac{NT}{N+T}\right)\frac{N+T}{NT}$, $g_2(N,T) = \log(\min[N,T])\frac{N+T}{NT}$.

[14] As suggested by a referee, a variable could be a fundamental factor in one subsample and not the other, and therefore not in the full sample, which could account for the rejections in the full sample. Indeed, there is instability in δ_j.

eight observed factors considered cannot span the true factor space. Because the five non-zero canonical correlations are also far from unity, we can also conclude that measurement errors are significant enough that the eight observed variables cannot even span a five dimensional subspace of the true factors.

5.2 Stock Returns

We next apply our tests to monthly stock returns. Because portfolios are aggregated from stocks, individual stock returns should have larger idiosyncratic variances than portfolios. Thus, the common components in the individual returns can be expected to be smaller than those in the portfolios. It is thus of interest to see if good proxy variables can be found when the data have larger idiosyncratic noises.

Data for 190 firms are available from CRSP over the *entire* sample.[15] The two PCP criteria select 6 and 5 factors in this panel of data, while the ICP always selects 4. We set r to 6 in the analysis. The results are reported in Table 5. None of the observed variables can be considered an exact factor, though 'Market' is a strong proxy factor with a low noise to signal ratio and a high R^2. However, SMB, HML, and especially the macroeconomic variables are poor proxies for the factors in the returns data. The best of the proxy macroeconomic variable – 'RISK'– still has a noise-to-signal ratio that is ten times larger than the 'Market' factor.

The above analysis indicates that the factors in annual portfolios are better approximated by observed variables than the factors in monthly portfolios, and finding proxies for the factors in the monthly portfolios is in turn a less challenging task than finding observed variables to proxy the factors in individual returns. This is because high frequency and/or disaggregated data are more likely to be contaminated by noise. Thus, even though more data are available at the high frequency and disaggregated levels, they are less reliable proxies for the systematic variations in the data. Inference using observed variables as proxies for the common factors will likely be distorted by the measurement noise. Of all the variables considered, the most satisfactory proxy for the latent factors in *both* portfolios and individual stock returns appears to be the 'Market' factor as described in Fama and French (1993). Its signal to noise ratio is systematically high, and its coherence with the latent factors is robust across sample periods.

[15]There are data on many more firms that exist in a given year, but returns for only 190 firms are available continuously over the entire sample. The readers are cautioned that the results could be distorted by survivorship bias.

5.3 Macroeconomic Factors

A fundamental characteristic of business cycles is the comovement of a large number of economic series. Such a phenomenon can be rationalized by common factors being the driving force of economic fluctuations. This has been used as a justification to represent the unobserved state of the economy by variables thought to be dominated by variations arising from common sources. Industrial production, unemployment rate, and various interest rate spreads have been used for this purpose. However, there has been no formal analysis of how good the proxy variables are.

We estimate the latent factors from 150 monthly series considered in Stock and Watson (2002). This data consist of series from output, consumption, employment, investment, prices, wages, interest rate, and other financial series such as exchange rates. In addition to the five macroeconomic variables considered throughout, we also tested if unemployment rate (LHUR) is a common factor. For this application, we also consider a post Bretton Woods sub-sample. The results are reported in Table 6.

Given the noise in monthly data, rejecting the null hypothesis that the variables are exact factors is hardly surprising. Industrial production is a widely used indicator of economic activity. While it has one of the strongest relations with the latent factors, the R^2 is only around .5, indicating that the noise component is also non-trivial. Interestingly, 'RISK' bears a strong relation not just with the latent factors in portfolios, but also in the macroeconomic variables. As in the results for asset returns, the relations between the macroeconomic variables and the factors also appear to be unstable. While inflation has become less important since 1983, unemployment has become more important. The non-zero canonical correlations also reveal instability in the relations, as they are higher in the later than the earlier subsamples. One interpretation of this result is that idiosyncratic shocks were more important in the sixties and seventies, but common shocks have become more important in recent years as sources of economic fluctuations.

To the extent that the FF factors are good proxies for the factors in portfolios, one might wonder if the common shocks to economic activity are also the common shocks to portfolio returns. To shed some light on this issue, we also tested if the FF factors are related to the factors in macroeconomic data. As seen from Table 5, there is hardly any evidence for a relation between the FF factors and the panel of macroeconomic variables.

6 Conclusion

It is common practice in empirical work to proxy unobserved common factors by observed variables, yet hardly any formal procedures exist to assess if the observed variables equal, or are close to the factors. This paper exploits the fact that the space spanned by the common factors can be consistently estimated from large dimensional panels. We develop several tests that can serve as guides as to which variables are close to the factors. The tests have good properties in simulations. We estimate the common factors in portfolios, individual returns, as well as a large set of macroeconomic data. The Fama and French factors approximate the factors in portfolios and individual stock returns much better than any single macroeconomic variable. None of the macroeconomic variables considered, namely industrial production, inflation, unemployment rate, and the risk premium, have stronger coherence with the macroeconomic factors. Inflation was a good proxy of the factors in portfolios and macroeconomic data prior to the 1980s, but its importance has diminished since.

Proof of Proposition 1

For national simplicity, we omit the subscript j, which refers to the jth series. We will use results of Bai (2003) repeatedly, though it should be noted that the H matrix in the present paper is H' in Bai (2003), purely for notational simplicity.

Under the null hypothesis, $G_t = \delta' F_t$ for all t. Adding and subtracting terms, $G_t = \delta' H^{-1}\widetilde{F}_t + \delta' H^{-1}(HF_t - \widetilde{F}_t)$, or

$$G_t = \gamma'\widetilde{F}_t + \gamma'(HF_t - \widetilde{F}_t),$$

where $\gamma = H^{-1\prime}\delta$. In matrix notation,

$$G = \widetilde{F}\gamma + (FH' - \widetilde{F})\gamma. \tag{13}$$

The least squares estimator of γ is

$$\widehat{\gamma} = (\widetilde{F}'\widetilde{F}/T)^{-1}(\widetilde{F}'G/T) = \frac{1}{T}\widetilde{F}'G.$$

Substituting G of (13) into $\widehat{\gamma}$, we have

$$\begin{aligned}\widehat{\gamma} &= \gamma + \frac{1}{T}\widetilde{F}'(FH' - \widetilde{F})\gamma,\\ \sqrt{N}(\widehat{\gamma} - \gamma) &= \sqrt{N}\frac{1}{T}\widetilde{F}'(FH' - \widetilde{F})\gamma.\end{aligned}$$

From Lemma B.3 of Bai (2003), $\frac{1}{T}\widetilde{F}'(FH' - \widetilde{F}) = O_p(\min[N,T]^{-1})$. So $\sqrt{N}(\widehat{\gamma} - \gamma) = \sqrt{N}\cdot O_p(\min[N,T]^{-1}) \to 0$ provided $\sqrt{N}/T \to 0$.

