A Protocol for Factor Identification

Kuntara Pukthuanthong
University of Missouri

Richard Roll
California Institute of Technology

Avanidhar Subrahmanyam
Anderson School, UCLA

We propose a protocol for identifying genuine risk factors. A genuine risk factor must be related to the covariance matrix of returns, must be priced in the cross-section of returns, and should yield a reward-to-risk ratio that is reasonable enough to be consistent with risk pricing. A market factor, a profitability factor, and traded versions of macroeconomic factors pass our protocol, but many characteristic-based factors do not. Several of the underlying characteristics, however, do command premiums in the cross-section. (*JEL* G1, G11, G12)

Received July 20, 2017; editorial decision June 9, 2018 by Editor Andrew Karolyi. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

A factor model for asset returns can be expressed as

$$\mathbf{R}_{t} = E_{t-1}(\tilde{\mathbf{R}}_{t}) + \boldsymbol{\beta}_{t-1}\mathbf{f}_{t} + \boldsymbol{\gamma}_{t-1}\mathbf{g}_{t} + \boldsymbol{\varepsilon}_{t}, \tag{1}$$

where $\mathbf{R}_t$ is an column vector of securities' returns in period t.[1] Let N denote the number of securities. Suppressing time subscripts, assume that stochastic factors **f** command risk premiums and stochastic factors **g** do not. Given K risk

For insightful and constructive comments, we thank two anonymous referees; Andrew Karolyi (the editor), Dave Berger, Michael Brennan, Stephen Brown, Daniel Carvalho, Eric de Bodt, Wayne Ferson, Stefano Gubellinni, Campbell Harvey, Yan Liu, Peter Pope, Stephen Ross, Eduardo Schwartz, Kenneth Singleton, Ivo Welch, and Russ Wermers; and seminar participants at the 2012 Inquire UK conference, the 2013 Southwestern Finance Association annual meeting, the 2013 UCLA Brown Bag Seminar, the 2013 Rotman ICPM Seminar, the 2013 Australian Finance Conference, the 2014 Asian Finance Association Conference in Bali, the 2016 Dauphine-Amundi Conference in Paris, the 2016 Behavioral Finance Conference at the University of Miami, the 2016 Q-Group Conference in Phoenix, and the 2017 UBS/Maryland Conference in New York. We owe a great debt of gratitude to Olivier Ledoit for generous and unstinting guidance. We are thankful for financial support from Inquire UK, the Dauphine-Amundi Chair in Asset Management, the Rotman International Centre for Pension Management (ICPM) at the University of Toronto, and the Q-Group's 2016 Jack Treynor Award. Supplementary data can be found on *The Review of Financial Studies* Web site. Send correspondence to Send correspondence to Kuntara Pukthuanthong, University of Missouri Cornell, Columbia, MO 65211; telephone: 619-807-6124. E-mail: pukthuanthongk@missouri.edu.

[1] Hereafter, boldface indicates a vector or a matrix.

doi:10.1093/rfs/hhy093 Advance Access publication August 25, 2018

factors, **f** is a KX1 mean zero column vector; the true risk factor loadings are in a matrix, $\boldsymbol{\beta}$, with N rows and K columns. Similarly, if there are J unpriced factors, **g** is a JX1 mean zero column vector and the associated loadings, $\boldsymbol{\gamma}$, is a matrix with N rows and J columns. Finally, $\boldsymbol{\varepsilon}$ is an idiosyncratic NX1 mean zero column vector whose covariance matrix is diagonal. Notice that the loadings of both the true risk factors and the unpriced factors have time subscripts $t-1$ to allow for time variation. The loadings are assumed to be known one period in advance of the returns.

The expected returns as of $t-1$ conform to their own linear cross-sectional relation:[2]

$$\mathbf{E_{t-1}(R_t)=R_{F,t-1}}+\boldsymbol{\beta}_{\mathbf{t-1}}\boldsymbol{\lambda}_{\mathbf{t-1}}, \tag{2}$$

where the first term on the right is an NX1 column vector with the riskless rate at the beginning of the period in every position, $\boldsymbol{\lambda}$ is a possibly time-varying KX1 column vector of nonzero risk premiums corresponding to factor class **f**.[3] This implies that the factor set **g** is not priced in the cross-section of assets.

Empirically, how should one determine whether a particular candidate factor is in the set **f**, the set **g**, or neither of these sets? The job of any such procedure for factor identification should be to ascertain whether a particular factor candidate is in the **f** class and hence is unpredictable; is related to systematic volatility and has an associated risk premium; or is in the **g** class and hence is unpredictable; is related to volatility but does not earn a risk premium; or is neither priced nor related to asset volatility. A principal goal of our paper is to present a protocol for identifying whether a particular proposed factor is indeed a priced risk factor, that is, belongs to class **f**.

Note that Equation (2) holds in a market where arbitrage is perfect and assets are not mispriced because of behavioral biases and arbitrage constraints. If asset mispricing is allowed, then deviations from Equation (2) are permissible, and such deviation will be associated with "characteristics" that proxy for investor biases. Indeed, numerous factor candidates and firm-specific return predictors (characteristics) have been proposed in a voluminous literature. For example, Lewellen, Nagel, and Shanken (2010, p. 175) list several of the most prominent predictor candidates in their opening paragraph and note that although they explain some empirical regularities, they have "… little in common economically with each other." Subrahmanyam (2010) surveys more than fifty characteristics that various papers contend to be cross-sectionally related to mean returns. McLean and Pontiff (2016) examine 95 characteristics that were claimed in previous papers to explain returns cross-sectionally but find that predictability declines after publication. Lewellen (2015) finds

2 Please see Ross (1976) for an approximate version of this relation and Connor and Korajczyk (1988) for an equilibrium version.

3 The arbitrage pricing theory represented by (2) holds exactly in an economy with infinitely many assets and approximately otherwise.

strong predictive power of actual returns using 15 firm characteristics. Harvey, Liu, and Zhu (2016) enumerate 316 "factor" candidates suggested in 313 papers and suggest that any newly proposed factor should have to pass a much higher hurdle for statistical significance than the level habitually used in the literature, simply because of the extensive data mining. However, they do not attempt to relate the "factors" to the covariance matrix of returns, and do not draw a sharp distinction between firm-specific return predictors (characteristics) and priced factors. Green, Hand, and Zhang (2013) identify 330 firm characteristics, and Green, Hand, and Zhang (2017) test whether 100 of them are priced (i.e., are associated with risk premiums.) They find that only 24 characteristics are priced with an absolute t-statistic ≥ 3.0.

Something needs to be done when more than 300 candidates have been suggested in the factor literature and confusion arises between priced "factors" and predictor "characteristics." New return predictors seem to be proposed in every issue of the major finance journals, adding to the existing ones, but there is no well-accepted process for determining their qualities. In addition, sometimes characteristic predictors are converted to their factor counterparts by computing the return differential across long-short decile portfolios formed based on the extreme values of the characteristics (Fama and French 1993, 2008). At this point, no protocol has been proposed in the literature to separately classify priced factors and nonpriced factors. We need a process to evaluate them and to assess each additional predictor that will be inevitably nominated in the future.

Few topics in finance, arguably none, are more important than factor identification, because factors are the main principal determinants of investment performance and risk. Indeed, the comparative values of well-diversified portfolios are almost completely determined by their factor exposures. Whether or not investors know it, every diversified portfolio is completely determined by factor drivers. Moreover, there seem to be more than one of them.

The multiplicity of factors is strongly suggested by two striking empirical regularities about portfolios. First, even really well-diversified portfolios are quite volatile. The volatility of a large positively weighted portfolio is often around half as large as the average volatility of its constituents. For example, during the decade from 2001 through 2010, the monthly total return on the S&P 500 had an annualized volatility (standard deviation) of 16.3%. Over the same period, the average volatility for the S&P's constituents was 36.1%.

Second, although well-diversified portfolios are highly correlated within the same asset class, they are much less correlated across classes, for example, across bond versus equities versus commodities or across countries or across industry sectors. From 2001 through 2010, the monthly total return correlation between the S&P 500 and Barclay's Bond Aggregate Index was −0.0426. The return correlations between these two indexes and the Goldman Sachs Commodity index were 0.266 and 0.0113, respectively. Similarly, modest correlations are typical between real estate assets and assets in other classes.[4]

The first empirical fact indicates the existence of at least one common underlying systematic influence (or "risk driver" or "factor") that limit diversification within an asset class; otherwise, diversified portfolios would have much smaller volatilities. The second fact implies the presence of multiple systematic factors across assets; otherwise, diversified portfolios would be more correlated across asset classes, countries, and sectors.

Almost all academics and probably the vast majority of finance professionals now recognize that pervasive factors are among the main drivers of observed returns, but there is considerable disagreement about the identities of factors and even about whether they represent risks, anomalies, or something else.

Theory suggests that a true risk factor (in the class **f** in Equation (1)) has three fundamental attributes:

1. It varies unpredictably in a time-series sense.
2. Its variations induce changes in asset prices.
3. It is associated with a risk premium.

Quasi factors (in the set **g**) influence the returns of few securities and are unpriced in aggregate. A factor of this type possesses two attributes:

1. It varies unpredictably in a time-series sense.
2. Its variations do not affect expected returns.

Characteristics are sometimes associated with factors, but a characteristic

1. is known in advance.
2. might be cross-sectionally related to the expected returns of some assets.
3. might be cross-sectionally related to the loadings on true risk factors or the loadings on quasi factors.

Our main goal is to popularize a process to identify factors that will be broadly acceptable to both scholars and practitioners. We believe this is the first attempt to suggest a complete normative process for dealing with one of the most fundamental questions in finance: how to identify systematic risk factors that are reliably associated with expected returns. Our protocol has the potential to identify factors associated with risk premiums or true factors, but also factors that move some returns but do not have associated risk premiums, and characteristics that are associated with systematic return differences but are not related to risk. Characteristics that are reliably associated with returns, but

[4] Cotter and Roll (2015) report that real estate investment trusts have rather low betas against the S&P 500.

not risks, are perhaps the most interesting of all, because they offer potential profit opportunities.[5]

The protocol comprises a sequence of steps to check necessary conditions and one final step that examines a sufficient condition. From Equation (1) (suppressing time subscripts and assuming orthogonal factors), we have the familiar relation

$$\mathbf{cov(R)} = \boldsymbol{\beta\beta'}\mathbf{var(f)} + \boldsymbol{\gamma\gamma'}\mathbf{var(g)} + (\boldsymbol{\gamma\beta'} + \boldsymbol{\beta\gamma'})\mathbf{cov(f,g)} + \mathbf{E}(\boldsymbol{\varepsilon\varepsilon'}), \qquad (3)$$

where $\mathbf{E}(\boldsymbol{\varepsilon\varepsilon'})$ is diagonal. Equation (3) implies that a necessary condition for any factor candidate is that it must be related to the covariance matrix of returns. Although this condition for the factor existence (correlation with the assets in question) is well known and used to various extents in much of the empirical work, our protocol presents a more systematic treatment of the subject. Note that this necessary condition does not distinguish between pervasive priced factors (those with risk premiums) and unpriced factors. Our sufficient condition tests provide for this distinction, by estimating associate risk premiums, if any.