Next, $\sqrt{N}(\widehat{G}_t - G_t) = \sqrt{N}(\widehat{\gamma}' - \gamma')HF_t + \widehat{\gamma}'\sqrt{N}(\widetilde{F}_t - HF_t)$. The above analysis shows $\sqrt{N}(\widehat{G}_t - G_t) = \widehat{\gamma}'\sqrt{N}(\widetilde{F}_t - HF_t) + o_p(1)$. Let $Avar(\widetilde{F}_t)$ denote the asymptotic variance of $\sqrt{N}(\widetilde{F}_t - HF_t)$. Then Lemma 1(ii) implies that, as $N, T \to \infty$ with $\sqrt{N}/T \to 0$,

$$\frac{\sqrt{N}(\widehat{G}_t - G_t)}{(\widehat{\gamma}' Avar(\widetilde{F}_t)\widehat{\gamma})^{1/2}} \xrightarrow{d} N(0,1).$$

Let $\text{var}(\widehat{G}_t) = \frac{1}{N}\widehat{\gamma}' Avar(\widetilde{F}_t)\widehat{\gamma}$, then the above is the same as $\tau_t(j) \xrightarrow{d} N(0,1)$, where $\tau_t(j)$ is defined in (1). Recall $\widehat{\tau}_t(j)$ is obtained when $\text{var}(\widehat{G}_t)$ is replaced by a consistent estimator $\widehat{\text{var}}(\widehat{G}_t)$, thus $\widehat{\tau}_t(j) \xrightarrow{d} N(0,1)$. It follows that $P(|\widehat{\tau}_t(j)| > \Phi_\alpha) \to 2\alpha$. We next show that $\frac{1}{T}\sum_{t=1}^{T} 1(|\widehat{\tau}_t(j)| > \Phi_\alpha) \to 2\alpha$. The essential argument is the law of large numbers, but because the object of interest depends on both N and T, extra considerations are required.

Note first that in the preceding analysis, the limiting normality of $\widehat{\tau}_t(j)$ is derived from the limiting normality of $\widehat{\gamma}'\sqrt{N}(\widetilde{F}_t - HF_t)$. From Bai (2003),

$$\sqrt{N}(\widetilde{F}_t - HF_t) = \widetilde{V}^{-1}(\widetilde{F}'F/T)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{it} + o_p(1).$$

The normality of $\sqrt{N}(\widetilde{F}_t - HF_t)$ is therefore derived from the CLT for $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{it}$ when $N \to \infty$. Thus, we can write

$$\widehat{\tau}_t(j) = \tau_{t,N} + o_p(1) \tag{14}$$

where $\tau_{t,N} = \psi'\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{it}$, and ψ is a vector of constant, and $\tau_{t,N}\xrightarrow{d}N(0,1)$. For any fixed real number $x_0 > 0$, if $\tau_{t,N}$ has a continuous distribution, then $P(\tau_{t,N} = x_0) = 0$. Because the limit of $\tau_{t,N}$ is a continuous random variable, $P(\tau_{t,N} = x_0) \to 0$, as $N \to \infty$. This means that the indicator function $1(|x| \geq x_0)$ is (asymptotically) an almost surely continuous function with respect to the measure of $\tau_{t,N}$. It follows from (14) that

$$1(|\widehat{\tau}_t(j)| \geq x_0) = 1(|\tau_{t,N}| \geq x_0) + o_p(1).$$

Replacing x_0 by Φ_α, averaging over t, we obtain

$$\frac{1}{T}\sum_{t=1}^{T}1(|\widehat{\tau}_t(j)| \geq \Phi_\alpha) = \frac{1}{T}\sum_{t=1}^{T}1(|\tau_{t,N}| \geq \Phi_\alpha) + o_p(1).$$

Let $X_{T,N}$ denote the first term on the right hand side above. We will show the joint probability limit of $X_{T,N}$ is 2α. We use Theorem 1 of Phillips and Moon (1999), switching the role of N and T in their paper.

Define $Y_{t,N} = 1(|\tau_{t,N}| \geq \Phi_\alpha)$. From $\tau_{t,N}\xrightarrow{d}Z_t \sim N(0,1)$, we have $Y_{t,N}\xrightarrow{d}Y_t = 1(|Z_t| \geq \Phi_\alpha)$. For each fixed T, $X_{T,N}\xrightarrow{d}X_T = \frac{1}{T}\sum_{t=1}^{T}Y_t$ as $N \to \infty$. Because $E(Y_t) = 2\alpha$, by the law of large numbers, $X_T\xrightarrow{p}2\alpha$, as $T \to \infty$. To show this holds not just sequentially as $N \to \infty$ and then $T \to \infty$ but also as $N, T \to \infty$ jointly, we need to verify conditions (i)-(iv) of Phillips and Moon. Because $Y_{t,N}$ is bounded by 1, all conditions except (ii) of Phillips and Moon are trivially satisfied. For (ii) we need to show

$$\lim_{T,N\to\infty}\frac{1}{T}\sum_{t=1}^{T}|P(|\tau_{t,N}| \geq \Phi_\alpha) - P(|Z_t| \geq \Phi_\alpha)| = 0.$$

But by the Berry-Esseen Theorem, $|P(|\tau_{t,N}| \geq \Phi_\alpha) - P(|Z_t| \geq \Phi_\alpha)| \leq N^{-1/2}C$ for a constant C not depending on t given the moment conditions on e_{it}. Thus their condition (ii) is also met. This proves part (i) of Proposition 1.

Let $M = \max\{|Z_1|, |Z_2, ..., |Z_T|\}$, where Z_t are iid $N(0,1)$. Then $P(M < x) = (2\Phi(x) - 1)^T$. Assuming that the idiosyncratic errors e_{it} are serially uncorrelated, then $\tau_t(j)$ are asymptotically independent over t. This is because the limiting distribution of $\sqrt{N}(\widetilde{F}_t - HF_t)$ and thus $\tau_t(j)$ is determined by $N^{-1/2}\sum_{i=1}^N e_{it}$. Because $N^{-1/2}\sum_{i=1}^N e_{it}$ and $N^{-1/2}\sum_{i=1}^N e_{i,t+s}$ $(s \neq 0)$ are uncorrelated and asymptotically normal, they are asymptotically independent. This implies the asymptotic independence of $\tau_t(j)$ and $\tau_{t+s}(j)$ and thus $P(M(j) < x) \sim (2\Phi(x) - 1)^T$, proving Proposition 1.

The proof of Proposition 2 Follows from the same argument in the proof of Theorem 2 of Bai and Ng (2004). The details are omitted.