We note that our paper is not aimed at testing a particular asset-pricing model, in contrast to studies by Lewellen and Nagel (2006) and Campbell et al. (forthcoming), both of which examine the validity of the “conditional” capital asset pricing model (CAPM). For the reasons mentioned above, we think that any single-factor theory, albeit conditional, cannot explain why diversified portfolios in different asset classes are so weakly correlated.

A single stochastic discount factor (SDF) or a conditional mean/variance efficient portfolio, is always available to conditionally explain the cross-section of asset returns with a single exposure coefficient.[6] However, there is a mapping from the SDF to the factor model. To see this, let us suppress time subscripts and note that the Euler equation states the following:

$$\mathrm{E}(\mathbf{R}\mathrm{M}) = \mathbf{1},$$

where M is the SDF and $\mathbf{1}$ is the unit vector. Substituting for the factor model from (1), we have

$$\mathrm{E}(\mathbf{R}) = \mathrm{a}_1 + \mathrm{a}_2\boldsymbol{\beta} + \mathrm{k},$$

where a_1 and a_2 are constants and k=-E($\boldsymbol{\varepsilon}$M)/E(M). Now, the linear form of Equation (2) holds as long as k=0,[7] for which it suffices that E($\boldsymbol{\varepsilon}$M)=0. This is true from Equation (1) as long as the SDF M is a linear function of the factors.

[5] Engelberg, McLean, and Pontiff (forthcoming) show that profitability from anomalies is higher around earnings announcement days after controlling for risk a fact that indicates that the anomalies capture mispricing. Linnainmaa and Roberts (2017) argue that several recently discovered return predictors (“characteristics”) might be spurious because these return predictors do not survive in the data from earlier time periods.

[6] This is emphasized by Cochrane (2001) and Singleton (2006, chapter 8).

[7] With the additional observation that the risk-free rate is the reciprocal of the expected value of M (see, e.g., Campbell and Cochrane (2000) for details).

Thus, multiple factors are a practical way to explain unconditional returns over any finite sample.

From a practical perspective, either of an investor or a financial econometrician, incomplete information, via the finite sample problem, is inevitable. Our aim is to popularize an identification procedure for risk and nonrisk factors that is useful, though perhaps not theoretically pristine in the sense of being congruent with a SDF. We illustrate our protocol using popular factors, which are based on fundamentals-driven or characteristics-driven arguments.

1. Related Research

One related study is by Charoenrook and Conrad (2008) (hereafter CC.) Their approach is motivated by Section 6.3 in Cochrane (2001), which derives a relation between the conditional variance of a true factor and that factor's associated risk premium. CC notice an important implication, viz., that time variation in a factor candidate's volatility should be positively correlated with time variation in its expected return. Consequently, if (a) a proposed factor has significant intertemporal variation, (b) its mean return also has significant variation, and (c) the two are positively correlated, then the factor candidate satisfies a necessary condition to be proxying for a true underlying priced factor. As CC emphasize, such an empirical finding is not a sufficient condition.

CC find empirically that several proposed factor candidates, including size, the book-to-market ratio, and a liquidity construct, satisfy the above necessary condition. Momentum[8] does not. Momentum's estimated mean/volatility relation has the wrong sign. If this finding is upheld, it implies strongly that the momentum characteristic offers a free lunch, supposedly an arbitrage opportunity.

In the last section of their paper, CC, motivated by the recognition that their empirical condition is only necessary, examine whether the risk premiums associated with size, the book-to-market ratio, and liquidity are in a plausible range. They find that the Sharpe ratios of size and the book-to-market ratio are plausible, but the Sharpe ratio for liquidity is not. We are left in doubt as to which of these are priced factors. We note that although size, the book-to-market ratio, and liquidity satisfy a necessary condition to be risk factors, a test of sufficiency would build on their work. Also, because time variation in risk premiums is required for the CC necessary condition, a method that identifies factor candidates with stable risk premiums or with statistically small variation would also be complementary to their work.

Another related and recent paper is Harvey and Liu (2016). They propose a bootstrap method to select among a large group of candidate factors. They

[8] Carhart (1997) originally proposed this factor candidate.

ascertain the factor from a pool of candidates that yields the lowest intercept in a cross-sectional model. They then find a second factor that yields the lowest intercept from the base model that has the first successful factor. The process is repeated until no further factor passes a significant hurdle.

To illustrate, suppose they have seven candidate factors and the market factor is the first successful factor, so it is added to the model. Then, in the second round, HML is the best factor, so it is added to the model that already has the market factor. In the third round, suppose SMB is the best factor but it has an insignificant p-value. They stop at this point and declare that the market and HML are the only significant factors.

Fama and French (2018) propose squared Sharpe ratios to select factors. But an anomaly can have a high Sharpe ratio and not be risk related. Both the Harvey and Liu (2016) and the Fama and French (2018) protocols are useful, but linking the candidate factors to the sample covariance matrix is an additional step that would advance their work.

Barillas and Shanken (2018) propose a Bayesian asset pricing test that allows comparison of all possible asset pricing models from subsets of given factors. Feng, Giglio, and Xiu (2017) propose the combination of the double-selection LASSO method of Belloni et al. (2014) with two-pass regressions such as Fama-MacBeth to systematically select the best control model out of the large set of factors, while explicitly taking into account that in any finite sample we cannot be sure to have selected the correct model. Applying the key principle that true factors have to be related to the covariance matrix, which these papers do not do, would again be a useful exercise that would supplement their work.

Our protocol identifies not only factors associated with risk premiums or true factors, but also factors that move some returns but do not have associated risk premiums, and factors or characteristics that are associated with systematic return differences but not risks. Factors or characteristics that are reliably associated with returns but not risks are perhaps the most interesting of all, because they offer potential profit opportunities. Although the papers mentioned above have the same goal as ours, they do not distinguish among these categories of factors.

2. Factors and the Covariance Matrix

A necessary condition for any empirically measurable candidate (like Fama and French's (1993) HML) to be a factor is that it be related to the principal components of the covariance matrix. This condition represents the motivation for the analysis in Moskowitz (2003), who checks it for three candidates, size, the book-to-market ratio, and momentum. Moskowitz finds that size satisfies the condition; it is related to covariation and its associated risk premium is positively associated with its volatility. The book-to-market ratio is close to satisfying but momentum is not. This agrees with the results of CC discussed above in the case of momentum, and it more or less agrees with CC in the case of the book-to-market ratio.

Unfortunately, in our imperfect world, factor extraction from the covariance matrix faces a number of serious difficulties, including

a. It produces only estimates for linear combinations of the true underlying factors, not the factors themselves.
b. It is compromised by nonstationarity because there is no plausible reason the number of factors or their relative importance should be constant through time.[9]
c. It includes true risk drivers, pervasive priced factors (or linear combinations thereof) along with unpriced factors.

Fortunately, there is a remedy, perhaps imperfect, for each of these conundrums. For (a), the linear combinations extracted by PCA could be related to other candidate factors, such as macroeconomic variables, through canonical correlation or a similar method. This wouldn't prove anything but it would at least give some reassurance or raise some serious doubt. For (b), PCAs could be estimated for subperiods or with models that accommodate nonstationarity. For (c), a second-stage method like in Fama and MacBeth (1973) could be employed to distinguish priced factors from one another. Needless to say, none of these cures is without its own problems.[10]

3. What Are the Underlying Factors?

What are the salient features of factors? What are the underlying risk drivers? Cochrane (2001, p. xiv) says unequivocally, "The central and unfinished task of absolute asset pricing[11] is to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices." He particularly has in mind aggregate consumption as a driver and even goes so far as to say that "… the only consistent motivation for factor models is a belief that consumption *data* are unsatisfactory" (p. 170, emphasis in original). In other words, if we only had adequate measures of aggregate consumption, we wouldn't need much else for risk modeling. The absence of adequate consumption data motivates the study of other indicators of macroeconomic activity, even hundreds of such indicators.

[9] Moreover, it seems that nonstationarity is an empirical fact. Moskowitz (2003, p. 436) finds that "…significant time variation in the covariance structure of asset returns distorts the ability of these time-invariant factors (principal components extracted from the unconditional covariance matrix) to capture second moments, suggesting that unconditional factors miss important dynamics in return volatility.".

[10] It is known that PCA will do a rotation that makes it seem that the first factor is more dominant than in the true underlying structure. Brown (1989) offers a remedy. However, this problem is not all that troubling for our protocol because we do not need to determine the true number of underlying factors. Instead, we merely need to confirm that a factor candidate is related to some PCA extracted from the covariance matrix. Just one, the first PCA, is sufficient for a factor candidate to pass the necessary conditions.

[11] This is as opposed to relative asset pricing, such as comparing an option price to the underlying stock price.

The underlying drivers cannot be the infrequently published official numbers about macroeconomic variables because market prices move around much too rapidly. Instead, the drivers must be high-frequency changes in privately held market perceptions of pervasive macroeconomic conditions. Perceptions could include (a) rational anticipations of change in macro conditions, such as real output growth, real interest rates, inflation, or energy, that are truly pervasive and (b) behavior-driven pervasive shocks in confidence or risk perceptions, such as panics or liquidity crises.

To do a really good job, we must be able to identify and measure the pervasive factor perceptions and then to estimate factor sensitivities (betas) for every real asset. The first job is to identify and measure the factors. Existing literature has studied several alternative approaches. As discussed in the previous section, one approach relies on an entirely statistical method, such as principal components or factor analysis (e.g., Roll and Ross 1980; Connor and Korajczyk 1988.) A second approach prespecifies macroeconomic variables that seem likely to be pervasive and then prewhitens the official numbers pertaining to such low frequency constructs as industrial production, inflation, and so on (e.g., Chen, Roll, and Ross 1986). Then there is the approach of relying on asset characteristics to develop proxies that are empirically related to average returns (e.g., Fama and French 1993; Carhart 1997).

4. Putting It All Together: Linking Proposed Factors to the Covariance Matrix

Given the discussion above, we are ready to outline the first stage of our protocol for identifying factors. This stage identifies factors that move asset prices systematically, but it does not distinguish between pervasive priced factors (with risk premiums) and unpriced ones. That crucial information is postponed to a later stage. The recommended steps for this first stage are as follows.

First, collect a set of N equities for the factor candidates to explain. The test assets should belong to different industries and have enough heterogeneity so that the underlying risk premium associated factors can be detected.