Proof of Proposition 3

Let $\widetilde{B} = S^{-1}_{\widetilde{F}\widetilde{F}} S_{\widetilde{F}G} S^{-1}_{GG} S_{G\widetilde{F}}$, and $B = S^{-1}_{FF} S_{FG} S^{-1}_{GG} S_{GF}$. Let $\widetilde{\rho}_k$ and $\widehat{\rho}_k$ be the canonical correlations of $\widetilde{B}$ and B, respectively. Because eigenvalues are continuous functions, we will have $\sqrt{T}(\widetilde{\rho}_k^2 - \widehat{\rho}_k^2) \xrightarrow{p} 0$ provided that $\sqrt{T}(\widetilde{B} - B) \xrightarrow{p} 0$. That is, the canonical correlations of $\widetilde{B}$ have the same limiting distributions as those of B when $\widetilde{B}$ and B are asymptotically equivalent. We next establish $\sqrt{T}(\widetilde{B} - B) \xrightarrow{p} 0$. First note that because H is full rank, the canonical correlations of B is the same as the canonical correlations of B^*, where

$$B^* = S^{-1}_{HF\,HF} S_{HF\,G} S^{-1}_{GG} S_{G\,HF}.$$

Thus, it suffices to show that $\sqrt{T}(\widetilde{B} - B^*) \xrightarrow{p} 0$. But this is implied by

$$\sqrt{T}(S^{-1}_{\widetilde{F}\widetilde{F}} - S^{-1}_{HF\,HF}) \xrightarrow{p} 0 \tag{15}$$

$$\sqrt{T}(S_{\widetilde{F}G} - S_{HF\,G}) \xrightarrow{p} 0. \tag{16}$$

Consider (15). Now

$$(S^{-1}_{\widetilde{F}\widetilde{F}} - S^{-1}_{HF\,HF}) = S^{-1}_{\widetilde{F}\widetilde{F}}(S_{HF\,HF} - S_{\widetilde{F}\widetilde{F}}) S^{-1}_{HF\,HF}.$$

Thus, (15) is implied by $\sqrt{T}(S_{HF\,HF} - S_{\widetilde{F}\widetilde{F}}) \xrightarrow{p} 0$, or that

$$\sqrt{T}\left(\frac{H'F'FH}{T} - \frac{\widetilde{F}'\widetilde{F}}{T}\right) \xrightarrow{p} 0. \tag{17}$$

By Lemma B.2 and B.3 of Bai (2003),

$$\begin{aligned} \frac{\widetilde{F}'(\widetilde{F} - FH')}{T} &= O_p(\min[N,T]^{-1}) \\ \frac{F'(\widetilde{F} - FH')}{T} &= O_p(\min[N,T]^{-1}). \end{aligned}$$

Adding and subtracting terms, (17) becomes

$$\frac{-\sqrt{T}(\widetilde{F}-FH')'\widetilde{F}}{T}-\frac{-\sqrt{T}HF'(\widetilde{F}-FH')}{T}=\sqrt{T}O_p(\min[N,T]^{-1})\rightarrow 0$$

if $\sqrt{T}/N \rightarrow 0$. For (16),

$$\sqrt{T}\left(\frac{\widetilde{F}'G}{T}-\frac{HF'G}{T}\right)=\sqrt{T}\left(\frac{(\widetilde{F}-FH')'G}{T}\right).$$

But $\frac{1}{T}(\widetilde{F}-FH')'G = O_p(\min[N,T]^{-1})$, see Lemma B.2 of Bai (2003). Thus, (16) is $O_p(\sqrt{T}/\min[N,T])\xrightarrow{p}0$, establishing $\sqrt{T}(\widetilde{B}-B^*)=o_p(1)$ and thus the proposition.

References

Anderson, T. W. 1984, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

Bai, J. 2003, Inferential Theory for Factor Models of Large Dimensions, *Econometrica* **71:1**, 135–172.

Bai, J. and Ng, S. 2002, Determining the Number of Factors in Approximate Factor Models, *Econometrica* **70:1**, 191–221.

Bai, J. and Ng, S. 2004, Confidence Intervals for Diffusion Index Forecasts with a Large Number of Predictors, mimeo, University of Michigan.

Breeden, D. 1989, An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities, *Journal of Financial Economics* **7**, 265–296.

Breeden, D., Gibbons, M. and Litzenberger, R. 1989, Empirical Tests of the Consumption-Oriented CAPM, *Journal of Finance* **XLIV:2**, 231–262.

Chamberlain, G. and Rothschild, M. 1983, Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets, *Econometrica* **51**, 1305–1324.

Chen, N., Roll, R. and Ross, S. 1986, Economic Forces and the Stock Market, *Journal of Business* **59**, 383–403.

Connor, G. and Korajzcyk, R. 1998, Risk and Return in an Equilibrium APT Application of a New Test Methodology, *Journal of Financial Economics* **21**, 225–289.

Fama, E. and French, K. 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics.*

Johansen, S. 1988, Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control* **12**, 231–254.

Kandel, S. and Stambaugh, R. 1987, On Correlations and Inference About Mean-Variance Efficiency, *Journal of Financial Economics* **18**, 61–90.

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

Lettau, M. and Ludvigson, S. 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time Varying, *Journal of Politcal Economy* **109:6**, 1238–1287.

Lintner, J. 1965, Security Prices, Risk, and Maximal Gains from Diversification, **20**, 587–615.

Lucas, R. E. 1978, Asset Prices in an Exchange Economy, *Econometrica* **46**, 1429–1445.

Merton, R. 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* **41**, 867–886.

Muirhead, R. J. and Waternaux, C. 1980, Asymptotic distribution in canonical correlation analysis and other multivariate procedures for nonnormal populations, *Bimetrika* **67:1**, 31–43.

Phillips, P. and Moon, R. 1999, Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica* **67**, 1057–1111.

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

Ross, S. 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Finance* **13**, 341–360.

Shanken, J. 1987, Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique, *Journal of Financial Economics* **18**, 91–110.

Shanken, J. 1992, On the Estimatin of Beta-Pricing Models, *Review of Financial Studies* **5:1**, 1–33.

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

Stambaugh, R. 1982, On the Exclusion of Assets from Tests of the Two Parameter Model, *Journal of Financial Economics* **10**, 235–168.

Stambaugh, R. 1988, The Information in Forward Rates: Implications for Models of the Term Structure, *Journal of Financial Economics* **21**, 41–70.

Stock, J. H. and Watson, M. W. 2002, Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics* **20:2**, 147–162.

White, H. 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* **48**, 817–38.

Yuan, K. H. and Bentler, P. 2000, Inferences on Correlation Coefficients in Some Classes of Non-Normal Disturbances, *Journal of Multivariate Analysis* **72**, 230–248.

Table 1a: Tests for G_{jt}: $\widetilde{\Gamma}_t$ defined in (6)