Second, extract L principal components from the return series, using the asymptotic approach of Connor and Korajczyk (CK) (1988). With T time-series units up to time t, the procedure involves computing the TxT matrix $\boldsymbol{\Omega}_{\mathbf{t}} = \mathbf{(1/T)\ RR'}$, where $\mathbf{R}$ is the (de-meaned) return matrix. CK show that for large N, analyzing the eigenvectors of $\boldsymbol{\Omega}_{\mathbf{t}}$ is asymptotically equivalent to factor analysis. The first L eigenvectors of $\boldsymbol{\Omega}_{\mathbf{t}}$ form the factor estimates. The cutoff point for L $<$ N should be designated in advance; for instance, L could be chosen so that the cumulative variance explained by the principal components is at least 90%. Note that because, in most finance applications, N$>>$T, the approach has the virtue of allowing us to work with the smaller-dimension TxT matrix $\boldsymbol{\Omega}_{\mathbf{t}}$, as opposed to the traditional NxN covariance matrix used for factor analysis.

Third, collect a set of K factor candidates. These could be well-known characteristics-based candidates, such as size, the book-to-market ratio, momentum, or any of the 50 or so documented in Subrahmanyam (2010), the 316 from Harvey, Liu, and Zhu (2016), or any new candidate yet to be suggested.

Fourth, compute canonical correlations between the factor candidates and the corresponding eigenvectors from the second step. To do this, first certain matrix computations are necessary. First, using the L eigenvectors from step #2 and the K factor candidates from step #3, calculate the covariance matrix over a period up to time t, $\mathbf{V_t}$ **(L+K x L+K)**. Next, from the covariance matrix $\mathbf{V_t}$, in each period t, break out a submatrix, the cross-covariance matrix, which we denote $\mathbf{C_t}$. It has K rows and L columns (i.e., K x L); the entry in the ith row and jth column is the covariance between factor candidate i and eigenvector j. It also will be necessary to break out the covariance submatrix of the factor candidates, $\mathbf{V_{f,t}}$ **(K x K)** and the covariance submatrix of the real eigenvectors, $\mathbf{V_{e,t}}$ **(L x L)**. These computations allow us to then find two weighting column vectors, $\boldsymbol{\lambda}_t$ and $\boldsymbol{\kappa}_t$, on the factor candidates and eigenvectors, respectively ($\boldsymbol{\lambda}_t$ has K rows and $\boldsymbol{\kappa}_t$ has L rows), that maximize the correlation between the two weighted vectors. The covariance between the weighted averages of factor candidates and eigenvectors is $\boldsymbol{\lambda}_t' C_t \boldsymbol{\kappa}_t$, and their correlation is

$$\rho = \frac{\lambda_t' C_t \kappa_t}{\sqrt{\lambda_t' V_{f,t} \lambda_t \kappa_t' V_{e,t} \kappa_t}}$$

The correlation is maximized over all choices of λ_t and κ_t. The maximum occurs when the weights satisfy $\lambda_t = V_{f,t}^{-1/2} h_t$ where h_t is the eigenvector corresponding to the maximum eigenvalue in the matrix $V_{f,t}^{-1/2} C_t V_{e,t}^{-1} C_t' V_{f,t}^{-1/2}$ The vector κ_t is proportional to h_t One then maximizes the correlation again, subject to the constraint that the new vectors are orthogonal to the old one, and so on. This way, there are min(L,K) pairs of orthogonal canonical variables sorted from the highest correlation to the smallest. Each correlation can be transformed into a variable that is asymptotically distributed as chi-squared under the null hypothesis that the true correlation is zero.[12] This provides a method of testing whether the factor candidates as a group are conditionally related (on date t) to the covariance matrix of real returns (as represented by Equation (3)). Also, by examining the relative sizes of the weightings in λ_t, one can obtain an insight into which factor candidates, if any, are more related to real return covariances. We describe the latter procedure in detail within our empirical application in Section 7.

The intuition behind the canonical correlation approach is straightforward. The true underlying drivers of real returns are undoubtedly changes in perceptions about macroeconomic variables (see Section 3 above). But

12 See Anderson (1984, chap. 12) or Johnson and Wichern (2007).

the factor candidates and the eigenvectors need not be isomorphic to a particular macro variable. Instead, each candidate or eigenvector is some linear combination of all the pertinent macro variables. This is the well-known "rotation" problem in principal components or factor analysis.[13] Consequently, the best we can hope for is that some linear combination of the factor candidates is strongly related to some different linear combination of the eigenvectors that represent the true factors in Equation (1). Canonical correlation is intended for exactly this application.

Any factor candidate that does not display a significant (canonical) correlation with its associated best linear combination of eigenvectors can be rejected as a viable factor. It is not significantly associated with the covariance matrix of real asset returns.

5. Putting It All Together: Testing for Whether a Risk Factor Is Priced

In principle, the sufficiency stage of ascertaining whether factor candidates command risk premiums is easy. We simply run a pooled cross-section/time-series panel with real returns as dependent variables and betas on surviving factors as the explanatory variables, taking account of correlations across assets and time (cf. Petersen, 2009). This should be done with individual real asset returns on the left side, not with portfolio returns, because portfolios might diversify away and thus mask relevant risk- or return-related features of individual assets. Diversification into portfolios can mask cross-sectional phenomena in individual assets that are unrelated to the portfolio grouping procedure. Roll (1977) argues that the portfolio formation process makes it difficult to reject the null hypothesis of no effect on security returns. Advocates of fundamental indexation (Arnott, Hsu, and Moore 2005) argue that high market value assets are overpriced and vice versa, but any portfolio grouping by an attribute other than market value itself could diversify away such mispricing, making it undetectable.

Second, test portfolios are typically organized by firm characteristics related to average returns, for example, size and book-to-market ratio. Sorting on characteristics that are known to predict returns helps generate a reasonable variation in average returns across test assets. However, Lewellen, Nagel, and Shanken (2010) point out that sorting on characteristics also imparts a strong factor structure across test portfolios. Lewellen, Nagel, and Shanken (2010) show that even factors that are weakly correlated with the sorting characteristics would explain the differences in average returns across test portfolios regardless of the economic theories underlying the factors. They caution about the low dimensionality issue when portfolios are used, decreasing test power

[13] The rotation problem is resolved by placing restrictions on the extracted factors. In principal components, the restriction is that successive factors explain the maximum amount of remaining variance. In factor analysis, restrictions are imposed on the factor covariance matrix (e.g., it is diagonal or lower triangular.)

because there are fewer observations with portfolios than with individual assets. Lo and MacKinlay (1990) support this strand of the argument and show that, in contrast to Roll (1977), forming portfolio on characteristics makes it likely to reject the null hypothesis too often because of a "data-snooping" bias.

Third, forming portfolios might mask cross-sectional relation between average returns and factor exposures ("betas"). To illustrate, the cross-sectional relation between expected returns and betas under the single-factor CAPM holds exactly if and only if the market index used for computing betas is on the mean/variance frontier of the individual asset universe. Errors from the beta/return line, either positive or negative, imply that the index is not on the frontier. But if the individual assets are grouped into portfolios sorted by portfolio beta and the individual errors are not related to beta, the analogous line fitted to portfolio returns and betas will display much smaller errors. This could lead to a mistaken inference that the index is on the efficient frontier.

Finally, the statistical significance and economic magnitudes of risk premiums could depend on the choice of test portfolios. For example, the Fama and French size and book-to-market risk factors are significantly priced when test portfolios are sorted based on corresponding characteristics, but they do not command significant risk premiums when test portfolios are sorted only based on momentum. Brennan, Chordia, and Subrahmanyam (1998) also show different results for different sets of portfolios depending on characteristics used to form such portfolios.

The preceding discussion indicates that a properly specified regression analysis based on individual securities is more desirable than a portfolio approach to identify risk premiums. A variant of the panel approach of Petersen (2009) is standard in finance; it was originated by Fama and MacBeth (FM) (1973). The only real difficulty is that the regression betas, the factor loadings, are not known quantities and must be estimated. This implies a classic errors-in-variables (EIV) problem because the betas are the explanatory variables in each FM cross-section. Because the estimated betas inevitably contain measurement errors, the cross-sectional regressions have biased coefficients.

The error variances for individual assets are almost certainly greater than they are in betas estimate for portfolios, which explains why Fama and French (1992) use the latter. In our analysis, we adopt their procedure in using portfolios to obtain beta estimates, assigning portfolio betas to the constituent individual stocks, and then checking to see if the factor is priced via FM regressions. While this exercise is performed on the market factor in Fama and French (1992), it has not been performed consistently on other factors. Indeed, Fama and French (1993) do not perform this second stage exercise on individual securities for their SMB and HML factors.

To specifically address the EIV problem, we do a double sorting. First, we sort stocks based on size into 10 portfolios; then, within each size decile, we sort stocks into 10 portfolios by market beta. Next, we independently do the same double sorts but instead of sorting by market beta, we sort by HML beta. We do the same double sort for every factor that passes our necessary conditions. Then we assign each of these portfolio betas to the stocks that the portfolio contains. Following this assignment procedure, we consider the significance of the betas in FM regressions.

As a final check following the FM regression, we propose that for a genuine risk factor, its reward-to-risk ratio must be within reasonable limits. As an extreme example, if a candidate risk factor delivers a Sharpe ratio of three, it would be difficult to accept such a magnitude as indicative of a priced risk, given that Sharpe ratios for most well-diversified market indices are usually less than unity over periods of a decade or more (MacKinlay 1995).

Thus, we propose an investigation of whether factor-based Sharpe ratios exceed a "reasonable bound." Our bound is the one proposed by MacKinlay (1995). He argues that based on the historical mean excess return and volatility of the CRSP value-weighted index, a reasonable annualized Sharpe ratio for a risk factor is 0.6 (corresponding, e.g., to an annualized excess return of 10% and a standard deviation of 16%). We propose to test whether each individual proposed factor delivers a Sharpe ratio is statistically higher than the proposed MacKinlay bound.

6. Simulation

We perform a simulation exercise to assess whether our protocol reliably identifies risk factors that are priced cross-sectionally and explains covariation in asset returns. In the basic setup, we simulate five factors associated with risk premiums (f), three factors that do not command risk premiums (g), and five characteristics (Z). Our simulations are all based on the Gaussian cross-sectional or time-series distributions.

We use the actual mean excess return and standard deviation of the value-weighted CRSP index to generate the time-series of the first priced factor f_1. The other four priced factors are generated so that they are mutually independent and independent of f_1, from a distribution with the actual mean (0.354%) and standard deviation (2.897%) of a well-known factor, the Fama and French HML, over our sample period. Because the g factors are unpriced, each has a mean of zero. The standard deviation of each of the g's is set equal to that of HML. The five characteristics (Zs) represent firm-specific information, which varies across time; therefore, they are generated during each time period from the uniform distribution [-0.1%, 0.1%].[14]

We draw 5,000 data sets from the data-generating process (DGP). For each data, we check both necessary and significant condition of our protocol from a pool of 13 factor candidates. To make our simulation more realistic, our

[14] An alternative way to simulate characteristics would be to draw them from autocorrelated (e.g., autoregressive) processes for each firm. Because it is not immediately obvious how the autocorrelation should vary in the cross-section, we leave this exercise for future work.

candidate factors are noisy versions of f and g. In particular, we add estimation errors (or noise) η and ι to the factors f and g respectively and denote the candidate factors as $\tilde{\mathrm{f}}$ and $\tilde{\mathrm{g}}$. These errors have zero mean and a standard deviation that is parametrically varied in the simulations.[15]

More specifically, our DGP is based on the following equation:

$$R_t = r_f + \beta f_t + \gamma g_t + k Z_t + \varepsilon_t, \tag{4}$$

where R_t is a N-asset vector of returns in period t, t = 1, 2, …, T Furthermore,

- r_f is a time-invariant risk-free rate set to 0.1%.
- $f_t = (\mathrm{f}_{1,t}, \mathrm{f}_{2,t}, \mathrm{f}_{3,t}, \mathrm{f}_{4,t}, \mathrm{f}_{5,t})$ is a 5x1 vector of independent risk factors at time t.
- $\mathrm{f}_{1,t}$is normally distributed with the mean and standard deviation of $\mathrm{R_m}$-$\mathrm{R_f}$. $\mathrm{f}_{i,t}$ $i \neq 1$ is also normally distributed, but with the mean and standard deviation of the HML factor.
- β is a N x 5 asset-specific factor loading matrix for the five factors that carry a risk premium. β is independently drawn from N(1,1) for each firm.
- $g_t = (\mathrm{g}_{1,t}, \mathrm{g}_{2,t}, \mathrm{g}_{3,t})$ is a 3x1 vector of independent unpriced factors at time t. $g_{i,t}$ is normally distributed with a mean of zero and a standard deviation equal to that of HML where i = 1, 2, 3. The loadings on g, i.e, γ, are drawn in a manner similar to β.
- $Z_t = (\mathrm{Z}_{1,t}, \mathrm{Z}_{2,t}, \mathrm{Z}_{3,t}, \mathrm{Z}_{4,t}, \mathrm{Z}_{5,t})$ is a N x 5 matrix of characteristics at time t. In particular, $\mathrm{Z}_{i,t}$ is an Nx1 asset-specific vector of the i^{th} characteristic. $\mathrm{Z_{i,t}}$ is drawn at each time period t from a uniform distribution [−0.1%, 0.1%].
- k = (1, 1, 1, −1, −1) is a 5 x 1 constant vector.
- ε_t is an Nx 1 vector of the firm-specific return components. It is assumed to follow a normal distribution with mean zero. The variance of this term is chosen such that for a “typical” stock whose betas equal their cross-sectional means, the R^2 from a regression of individual stock returns on the factors equals 25%.[16]

Our factor candidates (FC) are denoted by

- $\tilde{\mathrm{f}}_{i,t} = \mathrm{f}_{i,t} + \eta_{\mathrm{i,t}}$ is a noisy factor with measurement error $\eta_{(i,t)} \sim N(\mu = 0, \sigma)$, where i=1,2,3,4,5.
- $\tilde{g}_{i,t} = g_{i,t} + \iota_{i,t}$ is a noisy factor with measurement error $\iota_{(i,t)} \sim N(\mu = 0, \sigma)$, where i=1,2,3.

[15] We have tried other DGPs for returns. Although they lead to broadly similar results, we desist from reporting results from other variants on the rationale that this exercise might obscure the nature of our protocol.

[16] This threshold corresponds to the higher end of the estimates in Campbell et al. (2001) that use but a single (market) factor; however, our results are not particularly sensitive to the specific assumption.

For each of the 5,000 drawn data sets, we apply the asymptotic approach of Connor and Korajczyk (1988) to extract 10 principal components from the simulated return series. We compute canonical correlations (i.e., ρ) between the 10 principal components and the 13 factor candidates and test the significance of these canonical correlations by the chi-squared statistic described in the fourth step of Section 4. We proceed only with statistically significant canonical correlations and test the significance of each factor candidate as follows: First, for each of the significant canonical correlations, the corresponding canonical variate, say U, is constructed from a linear combination of PCs, whose coefficients come from the weighting vector κ in the fourth step of Section 4. Second, we run a regression of U on the all factor candidates. The t-statistics from the regression then present the significance level of each factor candidate. Any given factor candidate passes the necessary condition of our protocol if the mean t-statistic over all significant canonical correlations is significant at the 5% level (i.e., exceeds 1.96).

Next, we examine the sufficient condition for factor candidates that satisfy the necessary condition. Specifically, we ask if a factor is priced cross-sectionally. We perform a Fama-MacBeth (FM) regression as follows: first, each asset return is regressed on the time series of all factor candidates surviving the necessary condition. Second, the stock return is regressed cross-sectionally on the factor loadings (or betas), resulting in a series of risk premium coefficients, say ν_{it}, for each surviving factor. The risk premium for a given factor is computed by averaging the ν_{it} over T. We test for the significance of the risk premiums via their usual time-series t-statistics (i.e., the mean of each ν_{it} times the square root of T divided by the standard deviation of ν_{it}).

We also control for asset-specific characteristics. Thus, in a variant of the FM regression above, instead of using factor loadings (or betas) alone in the sufficient condition, we also include the five characteristics Z in the DGP and compute the risk premiums for the surviving factors and the premiums for these characteristics. Table 1 presents the percentages that pass the necessary conditions and significance rate of each factor in the sufficient condition after 5,000 simulations. The number of stocks N ranges from 2000 to 6000, whereas the number of time periods T ranges from 600 to 840. We also consider two values for the standard deviation of the measurement errors in the factors (σ), 0.25% and 1.0%. The second column presents the passing percentages of factors (f and g) for the necessary condition hurdle. The third column presents the passing percentages for the sufficient condition, while including the characteristics, Z, in the monthly cross-sectional regressions.[17]

We find that the f and g factors all pass the necessary condition at rates in excess of 95%. The average pass rate for the priced f factors at the sufficient condition (pricing) stage is about 80% when the number of time periods is low

[17] The second-stage pass rates are virtually unaffected by whether the characteristics Z are included in the regression.

Table 1
Simulation results

Factor candidates	Periods of time = 600		Periods of time = 720		Periods of time=840		Periods of time=600		Periods of time=720		Periods of time=840	
	SD of measurement errors = 1.00%						SD of measurement errors = 0.25%					
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
A. Number of stocks = 2,000												
f.1	100	73	100	82	100	86	100	75	100	80	100	90
f.2	100	83	100	88	100	90	100	81	100	89	100	92
f.3	100	81	100	87	100	91	100	83	100	87	100	91
f.4	100	82	100	89	100	94	100	82	100	89	100	93
f.5	100	83	100	89	100	92	100	84	100	88	100	94
g.1	100	2.6	100	6.2	100	4.2	100	4.0	100	4.8	100	3.2
g.2	100	3.6	100	3.6	100	4.8	100	4.8	100	5.4	100	5.8
g.3	100	3.0	100	4.8	100	5.2	100	6.2	100	3.8	100	4.8
Z_1	–	98	–	100	–	99	–	99	–	99	–	100
Z_2	–	98	–	99	–	100	–	97	–	100	–	100
Z_3	–	99	–	100	–	99	–	98	–	100	–	100
Z_4	–	99	–	99	–	100	–	98	–	99	–	99
Z_5	–	98	–	99	–	100	–	99	–	99	–	100
B. Number of stocks = 3,000												
f.1	100	76	100	79	100	89	100	75	100	81	100	90
f.2	100	83	100	89	100	93	100	81	100	90	100	93
f.3	100	82	100	88	100	94	100	82	100	89	100	93
f.4	100	82	100	88	100	94	100	79	100	88	100	92
f.5	100	84	100	89	100	92	100	81	100	87	100	91
g.1	100	2.0	100	4.2	100	6.6	100	4.0	100	3.6	100	5.4
g.2	100	3.2	100	3.8	100	5.0	100	5.2	100	6.2	100	4.4
g.3	100	3.2	100	4.6	100	5.2	100	4.6	100	6.4	100	6.2
Z_1	–	100	–	100	–	100	–	100	–	100	–	100
Z_2	–	100	–	100	–	100	–	100	–	100	–	100
Z_3	–	100	–	100	–	100	–	100	–	100	–	100
Z_4	–	100	–	100	–	100	–	100	–	100	–	99
Z_5	–	100	–	100	–	100	–	100	–	100	–	100
C. Number of stocks = 6,000												
f.1	100	75	100	81	100	88	100	73	100	80	100	89
f.2	100	83	100	89	100	91	100	78	100	91	100	91
f.3	100	82	100	88	100	95	100	85	100	88	100	94
f.4	100	81	100	88	100	94	100	87	100	88	100	92
f.5	100	83	100	88	100	92	100	78	100	85	100	93
g.1	100	5.6	100	4.4	100	3.0	100	3.0	100	3.0	100	4.0
g.2	100	4.8	100	3.6	100	5.2	100	4.4	100	2.2	100	3.2
g.3	100	4.6	100	3.4	100	4.0	100	5.6	100	2.6	100	2.6
Z_1	–	100	–	100	–	100	–	100	–	100	–	100
Z_2	–	100	–	100	–	100	–	100	–	100	–	100
Z_3	–	100	–	100	–	100	–	100	–	100	–	100
Z_4	–	100	–	100	–	100	–	100	–	100	–	100
Z_5	–	100	–	100	–	100	–	100	–	100	–	100

This table presents the percentage pass rates for the necessary condition (Columns 1) and the sufficient condition (Columns 2) from 5,000 simulations. f.1 to f.5 are factors associated with risk premiums. Factors g.1, g.2, and g.3 do not command risk premiums. Z_i ($i=1,\dots,5$) represent five characteristics. f.1 has a mean and standard deviation equal to those of the excess return on the value-weighted CRSP index. The factors f.2 to f.5 are independently generated from a distribution with the actual mean and standard deviation of HML. The g factors are independent and have mean zero, with standard deviations equal to that of HML. The five characteristics represent firm-specific information that varies across time, and they are generated at each time period from the uniform distribution [−0.1%, 0.1%]. The standard deviation for the error term is computed to ensure that, for a "typical" stock whose betas equal their cross-sectional means, the R^2 from a regression of individual stock returns on the factors equals 25%. The simulation is based on 600, 720, and 840 periods of time and the standard deviation of measurement errors $\eta_{i,t}$ and $\iota_{i,t}$ of $\tilde{f}_{i,t}$ and $\tilde{g}_{i,t}$, respectively (see Section 6), of 1% and 0.25% for 2,000 stocks (panel A), 3,000 stocks (panel B), and 6,000 stocks (panel C).

but rises to well above 90% for the lower value of the measurement error as we increase the number of time periods to the upper value. The pass rate for the nonpriced g factors is only about 2%–6% (or lower) across all cases. Finally, the priced characteristics, Z, when included in the sufficient condition regressions are significant more than 95% of the time. The general finding is that the pass rate of false factors is consistently low across all cases, while the pass rate of true factors is more sensitive to the number of periods than to the number of stocks. Overall, the simulation performs satisfactorily.[18]

In the Internet Appendix, we test out other scenarios, such as how often the protocol accepts false factors formed by portfolios that randomly pick out long-short deciles from within the cross-section of stocks. The pass rate for these portfolios ranges between just 4% and 6%. In another scenario, we assume the econometrician only observes linear combinations of the true factors, and find that these linear combinations generally have pass rates exceeding 90%.