N	T	j	$A(j)$	$M(j)$	$NS(j)$	$CI_\varepsilon(j)$	$R^2(j)$	$R^{2-}(j)$	$R^{2+}(j)$
50	50	1	0.03	0.02	0.03	0.97	0.97	0.96	0.99
50	50	2	0.03	0.03	0.03	0.97	0.97	0.95	0.99
50	50	3	0.17	0.64	0.07	0.97	0.94	0.90	0.97
50	50	4	0.16	0.60	0.07	0.97	0.94	0.90	0.97
50	50	5	0.85	1.00	5.05	0.95	0.22	0.03	0.41
50	50	6	0.85	1.00	5.08	0.95	0.22	0.03	0.41
50	50	7	0.95	1.00	339.24	0.96	0.04	0.00	0.13
100	50	1	0.03	0.01	0.01	0.97	0.99	0.98	0.99
100	50	2	0.03	0.01	0.01	0.97	0.99	0.98	0.99
100	50	3	0.27	0.94	0.05	0.97	0.95	0.92	0.98
100	50	4	0.27	0.94	0.05	0.97	0.95	0.92	0.98
100	50	5	0.89	1.00	4.79	0.95	0.23	0.04	0.42
100	50	6	0.89	1.00	4.77	0.95	0.22	0.03	0.42
100	50	7	0.96	1.00	290.41	0.94	0.04	0.00	0.14
50	100	1	0.03	0.01	0.03	0.97	0.97	0.96	0.98
50	100	2	0.03	0.01	0.03	0.97	0.97	0.96	0.98
50	100	3	0.17	0.74	0.07	0.97	0.94	0.91	0.96
50	100	4	0.17	0.73	0.07	0.97	0.94	0.91	0.96
50	100	5	0.85	1.00	4.52	0.96	0.21	0.07	0.35
50	100	6	0.86	1.00	4.49	0.96	0.21	0.07	0.34
50	100	7	0.96	1.00	1780.38	0.97	0.02	0.00	0.07
200	100	1	0.03	0.01	0.01	0.97	0.99	0.99	1.00
200	100	2	0.03	0.01	0.01	0.97	0.99	0.99	1.00
200	100	3	0.40	1.00	0.05	0.97	0.96	0.94	0.97
200	100	4	0.40	1.00	0.05	0.97	0.96	0.94	0.97
200	100	5	0.93	1.00	4.32	0.95	0.21	0.07	0.35
200	100	6	0.92	1.00	4.27	0.95	0.21	0.08	0.35
200	100	7	0.98	1.00	431.39	0.96	0.02	0.00	0.07
100	200	1	0.03	0.01	0.01	0.97	0.99	0.98	0.99
100	200	2	0.03	0.00	0.01	0.97	0.99	0.98	0.99
100	200	3	0.27	1.00	0.05	0.97	0.95	0.94	0.96
100	200	4	0.27	1.00	0.05	0.97	0.95	0.94	0.96
100	200	5	0.90	1.00	4.23	0.96	0.20	0.11	0.30
100	200	6	0.90	1.00	4.15	0.96	0.21	0.11	0.30
100	200	7	0.98	1.00	946.21	0.96	0.01	0.00	0.03

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the critical value of 1.96 in the sample of size T. the column labelled $M(j)$ reports the frequency that $\max_{1\le t\le T}|\widehat{\tau}_t(j)|$ exceeds the critical value for a sample of size T. CI_ε is the coverage frequency for ε, the measurement error. $NS(j)$ is the noise to signal ratio, see (8). R^2 is defined in (9), R^{2-} and R^{2+} are the lower and upper 95% confidence interval, averaged over 1000 replications.

Table 1b: Tests for G_{jt}: $\widetilde{\Gamma}_t$ defined in (5)

N	T	j	$A(j)$	$M(j)$	$NS(j)$	$CI_\varepsilon(j)$	$R^2(j)$	$R^{2-}(j)$	$R^{2+}(j)$
50	50	1	0.05	0.07	0.01	0.95	0.99	0.99	1.00
50	50	2	0.05	0.07	0.01	0.95	0.99	0.99	1.00
50	50	3	0.41	1.00	0.05	0.95	0.96	0.93	0.98
50	50	4	0.42	1.00	0.05	0.95	0.95	0.93	0.98
50	50	5	0.92	1.00	4.92	0.94	0.23	0.03	0.42
50	50	6	0.92	1.00	4.89	0.94	0.23	0.03	0.42
50	50	7	0.97	1.00	193.91	0.94	0.04	0.00	0.13
100	50	1	0.05	0.05	0.00	0.95	1.00	0.99	1.00
100	50	2	0.05	0.05	0.00	0.96	1.00	0.99	1.00
100	50	3	0.55	1.00	0.04	0.95	0.96	0.94	0.98
100	50	4	0.55	1.00	0.04	0.95	0.96	0.94	0.98
100	50	5	0.95	1.00	4.75	0.94	0.23	0.04	0.43
100	50	6	0.95	1.00	4.69	0.95	0.23	0.03	0.42
100	50	7	0.98	1.00	177.17	0.93	0.04	0.00	0.14
50	100	1	0.05	0.07	0.01	0.95	0.99	0.99	0.99
50	100	2	0.05	0.06	0.01	0.95	0.99	0.99	0.99
50	100	3	0.42	1.00	0.05	0.95	0.95	0.94	0.97
50	100	4	0.42	1.00	0.05	0.95	0.95	0.94	0.97
50	100	5	0.93	1.00	4.41	0.95	0.21	0.07	0.35
50	100	6	0.93	1.00	4.39	0.95	0.21	0.07	0.35
50	100	7	0.98	1.00	408.56	0.95	0.02	0.00	0.07
200	100	1	0.05	0.05	0.00	0.95	1.00	1.00	1.00
200	100	2	0.05	0.04	0.00	0.95	1.00	1.00	1.00
200	100	3	0.67	1.00	0.04	0.95	0.96	0.95	0.98
200	100	4	0.67	1.00	0.04	0.95	0.96	0.94	0.98
200	100	5	0.96	1.00	4.29	0.95	0.21	0.07	0.35
200	100	6	0.96	1.00	4.24	0.94	0.22	0.08	0.35
200	100	7	0.99	1.00	465.28	0.95	0.02	0.00	0.07
100	200	1	0.05	0.07	0.00	0.95	1.00	0.99	1.00
100	200	2	0.05	0.07	0.00	0.95	1.00	0.99	1.00
100	200	3	0.55	1.00	0.04	0.95	0.96	0.95	0.97
100	200	4	0.56	1.00	0.04	0.95	0.96	0.95	0.97
100	200	5	0.95	1.00	4.17	0.95	0.21	0.11	0.30
100	200	6	0.95	1.00	4.09	0.95	0.21	0.11	0.31
100	200	7	0.99	1.00	751.77	0.95	0.01	0.00	0.03

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the critical value of 1.96 in the sample of size T. the column labelled $M(j)$ reports the frequency that $\max_{1\le t\le T}|\widehat{\tau}_t(j)|$ exceeds the critical value for a sample of size T. CI_ε is the 95% confidence interval for ε, the measurement error. $NS(j)$ is the noise to signal ratio, see (8). R^2 is defined in (9), R^{2-} and R^{2+} are the lower and upper 95% confidence interval.

Table 1c: Tests for G_{jt}: $\widetilde{G}_t$ defined in (4)