7. An Empirical Analysis

This section presents an application of the suggested protocol using simultaneous monthly return observations over a half century, 1965–2014 inclusive. The sample assets are individual U.S. equities listed on CRSP. We select stocks based on Fama and French (1992).

As candidate factors, we include the five Fama and French (2015) factors: market (Rm-Rf), SMB, HML, profitability (RMW), and investment (CMA). We also include the Carhart (1997) momentum factor (MOM), the risk-free rate (Rf),[19] a traded liquidity factor (LIQ), the factors based on short-term (monthly) and long-term reversals, ST_REV, and LT_REV. We also include the Chen, Roll, and Ross (1986) factors. Specifically, ΔDP, ΔIP, ΔTS, UNEXPI, and ΔEI are the traded versions of the Chen, Roll, and Ross (1986) factors.

We obtain Rm-Rf, SMB, HML, RMW, CMA, MOM, ST_REV, and LT_LEV from Kenneth French's data library. We construct the traded liquidity factor, and Chen, Roll, and Ross's (1986) five factors using Cooper and Priestley's (2011, henceforth CP) methodology. We first obtain the raw CRR factors as follows. The default premium, ΔDP, is the yield spread between Moody's Baa and Aaa corporate bonds. The growth rate of industrial production, ΔIP is log (IP_t) subtracted by log (IP_{t-1}), where IP_t is the index of industrial production in month t. ΔTS is the term premium defined as the yield spread between the long-term (10-year) and the shorter-term (1-year) Treasury bond. UNEXPI and ΔEI are unexpected inflation and change in expected inflation, respectively.

[18] As is the case in much of the literature (e.g., Bai 2003), we assume in the simulation that the true number of factors is known and matches the number of factor candidates. See Connor and Korajczyk (1993) and Onatski (2009) for econometric approaches to scenarios in which the number of factors is unknown.

[19] The risk-free rate proxy (obtained from Federal Reserve Bank of St. Louis) is the 3-month Treasury-bill rate. Fluctuations in this proxy can be priced when investors have longer horizons, and these fluctuations are not readily diversifiable.

Table 2
Summary statistics for the candidate factors

	Mean	Median	Sigma	Skewness	Kurtosis	Maximum	Minimum
Rm-Rf	0.492	0.840	4.513	−0.524	1.823	16.100	−23.240
SMB	0.289	0.115	3.116	0.380	3.494	19.180	−15.360
HML	0.354	0.335	2.897	0.000	2.583	13.910	−13.110
RMW	0.253	0.165	2.163	−0.403	11.22	12.190	−17.570
CMA	0.324	0.195	2.038	0.269	1.585	9.510	−6.810
Rf	0.411	0.410	0.261	0.554	0.815	1.350	0.000
MOM	0.690	0.775	4.283	−1.401	10.77	18.380	−34.580
ST_Rev	0.497	0.330	3.180	0.367	5.459	16.200	−14.580
LT_Rev	0.301	0.185	2.534	0.629	2.628	14.490	−7.790
ΔDP	−0.061	−0.066	0.267	−0.017	1.913	1.050	−1.361
ΔIP	0.375	0.419	2.893	−0.316	1.496	12.670	−12.810
ΔTS	−0.099	−0.098	0.876	0.143	0.590	3.093	−3.190
UNEXPI	−0.203	−0.203	0.783	−0.025	0.487	3.011	−3.121
ΔEI	0.012	0.011	0.118	0.062	1.401	0.575	−0.464
LIQ	1.610	2.344	16.90	−0.096	2.807	91.310	−71.283

This table provides summary statistics for candidate factor realizations expressed as percentage per month. The sample period spans 600 months from January 1965 through December 2014. See Table A1 for variable definitions.

Following Chen, Roll, and Ross (1986), we derive UNEXPI from the total seasonally adjusted consumer price index (CPI). We collect the inputs for these five factors from the Web siteof the Federal Reserve Bank of St. Louis. We obtain Pastor and Stambaugh's (2003) innovation (INNOV) series from Lubos Pastor's Web site to construct liquidity-traded portfolio. We do not apply their traded factor (VWF) because it is in fact a zero net investment portfolio formed by longing stocks with high loadings on INNOV and shorting stocks with low loadings. A properly specified factor loading requires a loading of one on the actual factor (the INNOV series) and zero on other factors. This concept is consistent with CP, so we use the CP method to construct a traded version of the Pastor and Stambaugh's factor using the INNOV series as the input.

In Table 2, we present summary statistics associated with these factor candidates. The liquidity, market, and MOM factors tend to be the most volatile, whereas Rf and the Chen, Roll, and Ross (1986) factors tend to exhibit the least variation (except for industrial production.) The liquidity, momentum, and short-term reversal factors tend to exhibit the highest mean returns. The momentum factor has negative skewness, as does the market factor.

To construct mimicking portfolios for the six factors (the five CRR-based ones and the liquidity factor), we collect the return of 50 portfolios (the returns of the 10 equally weighted size portfolios, 10 equally weighted book-to-market portfolios, 10 value-weighted momentum portfolios, 10 equally weighted investment portfolios and 10 equally weighted operating profitability portfolios) from Kenneth French's Web site. We apply CP, who adopt the (Lehmann and Modest, 1988, section II) approach as follows. Returns of each of the 50 test assets are regressed on the five CRR factors and INNOV; that is, we perform 50 time-series regressions producing a (50 x 6) matrix **B** of slope coefficients against the five CRR factors and INNOV. We generate the variance-covariance matrix of the error terms for these regressions, which is denoted as $\mathbf{V}$. The weight on the mimicking portfolios ($\mathbf{W}$), a 6 x 50 matrix, is computed as $(\mathbf{B'V^{-1}B})^{-1}\mathbf{B'V^{-1}}$ and $\mathbf{R}$, the returns of 50 portfolios, is a Tx50 matrix where T is a number of months. The return of the CRR mimicking portfolios is $\mathbf{WR'}$, a 6 x T matrix, where each row represents the relevant mimicking portfolio return over the sample period. The CP procedure thus generates mimicking portfolios for each factor, where the beta with respect to a particular factor is unity.

The next step in our protocol is to compute asymptotic principal components that represent the covariance matrix. Because of possible nonstationarity, we split the overall sample into five subsamples with 10 years each, while the first spans 7 years because one of the potential factors was unavailable for the first 3 years (1965–1967 inclusive). For each subsample, we extract 10 principal components from the return series, using the method of Connor and Korajczyk (CK) (1988). In Table 3 we present the summary statistics for the principal components for each decade of our sample period. The components tend to be skewed, but cases of negative and positive skewness are about equally common. We retain only the first 10 PCs because they account for close to 90% of the cumulative eigenvalues or the total volatility in the covariance matrix, suggesting these 10 PCs capture most of the stock variations. We admit that the number of retained PCs is somewhat arbitrary. If something is omitted, it is omitted for all stocks and should not have impact on the pattern of detected factors. Considering the average across 50 sample years of the cumulative percentage of variance explained within the estimation year, the first principal component explains about 38% of the variance and five PCs explain over 75%. Thus, this is evidence of multiple factors, not just one. We also find there is some variation in cumulative variance explained within each estimation year by the first 10 PCs from year to year, and the total of variance explained is 90%.

Our protocol then proceeds to calculate canonical correlations. Because we have several factor candidates, there are several pairs of canonical variates, each pair being orthogonal to the others and having a particular intercorrelation. Canonical correlation sorts these pairs from largest to smallest based on their squared correlation. Panel A of Table 4 reports, in the second and third columns, the canonical correlations for the covariance matrices, and associated z-statistics, covering 1968–1974 monthly. The next few columns provide the correlations and z-statistics for subsequent periods.

As indicated by these results, the first and largest canonical correlation is dominant. Its mean conditional value is close to unity and strongly significant. Across all subperiods, only two correlations fall below 0.2 in absolute terms. The top five canonical correlations are significant in every subperiod we consider.

Information on significant relations between factor candidates and the principal components is reported in panel B of Table 4. We use the following procedure to derive the significance levels of each factor candidate. First, for

Table 3
Summary statistics for principal components

	1965–1974				1975–1984				1985–1994				1995–2004				2005–2014			
	Median	Skewness	Max	Min	Median	Skewness	Max	Min	Median	Skewness	Max	Min	Median	Skewness	Max	Min	Median	Skewness	Max	Min
PC1	0.005	−0.053	0.258	−0.245	0.038	−1.200	0.115	−0.206	−0.002	−3.851	0.360	−0.733	0.002	0.477	0.387	−0.263	−0.004	0.087	0.358	−0.389
PC2	−0.005	0.327	0.359	−0.311	0.004	−0.784	0.313	−0.460	−0.009	3.678	0.657	−0.171	0.012	−3.941	0.208	−0.696	−0.006	−0.011	0.454	−0.433
PC3	0.019	−2.675	0.155	−0.602	0.001	2.949	0.666	−0.246	−0.009	10.063	0.968	−0.051	−0.003	3.068	0.627	−0.306	0.004	−1.214	0.548	−0.655
PC4	−0.001	0.900	0.455	−0.309	−0.004	−0.321	0.334	−0.442	−0.014	3.945	0.506	−0.168	−0.008	3.510	0.689	−0.281	0.008	−0.849	0.551	−0.593
PC5	0.002	1.157	0.533	−0.308	−0.006	0.707	0.359	−0.222	0.004	−0.244	0.304	−0.284	0.003	−3.006	0.455	−0.709	−0.005	3.182	0.619	−0.355
PC6	−0.006	0.550	0.286	−0.226	0.002	−1.608	0.328	−0.579	0.001	−2.508	0.236	−0.589	0.005	0.507	0.450	−0.268	−0.006	1.354	0.443	−0.389
PC7	0.009	−0.736	0.271	−0.343	−0.011	0.709	0.381	−0.247	0.002	−0.856	0.434	−0.503	−0.010	1.303	0.513	−0.324	−0.001	0.391	0.523	−0.503
PC8	−0.008	0.970	0.396	−0.260	−0.004	2.366	0.599	−0.271	0.001	−0.083	0.395	−0.344	−0.005	4.901	0.780	−0.268	−0.002	0.529	0.387	−0.327
PC9	0.005	−0.320	0.295	−0.326	0.001	1.807	0.630	−0.366	0.006	−0.274	0.364	−0.384	0.002	−0.932	0.305	−0.428	−0.004	−1.519	0.466	−0.620
PC10	0.007	−0.177	0.271	−0.247	−0.001	−0.102	0.297	−0.307	0.001	−0.021	0.412	−0.390	0.005	−0.825	0.430	−0.509	0.006	−0.943	0.354	−0.386

This table presents summary statistics over 600 months for principal components (PCs) extracted using the Connor and Korajczyk (CK) (1988) method. The entire data period spans January 1965 through December 2014 and includes 50 years of monthly observations. For each decade within the 50 years, the CK method is applied to all available stocks with full records, and 10 principal components are extracted. The number of stocks included is 1,259 for 1965–1974, 2,331 for 1975–1984, 2,660 for 1985–1994, 3,145 for 1995–2004, and 3,349 for 2005–2014. Each PC has a mean of exactly zero and is normalized to have the same standard deviation. The standard deviation of each PC is 0.092.

Table 4
Canonical correlations with asymptotic PCs and significance levels of factor candidates

A. Canonical correlations

	1968–1974		1975–1984		1985–1994		1995–2004		2005–2014	
Canonical variate	Canonical correlation	z-stat	Canonical correlation	z-stat	Canonical correlation	z-stat	Canonical correlation	z-stat	Canonical correlation	z-stat
1	0.999	***25.330***	0.998	***28.573***	0.995	***25.478***	0.989	***26.918***	0.997	***30.559***
2	0.967	***15.512***	0.906	***14.868***	0.869	***13.204***	0.943	***17.926***	0.913	***19.063***
3	0.885	***10.096***	0.833	***9.658***	0.727	***8.980***	0.791	***11.585***	0.858	***14.581***
4	0.779	***6.721***	0.649	***5.153***	0.705	***6.740***	0.760	***8.670***	0.756	***10.666***
5	0.641	***4.534***	0.602	***3.076***	0.566	***4.075***	0.633	***5.384***	0.746	***7.970***
6	0.560	***3.509***	0.428	0.792	0.522	***2.659***	0.549	***3.274***	0.546	***4.352***
7	0.530	***2.913***	0.409	0.091	0.422	1.142	0.387	1.510	0.514	***2.947***
8	0.508	***2.224***	0.261	−1.048	0.357	0.211	0.363	1.162	0.385	1.162
9	0.390	1.179	0.228	−0.869	0.283	−0.609	0.307	0.657	0.298	0.308
10	0.344	0.876	0.185	−0.594	0.125	−1.485	0.246	0.324	0.221	−0.060

Table 4
Continued

B. Significance levels for factor candidates

		Factor candidates													
	Rm-Rf	SMB	HML	RMW	CMA	RF	MOM	ST_ Rev	LT_ Rev	ΔDP	ΔIP	ΔTS	UNEXPI	ΔEI	LIQ
Mean_t	**5.787**	**3.921**	**2.034**	1.519	1.357	1.320	1.950	1.332	1.634	1.640	1.138	1.512	1.602	1.304	1.368
Mean_t of significant Canonical Corr	***8.649***	***5.616***	***2.924***	**2.002**	1.561	1.809	**2.455**	1.596	**2.066**	**2.141**	1.384	1.940	**2.064**	1.535	1.749
Decade #	Number of t-statistics $\geq$ 1.96 out of 10 for each decade														
1	4	4	5	4	2	3	2	3	3	2	3	4	1	1	4
2	3	2	4	1	3	1	2	3	4	1	2	2	5	1	2
3	3	4	4	2	3	1	3	1	1	2	1	3	3	2	3
4	6	4	5	3	3	2	5	3	2	5	1	4	4	3	4
5	3	2	4	3	2	4	4	4	3	2	4	2	3	3	3
Mean	3.8	3.2	4.2	3.0	2.0	2.2	3.6	2.4	2.6	3.6	1.4	2.8	3.4	1.6	2.0

This table reports canonical correlations between Factor Candidates and Principal Components. The factor candidates include the five Fama-French (2015) factors, Rm-Rf, SMB, HML, RMW, and CMA along with RF, MOM, ST_REV, LT_REV, LIQ, ΔDP, ΔIP, ΔTS, UNEXPI, and ΔEI. See Table A1 for the variable definitions. The principal components are extracted as explained in Table 3 and the text using the Connor and Korajczyk (1988) cross-sectional method. Panel A reports 10 canonical correlations for each decade, sorted in descending order by their estimated squares. Corresponding z-statistics for the correlations are also reported. Panel B summarizes significance levels for factor candidates. The following procedure is implemented to derive the significance levels of each factor candidate: First, for each canonical pair, the eigenvector weights for the 10 CK PCs are taken and the weighted average CK PC (which is the canonical variate for the 10 CK PCs that produced the canonical correlation for this particular pair) is constructed. Then a regression using each CK PC canonical variate as the dependent variable and the candidate factor realizations as 15 independent variables is run over the sample months, including 120 months for the last four decades and slightly fewer for the first decade. The absolute t-statistics from the regression give the significance level of each candidate factor. There are 10 pairs of canonical variates in each decade and a canonical correlation for each one; thus, there are a total of 50 such regressions. In panel B, the first row presents the mean absolute t-statistic over all canonical correlations. The second row reports the mean absolute t-statistic when the canonical correlation itself is statistically significant. Rows 3 to 7 give the number of significant coefficients in each decade, and row 8 reports the average over the five decades. Critical rejection levels for the t-statistic are 1.65 (10%), 1.96 (5%), and 2.59 (1%). t-statistics breaching the 5% (1%) critical level are in boldface and italics.

each of the 10 canonical pairs,[20] the eigenvector weights for the 10 CK PCs are taken and the weighted average CK PC (which is the canonical variate for the 10 CK PCs that produced the canonical correlation for this particular pair) is constructed. Then, a regression using each CK PC canonical variate as the dependent variable and the actual candidate factors values as independent variables is run over the sample months in each subperiod. The absolute t-statistics from the regression then give the significance level of each candidate factor. Because there are 10 pairs of canonical variates in each of the five subperiods and a canonical correlation for each one, there are a total of 50 such regressions. The first row presents the mean absolute t-statistic of all canonical correlations. The second row shows the mean absolute t-statistic across cases in which the canonical correlation itself is statistically significant. The fifth through ninth rows present the number of significant coefficients in each decade, and the bottom row presents its average.

Because the absolute t-statistics are always positive, we use a one-tailed cutoff. We find that the mean t-statistics for the Fama-French three factors all exceed the one-tailed 2.5% level cutoff of 1.96. Further, the mean t-statistics for momentum are significant at the 5% levels for a one-tailed test. The average number of significant t-statistics exceeds two for all factors, except CMA, ΔIP, ΔEI, and LIQ.

We adopt the following screening criteria based on Table 4: a candidate is deemed a possible risk factor if in Table 4, panel B, the average absolute t-statistic for the significant canonical correlations in the second row exceeds the one-tailed, 2.5% cutoff, and the average number of significant t-statistics (last row of Table 4, panel B) exceeds 2.5. This is a conservative threshold to ensure we do not miss a true factor at our necessary condition stage. We focus on the significant canonical correlations, rather than all canonical correlations, because insignificant CCs imply that none of the factors matter, so using them would be over-fitting. Our criteria result in nine factors passing the hurdle, including the three original Fama and French (1992) market (Rm-Rf), SMB, HML factors and one of the two new Fama and French (2015) factors, RMW, followed by MOM, LT_REV, ΔDP, ΔTS, and UNEXPI. These factors pass the screen of being materially related to the covariance matrix of returns across the subperiods we consider.[21]

8. Are the Factors Priced?

8.1 Regression analysis

The next step in the protocol is to check whether a factor candidate that passes the necessary conditions is priced in the cross-section of returns. This procedure

[20] Recall that there are min(L,K) possible pairs, and, in our application, L = 10 and K = 15.

[21] Note that the investment factor (i.e., CMA) (also considered by Hou, Xue, and Zhang (2015)) does not pass the necessary condition, indicating that it is not materially related to the covariance matrix.

is as follows: we estimate two versions of factor betas, one uncorrected, and one corrected for EIV. For the non-EIV estimation, ordinary least-squares (OLS) multiple regressions are run for each stock on the nine accepted factors using all available observations for that stock. Then, for each calendar month in the sample, January 1965 through December 2014 inclusive, we multiply regress available individual stock returns cross-sectionally on the nine OLS beta estimates.[22] The time-series averages of the cross-sectional coefficients, termed the "risk premiums," along with associated sampling statistics, are then computed.

For the EIV calculations, stocks are sorted into ten groups (deciles) by market capitalization (Size), annually at the end of each June, based on NYSE size decile breakpoints. Then within each Size decile, stocks are sorted further by the OLS betas of the first factor (Rm-Rf) into 10 deciles, thus resulting in 100 Size/first factor beta groups. Within each of the 100 groups, the equal weighted average first factor beta of the group is assigned to each stock within that group. This is repeated for each of the eight additional betas whose factors pass the necessary conditions. Hence, for each of the nine factors, the individual stock beta is replaced by the equally weighted mean beta of the size/beta sorted group to which the stock belongs. Subsequently, over each of the following 12 calendar months, July through June of the next year, all available individual stock returns are multiply regressed cross-sectionally on the nine EW mean betas assigned to that stock. This is repeated for each June in the sample period, 1965–2014; then the time-series average of the cross-sectional coefficients, that is, the risk premiums, along with associated sampling statistics are computed. There are 594 months in the time series; the first 6 months are not used because the first sort is done in June 1965 and the last sort in June 2014 has only six available subsequent months.

We also control for stock-specific characteristics corresponding to some of the risk factors (specifically, those that are in fact associated directly with characteristics). The idea is to conduct a horse race between factor betas and the characteristics in the spirit of Daniel and Titman (1997).[23] The characteristics are RetLag1 (the 1-month lagged return), Lag2_12 (the 2- to 12-month lagged return), Lag13_36 (the 13- to 36-month lagged return), Size, Book/Mkt, ProfRatio (Profitability), AssetGrth (asset growth), and Amihud's (2002) illiquidity measure. These characteristics are defined in detail in Table A1.

[22] A possible check for a priced factor is (a) whether it is related to the covariance matrix and (b) whether it has a significant mean return. However, because the risk premiums emanate via multiple regression, whereas factor mean returns are in univariate settings (b) above is not directly comparable to the regression analysis. As an example, if one factor is similar to another factor, it is of interest to determine which of them carries stronger evidence of being priced, and this is best facilitated in a multiple regression.

[23] This horse race is relevant because, as Karolyi (2016) points out, characteristics may be related to factor loadings, and, without a proper setting that includes both betas and characteristics, one may misleadingly conclude that a characteristic represents market inefficiency.