N	T	j	$A(j)$	$M(j)$	$NS(j)$	$CI_\varepsilon(j)$	$R^2(j)$	$R^{2-}(j)$	$R^{2+}(j)$
50	50	1	0.01	0.01	0.03	0.99	0.97	0.96	0.99
50	50	2	0.01	0.01	0.03	0.99	0.97	0.95	0.99
50	50	3	0.07	0.21	0.07	0.99	0.94	0.90	0.97
50	50	4	0.07	0.20	0.07	0.99	0.94	0.90	0.97
50	50	5	0.66	1.00	5.05	0.98	0.22	0.03	0.41
50	50	6	0.66	1.00	5.08	0.98	0.22	0.03	0.41
50	50	7	0.72	1.00	339.24	0.99	0.04	0.00	0.13
100	50	1	0.04	0.08	0.01	0.96	0.99	0.98	0.99
100	50	2	0.04	0.08	0.01	0.97	0.99	0.98	0.99
100	50	3	0.22	0.63	0.05	0.97	0.95	0.92	0.98
100	50	4	0.22	0.62	0.05	0.97	0.95	0.92	0.98
100	50	5	0.77	1.00	4.79	0.98	0.23	0.04	0.42
100	50	6	0.77	1.00	4.77	0.98	0.22	0.03	0.42
100	50	7	0.80	1.00	290.41	0.97	0.04	0.00	0.14
50	100	1	0.00	0.00	0.03	1.00	0.97	0.96	0.98
50	100	2	0.00	0.00	0.03	1.00	0.97	0.96	0.98
50	100	3	0.03	0.03	0.07	1.00	0.94	0.91	0.96
50	100	4	0.03	0.03	0.07	1.00	0.94	0.91	0.96
50	100	5	0.64	1.00	4.52	0.99	0.21	0.07	0.35
50	100	6	0.64	1.00	4.49	0.99	0.21	0.07	0.34
50	100	7	0.72	1.00	1780.38	0.99	0.02	0.00	0.07
200	100	1	0.04	0.09	0.01	0.97	0.99	0.99	1.00
200	100	2	0.04	0.09	0.01	0.96	0.99	0.99	1.00
200	100	3	0.34	0.89	0.05	0.97	0.96	0.94	0.97
200	100	4	0.34	0.90	0.05	0.97	0.96	0.94	0.97
200	100	5	0.84	1.00	4.32	0.98	0.21	0.07	0.35
200	100	6	0.83	1.00	4.27	0.97	0.21	0.08	0.35
200	100	7	0.86	1.00	431.39	0.98	0.02	0.00	0.07
100	200	1	0.00	0.00	0.01	1.00	0.99	0.98	0.99
100	200	2	0.00	0.00	0.01	1.00	0.99	0.98	0.99
100	200	3	0.08	0.29	0.05	1.00	0.95	0.94	0.96
100	200	4	0.08	0.26	0.05	1.00	0.95	0.94	0.96
100	200	5	0.74	1.00	4.23	0.99	0.20	0.11	0.30
100	200	6	0.74	1.00	4.15	1.00	0.21	0.11	0.30
100	200	7	0.81	1.00	946.21	0.99	0.01	0.00	0.03

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the critical value of 1.96 in the sample of size T. the column labelled $M(j)$ reports the frequency that $\max_{1\leq t\leq T}|\widehat{\tau}_t(j)|$ exceeds the critical value for a sample of size T. CI_ε is the 95% confidence interval for ε, the measurement error. $NS(j)$ is the noise to signal ratio, see (8). R^2 is defined in (9), R^{2-} and R^{2+} are the lower and upper 95% confidence interval.

Table 2: Testing G_t Jointly:

	T	Set	$\widehat{\rho}_s^2$	$\widehat{\rho}_s^{2-}$	$\widehat{\rho}_s^{2-}$
50	50	1	0.85	0.78	0.93
50	50	2	0.09	0.00	0.23
50	50	3	0.96	0.94	0.98
50	50	4	0.03	0.00	0.13
100	50	1	0.87	0.80	0.94
100	50	2	0.10	0.00	0.24
100	50	3	0.98	0.97	0.99
100	50	4	0.04	0.00	0.13
50	100	1	0.85	0.80	0.91
50	100	2	0.08	0.00	0.18
50	100	3	0.97	0.95	0.98
50	100	4	0.02	0.00	0.06
200	100	1	0.88	0.83	0.92
200	100	2	0.08	0.00	0.18
200	100	3	0.99	0.99	1.00
200	100	4	0.02	0.00	0.06
100	200	1	0.87	0.83	0.90
100	200	2	0.07	0.01	0.14
100	200	3	0.98	0.98	0.99
100	200	4	0.01	0.00	0.03

$\widehat{\rho}_s^2$ is the smallest non-zero canonical correlation between $\widetilde{F}$ and G, $\widehat{\rho}_s^{2-}$ and ρ_s^{2+} define the 95% confidence interval.

Table 3: Testing the Factors in 100 FF Portfolios: Annual Data

Sample	j	$A(j)$	$M(j)$	$R^2(j)$	$NS(j)$	$\widehat{\rho}(k)^2$
60-96	Market	0.270	4.343	0.984 (0.974, 0.994)	0.016	0.995 (0.992, 0.998)
T=37	SMB	0.162	2.865	0.971 (0.952, 0.989)	0.030	0.970 (0.951, 0.989)
N=94	HML	0.081	2.479	0.962 (0.938, 0.986)	0.039	0.933 (0.891, 0.975)
	DC	0.730	20.216	0.238 (0.000, 0.478)	3.199	0.192 (0.000, 0.420)
	DIP	0.865	33.642	0.146 (0.000, 0.356)	5.845	0.149 (0.000, 0.360)
	DP	0.649	19.401	0.198 (0.000, 0.428)	4.047	0.111 (0.000, 0.302)
	TERM	0.838	16.800	0.249 (0.007, 0.490)	3.021	0.000 (0.000, 0.000)
	RISK	0.730	21.169	0.313 (0.066, 0.561)	2.190	0.000 (0.000, 0.000)

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the 5% asymptotic critical value. $M(j)$ is the value of the test. R^2 is defined in (9), NS(j) defined in (8), and $\widehat{\rho}(k)^2$ is vector of canonical correlations of G_t with respect to F_t. $\widetilde{\Gamma}_t$ is defined in (5).