Table 5
Summary statistics for the candidate factor betas and characteristics

	Mean	Median	Sigma	Skewness	Kurtosis	Maximum	Minimum
Rm-Rf	0.927	0.929	0.031	0.026	−0.658	0.994	0.861
SMB	0.800	0.808	0.045	−0.911	1.069	0.873	0.647
HML	0.151	0.153	0.067	−0.132	−0.874	0.276	0.023
RMW	0.083	0.089	0.061	−0.118	−0.474	0.218	−0.044
MOM	−0.111	−0.108	0.031	−0.180	−1.124	−0.055	−0.170
LT_Rev	−0.020	−0.017	0.018	−0.145	−0.309	0.031	−0.060
ΔDP	−1.009	−1.028	0.379	0.162	−1.190	−0.269	−1.709
ΔTS	−0.447	−0.433	0.179	0.003	−1.145	−0.141	−0.805
UNEXPI	0.435	0.474	0.192	−0.341	−1.259	0.703	0.112
RetLag1	1.069	1.503	5.426	−0.465	3.232	27.338	−27.660
Lag2_12	0.351	0.559	1.823	−0.518	0.932	5.866	−6.483
Lag13_36	0.437	0.567	1.134	−0.596	0.689	3.252	−3.342
SizeLag1	11.88	11.62	0.904	0.580	−0.892	13.810	10.385
Book/Mkt	0.874	0.797	0.319	1.722	4.112	2.221	0.453
ProfRato	0.274	0.208	0.457	6.199	38.34	3.332	−0.124
AsstGrth	0.170	0.174	0.054	−0.115	0.390	0.405	0.035

This table provides summary statistics for betas of the nine factors that pass our necessary condition, namely, Rm-Rf, SMB, HML, RMW, MOM, LT_REV, ΔDP, ΔTS, and UNEXPI. Betas (OLS slope coefficients) are computed for each individual stock in a multiple regression of the stock's monthly return on monthly factor realizations, using all observations available for each stock. To be included, a stock must have at least 24 monthly observations. Also included are corresponding summary statistics for individual stock characteristics (defined in Table A1). The summary statistics are for the time-series of cross-sectional means for each variable (beta or characteristic). The sample period spans 564 months from January 1968 through December 2014. The first 36 months are lost because of the lagged −13 to −36 return ("Lag13-36").

Table 5 reports summary statistics for the (non-EIV-corrected) betas as well as the characteristics. Variables are first averaged cross-sectionally, then in the time series. The default premium's innovation carries the lowest (most negative) mean betas whereas SMB and the market factor have the highest (most positive) mean betas. [24] Amongst the characteristics, the 1-month lagged returns are the most volatile, and prior (2-12) returns also exhibit relatively high variation. SMB also exhibits considerable negative skewness.

Table 6 presents estimated risk premiums for both the non-EIV-corrected and EIV-corrected betas. The first two models present the FM regression for the nine factors that pass our necessary conditions. Rm-Rf, RMW, Momentum and unexpected inflation are the only factors that command a risk premium. (Note that a negative premium of unexpected inflation is expected according to the way it is defined; high unexpected inflation is an adverse event that has a downward impact on stock prices.) Following the EIV correction, the excess market return, RMW, and unexpected inflation are significant.

In Models 3 and 4, we present results from FM regressions with the nine betas on the factors that pass necessary conditions and the characteristics. The results show that Rm-Rf, HML, RMW, LT_Rev, ΔDP, UNEXPI, and *all* characteristics command significant premiums. After correcting for EIV, the same factors except LT_Rev, UNEXPI, and lag13_36 remain significant.

24 Note that the mean multivariate beta for UNEXPI is positive, a finding that is counterintuitive because one would expect inflation shocks to negatively affect stock returns. In a univariate setting, the average UNEXPI beta is indeed negative; the positive sign is the result of the interaction of this factor with others.

Further, ΔTS becomes significant in this specification. Surprisingly, the HML risk premium is negative and significant, presumably because the beta estimates of HML (and CMA) are contaminated by multicollinearity. Indeed, we find that the correlation between HML and CMA is 0.71. It is noteworthy that in the EIV-corrected regression the insignificance result for lag13_36 is consistent with that of long-term reversals.

The Amihud measure causes a loss in sample size of more than 50%. Hence, we do not report results that include this measure. However, in unreported analyses, we do run regressions with this measure (along with all of the other betas and characteristics), and find that the results remain qualitatively unchanged. These results appear in the Internet Appendix.

We note that the characteristics in general are far more significant than the betas. While we cannot dismiss the possibility that characteristics instrument for unidentified risk exposures (Berk 2000; Zhang 2005), the results are consistent with market inefficiency.

We note that there is no *major* difference between correcting for EIV and not doing so. The results for the simplest cross-sectional regression risk premium estimates match the estimates using double sorts on all candidate factors (double sorts on size and beta and then portfolio betas replacing individual stock betas.) The Fama and French (1992) method introduces its own EIV problem; using portfolio betas in place of individual stock betas. This is an EIV because the true, but unknown, individual stock beta is not used.[25]

8.2 Hedge portfolio returns

While in Section 5, we describe the desirability of using individual securities as test assets, in this subsection we perform a robustness check with the standard methodology of forming hedge portfolios. These are formed by going long (short) in the portfolios with the highest (lowest) beta in deciles.[26] Table 7, panel A, presents the hedge portfolios that are long the top decile and short the bottom decile after sorting by individual stock (EIV-corrected) betas on each factor with replacement. For example, the individual stock betas on Rm-Rf are sorted and a portfolio is formed from the stocks with the largest 10%, equal weighted, and then another portfolio is formed from the stocks with the smallest 10%. The second portfolio's return is subtracted from the first one. Then, for SMB, the same procedure is repeated. Individual SMB betas are sorted and

[25] An instrumental variables (IV) approach represents another approach for resolving the EIV problem. It is potentially implementable by choosing odd-month betas for instruments and even-month betas for regressors, or vice versa (Jegadeesh et al. forthcoming). We did not employ it here because we are wary of the "weak instrument" issue in IV analysis, which might affect odd or even beta estimates used as instruments. To avoid any potential problem in this regard, we resort to other methods.

[26] Kozak, Nagel, and Santosh (2018) argue that a return spread from sorted betas does not necessarily reflect risk pricing, because it also reflects systematic sentiment that arbitrageurs leave unexploited because of a reluctance to take on systematic risk. Viewed in this way, sentiment is simply another source of priced risk that our protocol attempts to capture.

Table 6
Estimated risk premiums for factors candidates that satisfy the necessary conditions

	No EIV correction (Model 1)		EIV correction (Model 2)		No EIV correction (Model 3)		EIV correction (Model 4)	
	Mean	t-stat	Mean	t-stat	Mean	t-stat	Mean	t-stat
Constant	0.600	***9.467***	0.603	***6.895***	2.211	***9.014***	1.829	***5.726***
Rm-Rf	0.479	**2.503**	0.435	***2.872***	0.445	**2.214**	0.347	***2.713***
SMB	0.036	0.276	0.049	0.486	−0.240	−1.823	−0.077	−0.962
HML	−0.184	−1.484	−0.127	−1.456	−0.371	***−2.843***	−0.247	***−3.613***
RMW	0.244	***2.686***	0.190	***2.919***	0.250	**2.513**	0.098	**2.057**
MOM	0.480	***2.606***	0.264	1.875	0.303	1.620	0.190	1.653
LT_Rev	−0.197	−1.827	−0.113	−1.408	−0.259	**−2.334**	−0.112	−1.645
ΔDP	0.016	1.457	0.014	1.659	0.025	**2.104**	0.017	**2.211**
ΔTS	0.049	1.316	0.058	1.952	0.063	1.576	0.065	**2.462**
UNEXPI	−0.087	***−2.702***	−0.074	***−2.801***	−0.083	**−2.365**	−0.027	−1.189
RetLag1			−0.062	***−21.89***	−0.060	***−19.07***		
Lag2_12					0.106	***9.274***	11.759	***8.632***
Lag13_36					0.038	***2.878***	2.141	1.362
SizeLag1					−0.130	***−7.444***	−0.103	***−4.100***
Book/Mkt					0.210	***6.267***	0.176	***4.256***
ProfRato					0.044	**2.038**	0.262	***5.619***
AsstGrth					−0.272	***−6.832***	−0.337	***−6.403***
RSquare	0.125		0.082		0.124		0.084	
SamplSize	4,687		2,839		2,688		1,706	

Risk premiums (expressed as a percentage per month) are estimated from cross-sectional regressions computed using individual stock returns from 1968 to 2014 as dependent variables and, as explanatory variables, decile-sorted portfolio betas of the nine factors that pass necessary conditions including Rm-Rf, SMB, HML, RMW, MOM, LT_REV, ΔDP, ΔTS, and UNEXPI in Models 1 and 2, and the nine factors that pass necessary conditions and associated characteristics in Models 3 and 4. Characteristics include RetLag1, Lag2_12, Lag13_36, Size, Book/Mkt, ProfRato, and AsstGrth. See variable definitions in Table A1. These selected factors are those that are significantly related to any canonical variate in all decades in that they have mean t-statistics in the second row of Table 4, panel B, that exceed the one-tailed, 2.5% cutoff based on the chi-squared value and an average number of significant t-statistics exceeding 2.5 (see the bottom row of Table 4, panel B). For the non-EIV estimation, OLS multiple regressions are run for each stock on all (nine) factors using all available observations for that stock. For the EIV calculations, stocks are sorted into 10 groups (deciles) by market capitalization (Size), annually at the end of each June, based on NYSE size decile breakpoints. Then within each Size decile, stocks are sorted further by the OLS betas of the first factor (Rm-Rf) into 10 deciles, thus resulting in 100 Size/first factor beta groups. Within each of the 100 groups, the equally weighted average first factor beta of the group is assigned to each stock within that group. For each of the other eight factors, this procedure is repeated independently; ultimately, each stock's beta (for all nine betas) is replaced by the equally weighted portfolio beta of the double sorted size/beta group to which the stock belongs. This same procedure is redone every June from 1965 to 2014; then cross-sectional regressions are calculated in the 12 subsequent months of individual stock returns on the double-sorted portfolios betas (6 months only after the 2014 sort.) The time-series average over all months of the cross-sectional coefficients, termed the "risk premiums," along with associated sampling statistics, is reported. Critical rejections levels for the t-statistic are 1.65 (10%), 1.96 (5%), and 2.59 (1%). t-statistics breaching the 5% (1%) critical level are in boldface and italics.

then the hedge portfolio's return comes from the largest less the smallest decile. This is repeated for each candidate factor. The results, presented in Table 7, panel A, indicate that hedge portfolios for the market and RMW factors remain significant, whereas all of the Chen, Roll, and Ross (1986) factors, namely the innovations to the default premium, term spread, and inflation, also are significant. The sign of the hedged portfolio returns for unexpected inflation is negative. This is consistent with the sign of its risk premium in Table 6.