Table 4: Testing the Factors in 100 FF Portfolios: Monthly Data

Sample	j	$A(j)$	$M(j)$	$R^2(j)$	$NS(j)$	$\widehat{\rho}(k)^2$
60-96	Market	0.288	9.847	0.973 (0.968, 0.978)	0.028	0.992 (0.990, 0.993)
T=444	SMB	0.259	7.132	0.926 (0.913, 0.940)	0.079	0.917 (0.902, 0.932)
N=95	HML	0.182	5.351	0.886 (0.867, 0.906)	0.128	0.832 (0.804, 0.861)
	DC	0.928	103.708	0.035 (0.001, 0.068)	27.913	0.039 (0.004, 0.074)
	DIP	0.901	85.148	0.043 (0.006, 0.080)	22.158	0.019 (0.000, 0.044)
	DP	0.957	155.614	0.011 (0.000, 0.031)	86.468	0.005 (0.000, 0.019)
	TERM	0.919	115.840	0.040 (0.004, 0.076)	23.881	0.000 (0.000, 0.000)
	RISK	0.935	149.179	0.075 (0.028, 0.121)	12.423	0.000 (0.000, 0.000)
60-82	Market	0.380	9.041	0.965 (0.957, 0.973)	0.036	0.992 (0.991, 0.994)
T=276	SMB	0.341	6.488	0.913 (0.893, 0.933)	0.095	0.904 (0.883, 0.926)
N=95	HML	0.178	4.556	0.887 (0.862, 0.912)	0.128	0.792 (0.749, 0.836)
	DC	0.895	61.140	0.059 (0.005, 0.113)	15.844	0.136 (0.061, 0.211)
	DIP	0.906	47.236	0.087 (0.024, 0.151)	10.450	0.024 (0.012, 0.059)
	DP	0.924	121.266	0.051 (0.000, 0.101)	18.658	0.015 (0.000, 0.043)
	TERM	0.891	70.307	0.041 (0.000, 0.087)	23.271	0.000 (0.000, 0.000)
	RISK	0.913	74.944	0.097 (0.031, 0.164)	9.258	0.000 (0.000, 0.000)
83-96	Market	0.238	5.098	0.982 (0.977, 0.987)	0.018	0.993 (0.991, 0.995)
T=168	SMB	0.208	4.034	0.940 (0.923, 0.958)	0.063	0.922 (0.899, 0.944)
N=101	HML	0.190	4.340	0.917 (0.893, 0.941)	0.090	0.898 (0.868, 0.927)
	DC	0.857	47.584	0.090 (0.008, 0.173)	10.072	0.244 (0.131, 0.356)
	DIP	0.750	26.945	0.143 (0.045, 0.241)	6.003	0.086 (0.005, 0.168)
	DP	0.774	43.981	0.125 (0.032, 0.219)	6.971	0.016 (0.000, 0.053)
	TERM	0.881	45.842	0.070 (0.000, 0.144)	13.322	0.000 (0.000, 0.000)
	RISK	0.661	33.635	0.309 (0.193, 0.426)	2.231	0.000 (0.000, 0.000)
73-87	Market	0.383	9.075	0.970 (0.962, 0.979)	0.030	0.992 (0.990, 0.994)
T=180	SMB	0.428	6.840	0.903 (0.876, 0.930)	0.108	0.925 (0.904, 0.946)
N=98	HML	0.200	4.287	0.912 (0.888, 0.937)	0.096	0.804 (0.753, 0.856)
	DC	0.856	58.510	0.193 (0.089, 0.297)	4.183	0.276 (0.165, 0.387)
	DIP	0.917	71.336	0.045 (0.000, 0.104)	21.300	0.049 (0.000, 0.111)
	DP	0.817	48.647	0.147 (0.052, 0.243)	5.799	0.007 (0.000, 0.031)
	TERM	0.950	64.405	0.035 (0.000, 0.088)	27.622	0.000 (0.000, 0.000)
	RISK	0.939	66.935	0.036 (0.000, 0.090)	26.425	0.000 (0.000, 0.000)
88-96	Market	0.176	4.581	0.983 (0.977, 0.990)	0.017	0.991 (0.987, 0.994)
T=108	SMB	0.046	3.568	0.962 (0.948, 0.976)	0.039	0.960 (0.945, 0.975)
N=101	HML	0.148	4.075	0.909 (0.877, 0.942)	0.100	0.907 (0.873, 0.940)
	DC	0.750	21.920	0.130 (0.012, 0.248)	6.693	0.397 (0.253, 0.540)
	DIP	0.685	24.591	0.133 (0.014, 0.252)	6.530	0.081 (0.000, 0.179)
	DP	0.611	19.617	0.269 (0.126, 0.412)	2.715	0.016 (0.000, 0.063)
	TERM	0.833	44.342	0.079 (0.000, 0.177)	11.637	0.000 (0.000, 0.000)
	RISK	0.509	22.147	0.318 (0.173, 0.463)	2.147	0.000 (0.000, 0.000)

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the 5% asymptotic critical value. $M(j)$ is the value of the test. R^2 is defined in (9), NS(j) defined in (8), and $\widehat{\rho}(k)^2$ is vector of canonical correlations of G_t with respect to F_t. $\widetilde{\Gamma}_t$ is defined in (5).

Table 5: Testing the Factors in Monthly CRSP Returns

Sample	j	$A(j)$	$M(j)$	$R^2(j)$	$NS(j)$	$\widehat{\rho}(k)^2$
60-96	Market	0.255	6.685	0.967 (0.961, 0.973)	0.034	0.972 (0.966, 0.977)
T=444	SMB	0.468	9.677	0.603 (0.545, 0.660)	0.660	0.588 (0.529, 0.647)
N=190	HML	0.628	14.925	0.436 (0.367, 0.505)	1.294	0.359 (0.288, 0.431)
	DC	0.869	45.269	0.050 (0.010, 0.089)	19.046	0.042 (0.005, 0.078)
	DIP	0.883	70.418	0.032 (0.000, 0.065)	29.897	0.005 (0.000, 0.017)
	DP	0.919	87.720	0.021 (0.000, 0.048)	46.206	0.001 (0.000, 0.007)
	TERM	0.955	105.516	0.011 (0.000, 0.030)	90.426	0.000 (0.000, 0.000)
	RISK	0.899	63.344	0.059 (0.017, 0.102)	15.880	0.000 (0.000, 0.000)
60-82	Market	0.264	6.299	0.969 (0.961, 0.976)	0.032	0.974 (0.968, 0.980)
T=276	SMB	0.471	7.261	0.664 (0.600, 0.729)	0.505	0.635 (0.567, 0.704)
N=190	HML	0.562	11.003	0.525 (0.444, 0.606)	0.904	0.448 (0.361, 0.536)
	DC	0.880	32.922	0.085 (0.022, 0.147)	10.818	0.069 (0.011, 0.126)
	DIP	0.895	60.978	0.033 (0.000, 0.074)	29.732	0.012 (0.000, 0.038)
	DP	0.924	66.766	0.026 (0.000, 0.063)	37.155	0.006 (0.000, 0.024)
	TERM	0.924	63.639	0.025 (0.000, 0.062)	38.283	0.000 (0.000, 0.000)
	RISK	0.891	54.577	0.064 (0.008, 0.120)	14.625	0.000 (0.000, 0.000)
83-96	Market	0.286	4.747	0.967 (0.957, 0.977)	0.034	0.972 (0.963, 0.980)
T=168	SMB	0.530	10.600	0.556 (0.456, 0.656)	0.799	0.582 (0.486, 0.678)
N=190	HML	0.619	17.831	0.464 (0.354, 0.574)	1.155	0.390 (0.275, 0.506)
	DC	0.940	71.082	0.016 (0.000, 0.054)	60.212	0.026 (0.000, 0.073)
	DIP	0.911	88.930	0.021 (0.000, 0.064)	47.009	0.017 (0.000, 0.055)
	DP	0.952	123.024	0.008 (0.000, 0.035)	120.873	0.004 (0.000, 0.021)
	TERM	0.917	67.360	0.024 (0.000, 0.070)	40.716	0.000 (0.000, 0.000)
	RISK	0.714	37.568	0.101 (0.014, 0.187)	8.929	0.000 (0.000, 0.000)
73-87	Market	0.256	6.538	0.973 (0.965, 0.981)	0.028	0.976 (0.969, 0.983)
T=180	SMB	0.467	9.064	0.662 (0.582, 0.742)	0.510	0.661 (0.581, 0.742)
N=190	HML	0.656	16.801	0.553 (0.456, 0.650)	0.808	0.417 (0.307, 0.527)
	DC	0.850	54.418	0.107 (0.021, 0.192)	8.367	0.093 (0.012, 0.174)
	DIP	0.911	50.229	0.047 (0.000 0.107)	20.349	0.019 (0.000, 0.058)
	DP	0.950	91.785	0.037 (0.000, 0.091)	26.128	0.004 (0.000, 0.024)
	TERM	0.967	122.675	0.009 (0.000, 0.036)	112.690	0.000 (0.000, 0.000)
	RISK	0.928	55.851	0.039 (0.000, 0.095)	24.385	0.000 (0.000, 0.000)
88-96	Market	0.250	4.272	0.963 (0.950, 0.977)	0.038	0.967 (0.955, 0.979)
T=108	SMB	0.583	10.659	0.551 (0.425, 0.677)	0.815	0.621 (0.508, 0.733)
N=190	HML	0.648	18.149	0.437 (0.296, 0.577)	1.290	0.369 (0.225, 0.514)
	DC	0.898	62.467	0.034 (0.000, 0.101)	28.662	0.058 (0.000, 0.144)
	DIP	0.880	54.120	0.031 (0.000, 0.095)	31.255	0.017 (0.000, 0.066)
	DP	0.926	58.978	0.044 (0.000, 0.120)	21.639	0.010 (0.000, 0.047)
	TERM	0.870	51.365	0.034 (0.000, 0.102)	28.268	0.000 (0.000, 0.000)
	RISK	0.750	35.048	0.137 (0.017, 0.258)	6.274	0.000 (0.000, 0.000)

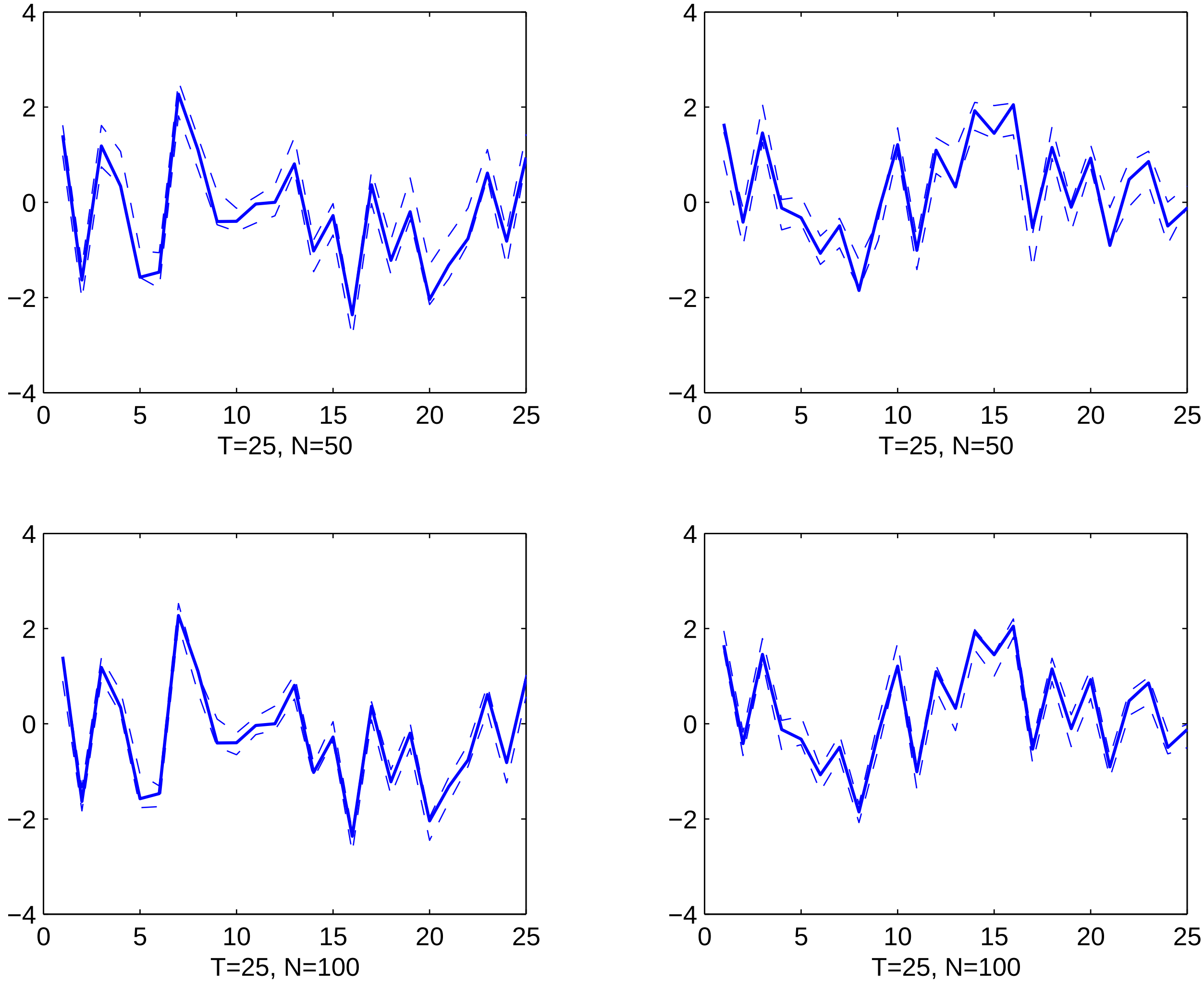
$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the 5% asymptotic critical value. $M(j)$ is the value of the test. R^2 is defined in (9), NS(j) defined in (8), and $\widehat{\rho}(k)^2$ is vector of canonical correlations of G_t with respect to F_t. $\widetilde{\Gamma}_t$ is defined in (5).

Table 6: Testing the Factors in Macroeconomic Data

Sample	j	$A(j)$	$M(j)$	$R^2(j)$	$NS(j)$	$\widehat{\rho}(k)^2$
60-96	Market	0.941	364.555	0.008 (0.000, 0.024)	124.359	0.624 (0.568, 0.679)
T=444	SMB	0.917	380.642	0.019 (0.000, 0.045)	50.978	0.335 (0.263, 0.406)
N=150	HML	0.926	466.551	0.010 (0.000, 0.029)	97.817	0.093 (0.042, 0.144)
	DC	0.833	28.531	0.283 (0.212, 0.354)	2.535	0.065 (0.021, 0.110)
	DIP	0.700	22.855	0.442 (0.373, 0.511)	1.261	0.021 (0.000, 0.048)
	DP	0.773	28.293	0.258 (0.188, 0.328)	2.879	0.002 (0.000, 0.009)
	TERM	0.768	36.474	0.274 (0.204, 0.345)	2.644	0.000 (0.000, 0.000)
	RISK	0.739	22.111	0.463 (0.395, 0.531)	1.162	0.000 (0.000, 0.000)
	UR	0.604	60.188	0.353 (0.281, 0.424)	1.835	0.000 (0.000, 0.000)
60-82	Market	0.924	337.746	0.011 (0.000, 0.035)	91.476	0.617 (0.547, 0.688)
T=276	SMB	0.906	230.985	0.023 (0.000, 0.059)	41.760	0.385 (0.294, 0.475)
N=150	HML	0.938	456.227	0.011 (0.000, 0.035)	92.165	0.113 (0.043, 0.184)
	DC	0.826	32.509	0.267 (0.178, 0.356)	2.746	0.056 (0.003, 0.109)
	DIP	0.717	29.180	0.400 (0.311, 0.490)	1.500	0.010 (0.000, 0.032)
	DP	0.750	28.225	0.290 (0.200, 0.380)	2.449	0.006 (0.000, 0.024)
	TERM	0.681	34.419	0.368 (0.278, 0.459)	1.716	0.000 (0.000, 0.000)
	RISK	0.754	26.045	0.466 (0.380, 0.552)	1.145	0.000 (0.000, 0.000)
	UR	0.663	43.591	0.325 (0.234, 0.416)	2.078	0.000 (0.000, 0.000)
83-96	Market	0.917	78.037	0.075 (0.000, 0.152)	12.277	0.728 (0.658, 0.798)
T=168	SMB	0.839	40.687	0.073 (0.000, 0.149)	12.683	0.581 (0.484, 0.677)
N=150	HML	0.905	136.574	0.036 (0.000, 0.091)	27.151	0.248 (0.135, 0.361)
	DC	0.899	53.801	0.207 (0.098, 0.316)	3.832	0.111 (0.022, 0.201)
	DIP	0.565	15.231	0.544 (0.442, 0.645)	0.839	0.043 (0.000, 0.103)
	DP	0.869	74.283	0.133 (0.037, 0.228)	6.538	0.013 (0.000, 0.048)
	TERM	0.762	19.521	0.383 (0.268, 0.499)	1.611	0.000 (0.000, 0.000)
	RISK	0.708	35.981	0.456 (0.345, 0.567)	1.191	0.000 (0.000, 0.000)
	UR	0.685	34.315	0.542 (0.440, 0.644)	0.844	0.000 (0.000, 0.000)
73-96	Market	0.924	90.926	0.037 (0.00, 0.079)	26.279	0.641 (0.575, 0.708)
T=288	SMB	0.903	162.320	0.065 (0.010, 0.120)	14.472	0.477 (0.393, 0.560)
N=150	HML	0.944	311.158	0.024 (0.000, 0.058)	41.259	0.225 (0.140, 0.309)
	DC	0.792	31.037	0.418 (0.332, 0.505)	1.390	0.113 (0.044, 0.182)
	DIP	0.670	19.184	0.510 (0.429, 0.591)	0.961	0.043 (0.000, 0.088)
	DP	0.753	30.671	0.353 (0.264, 0.441)	1.837	0.007 (0.000, 0.027)
	TERM	0.806	36.448	0.254 (0.167, 0.340)	2.942	0.000 (0.000, 0.000)
	RISK	0.771	43.639	0.334 (0.245, 0.423)	1.991	0.000 (0.000, 0.000)
	UR	0.795	98.987	0.209 (0.125, 0.292)	3.787	0.000 (0.000, 0.000)

$A(j)$ is the frequency that $|\widehat{\tau}_t(j)|$ exceeds the 5% asymptotic critical value. $M(j)$ is the value of the test. R^2 is defined in (9), NS(j) defined in (8), and $\widehat{\rho}(k)^2$ is vector of canonical correlations of G_t with respect to F_t. $\widetilde{\Gamma}_t$ is defined in (5).

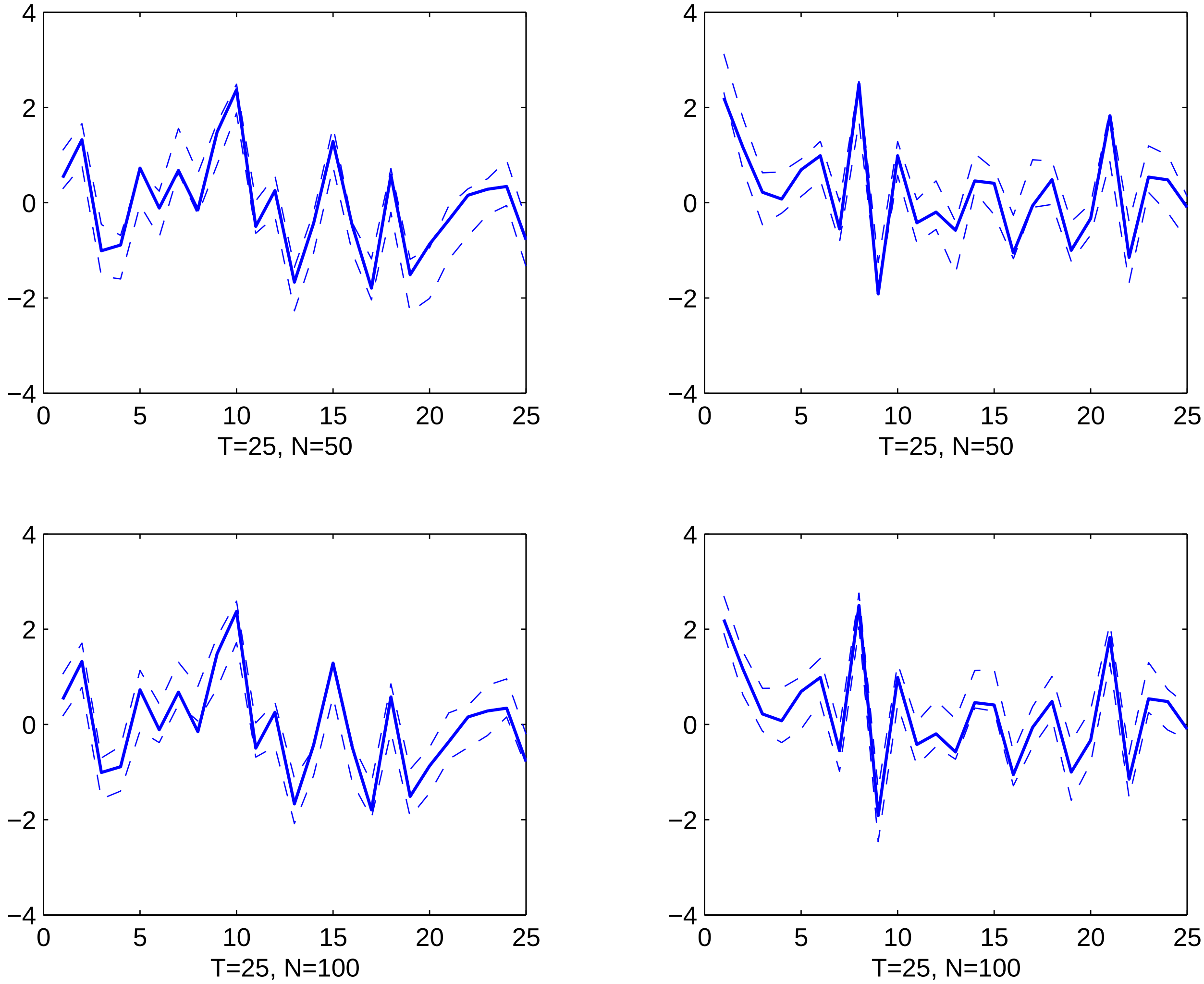
Figure 1: Factor Processes and Their Confidence Intervals



Solid Line: F_t

Dotted Lines: $\widehat{G}_t \pm 1.96\widehat{\text{var}}(\widehat{G}_{jt})$.

Figure 2: Figure 2: Measurement Errors and Their Confidence intervals



Solid Line: ε
Dotted Lines: $\widehat{\varepsilon} \pm 1.96\widehat{s}_{jt}$.