The final step in our protocol checks for whether the Sharpe ratios generated by the factors are within reasonable magnitudes. The traded versions of the factors are zero net investment portfolios. We thus combine each of them with

Table 7
Returns of hedge portfolios associated with factors that satisfy necessary conditions

A. Hedge portfolio from 10% top and bottom

	Rm-Rf	SMB	HML	RMW	MOM	LT_Rev	ΔDP	ΔTS	UNEXPI
Mean	0.691	−0.107	−0.415	0.667	0.360	−0.035	0.504	0.875	−0.994
SD	6.559	6.411	6.264	5.873	5.604	5.029	5.335	5.663	5.943
t(Mean)	**2.502**	−0.395	−1.572	***2.698***	1.527	−0.164	**2.245**	***3.670***	***−3.971***

B. Market return plus 10% top and bottom hedge portfolio returns

		Augmented returns								
	Rm-Rf	Rm-Rf	SMB	HML	RMW	MOM	LT_Rev	ΔDP	ΔTS	UNEXPI
Mean	0.490	1.181	0.383	0.075	1.157	0.850	0.455	0.995	1.365	−0.504
SD	4.584	10.739	8.711	7.191	7.140	6.595	7.032	7.207	7.200	7.821
t(Mean)	2.539	2.612	1.045	0.249	3.849	3.062	1.538	3.277	4.503	−1.529
Sharpe	0.370	0.381	0.152	0.036	0.561	0.447	0.224	0.478	0.657	0.223
Sharpe t	***−5.275***	***−5.02***	***−10.57***	***−13.38***	−0.850	***−3.472***	***−8.811***	***−2.744***	1.225	***−8.842***

C. Hedge portfolio 30% top and bottom

	Rm-Rf	SMB	HML	RMW	MOM	LT_Rev	ΔDP	ΔTS	UNEXPI
Mean	0.368	−0.154	−0.122	0.522	0.246	0.015	0.281	0.609	−0.664
SD	4.327	4.433	4.046	3.734	3.488	2.989	3.178	3.666	3.867
t(Mean)	**2.017**	−0.824	−0.715	***3.321***	1.675	0.123	**2.102**	***3.945***	***−4.080***

D. Market return plus 30% top-and-bottom hedge portfolio return

		Augmented returns								
	Rm-Rf	Rm-Rf	SMB	HML	RMW	MOM	LT_Rev	ΔDP	ΔTS	UNEXPI
Mean	0.490	0.858	0.336	0.368	1.012	0.736	0.506	0.771	1.099	−0.174
SD	4.584	8.669	7.335	5.319	5.589	4.920	5.756	5.540	5.588	6.469
t(Mean)	2.539	2.349	1.089	1.644	4.302	3.554	2.086	3.307	4.672	−0.640
Sharpe	0.370	0.343	0.159	0.240	0.627	0.518	0.304	0.482	0.681	0.093
Sharpe t	***−5.275***	***−5.939***	***−10.41***	***−8.433***	0.596	−1.821	***−6.865***	***−2.644***	1.743	***−12.007***

This table provides summary statistics for returns (expressed as a percentage per month) of hedge portfolios associated with the nine factors that satisfy the necessary conditions in Table 4. The factors include Rm-Rf, SMB, HML, RMW, MOM, LT_REV, ΔDP, ΔTS, and UNEXPI. For each candidate factor, hedge portfolios are formed by a long position in a group of stocks with the highest betas on the factor and a short position a group with the lowest betas; this is done with replacement. Panel A (C) shows the returns of the hedge portfolios with the top and bottom deciles (the top 30% and bottom 30%). Panel B (D) show the excess returns of the market (Rm-Rf) and augmented returns, which is the hedge portfolio return added to the market excess return (Rm-Rf). The augmented return of Rm-Rf is itself appended with its own hedge portfolio. The other eight to the right of it are for the other hedge portfolios. The Sharpe ratio (Sharpe), a t-statistic of the Sharpe ratio against 0.6 (Sharpe t), and the MacKinlay (1995) threshold are also reported. Critical rejections levels for the t-statistic are 1.65 (10%), 1.96 (5%), and 2.59 (1%). t-statistics breaching the 5% (1%) critical level are in boldface and italics.

a representative long-only portfolio in order to check if the Sharpe ratio is below a reasonable bound, relative to a representative long-only equity portfolio. Accordingly, in panel B, we present the mean, standard deviation, and Sharpe ratios for a portfolio that combines the market (i.e., the value-weighted CRSP index) with the zero net investment long-short portfolio from panel A and tests whether the resultant Sharpe ratio is statistically greater than the bound of 0.6 recommended by (MacKinlay, 1995, p. 13). None of the Sharpe ratios exceed this threshold; indeed, eight of nine are below the threshold. Thus, the priced factors command Sharpe ratios of a magnitude consistent with risk-based pricing.

Panels C and D repeat the analyses of panels A and B, except that the long-short portfolios are based on the top (bottom) 30% of stocks with the highest and lowest betas, instead of the highest and lowest deciles of betas. The results are similar. In panel D, almost all of the Sharpe ratios are lower than the MacKinlay (1995) threshold.

Table 8 presents hedge portfolio results similar to those in Table 7, but using characteristics rather than factor betas as the sorting criteria. Panel A demonstrates that RetLag1 (monthly reversals), Lag2_12 (momentum), Book/Mkt, ProfRato, and AsstGrth provide statistically significant average returns. The negative sign for RetLag1 and AssetGrth is consistent with their negative risk premium in Table 8. Panel B tests whether the Sharpe ratio from combining the market and the zero net investment portfolio from panel A is greater than the MacKinlay bound of 0.6. The results show that RetLag1, Lag2_12, Book/Mkt and ProfRato provide SRs that are statistically higher than the bound. That is, the strategies associated with these characteristics, although linked with premiums, provide abnormally high Sharpe ratios. When we construct hedge portfolios using the top and bottom 30% of stocks in panel C, the results are similar to those in panel A. The Sharpe ratios in panel D also have largely similar patterns as those in panel B, except that only Lag2_12 and Book/Mkt yield SR greater than 0.6.[27]

Table 9 presents the cross-correlations obtained using the hedge portfolios formed in panel A of Tables 7 and 8, across the factors and the characteristics. 86% of the correlations are below 0.5 in absolute magnitude. The hedge portfolio for the profitability characteristic is negatively correlated with SMB, whereas HML, RMW, and MOM, not surprisingly, are positively correlated with their characteristic-based counterparts. The hedge portfolios corresponding to some of the CRR factors are also positively cross-correlated with some characteristic-based portfolios, but there is no ready explanation for these results, so we leave a full explanation for future research.

Overall, when subjected to our protocol, across both the regression and hedge portfolio method, a market factor, a profitability factor, and factors based on credit spreads, term spread, and unexpected inflation are related to the covariance matrix, command statistically significant risk premiums in all specifications, and yield reasonable Sharpe ratios. Almost all characteristics are associated with statistically significant premiums, but only momentum and the book-to-market anomaly yield Sharpe ratios that exceed a reasonable bound to be considered an abnormal profit opportunity.

[27] Note that the sample observations used are different for different runs. The non-EIV sample sizes are different from the double-sorted EIV sample sizes and the double-sorted approach cannot be the same because it uses portfolio betas to replace the individual stock betas. This means that some observations are available after the double size sorts even when the individual betas are not available, for example, because we require 24 observations to compute them. The hedge portfolios are formed by sorting on betas or characteristics one variable at a time. Obviously, the beta sorted hedge portfolios and characteristics sorted hedge portfolios have different stocks, and, equally obvious, the particular stocks within the 10% and 30% portfolios are different for each beta and characteristic.

Table 8
Returns of hedge portfolios associated with characteristics

A. Hedge portfolio from 10% top and bottom

	RetLag1	Lag2_12	Lag13_36	SizeLag1	Book/Mkt	ProfRato	AsstGrth
Mean	−1.791	1.814	0.009	0.024	0.797	0.904	−0.510
SD	4.157	5.247	4.445	5.358	3.824	3.932	2.603
t(Mean)	***−10.23***	***8.212***	0.046	0.105	***4.951***	***5.458***	***−4.651***

B. Market return plus 10% top and bottom hedge portfolio returns

		Augmented returns						
	Rm-Rf	RetLag1	Lag2_12	Lag13_36	SizeLag1	Book/Mkt	ProfRato	AsstGrth
Mean	0.490	−1.300	2.304	0.499	0.514	1.287	1.394	−0.020
SD	4.584	5.255	6.775	6.819	7.333	4.859	5.788	5.951
t(Mean)	2.539	−5.877	8.077	1.737	1.664	6.292	5.719	−0.079
Sharpe	0.370	0.857	1.178	0.253	0.243	0.918	0.834	0.011
Sharpe t	***−5.275***	***5.223***	***10.55***	***−8.104***	***−8.363***	***6.331***	***4.791***	***−13.98***

C. Hedge portfolio 30% top and bottom

	RetLag1	Lag2_12	Lag13_36	SizeLag1	Book/Mkt	ProfRato	AsstGrth
Mean	−1.011	1.040	−0.035	0.097	0.562	0.420	−0.347
SD	2.870	3.643	3.032	3.958	2.643	2.533	1.738
t(Mean)	***−8.368***	***6.778***	−0.278	0.583	***5.045***	***3.940***	***−4.747***

D. Market return plus 30% top-and-bottom hedge portfolio return

		Augmented returns						
	Rm-Rf	RetLag1	Lag2_12	Lag13_36	SizeLag1	Book/Mkt	ProfRato	AsstGrth
Mean	0.490	−0.521	1.530	0.455	0.587	1.052	0.910	0.143
SD	4.584	4.640	5.562	5.591	6.302	4.330	5.094	5.505
t(Mean)	2.539	−2.666	6.533	1.931	2.213	5.768	4.244	0.616
Sharpe	0.370	0.389	0.953	0.282	0.323	0.841	0.619	0.090
Sharpe t	***−5.275***	***−4.833***	***6.950***	***−7.414***	***−6.416***	***4.926***	0.414	***−12.09***

Here are summary statistics for returns (expressed as a percentage per month) on hedge portfolios associated with the seven characteristics RetLag1, Lag2_12, Lag13_36, Size, Book/Mkt, ProfRato, and AsstGrth. See variable definitions in Table A1. For each characteristic, hedge portfolios are formed by a long position in a group of the stocks with high values of the characteristic and a short position a group with low characteristic values; this is done with replacement. Panel A (C) shows the returns of the hedge portfolios with the top and bottom deciles (the top 30% and bottom 30%). Panel B (D) show the returns of market (Rm-Rf) and augmented returns, which is the hedged portfolio returns added to the market return (Rm-Rf). The seven to the right of it are for the characteristics. The Sharpe ratio (Sharpe), t-statistic of the Sharpe ratio against 0.6 (Sharpe t), and the MacKinlay (1995) threshold are also reported. Critical rejections levels for the t-statistic are 1.65 (10%), 1.96 (5%), and 2.59 (1%). t-statistics breaching the 5% (1%) critical level are in boldface (boldface italic.)

9. Summary and Conclusions

Our goal in this paper is to suggest a protocol for categorizing factors that potentially are the drivers of asset returns and for determining whether they are associated with risk premiums. We are striving for a procedure that will be acceptable to scholars and practitioners; a standard for future factor identification. The protocol we present here is just an outline and it will undoubtedly be modified by others to render it more acceptable. Ours is just a first attempt.

Our suggested protocol has two stages. The first stage provides a sequence of steps that represent necessary conditions for factor candidates to be valid. A candidate that does not satisfy these conditions is not a risk factor, but this

Table 9
Correlations between factor- and characteristic-based hedge portfolios

	Rm-Rf	SMB	HML	RMW	MOM	LT_Rev	ΔDP	ΔTS	UNEXPI
RetLag1	−0.302	−0.305	0.117	0.049	0.437	−0.140	0.154	0.195	−0.288
Lag2_12	−0.115	−0.304	−0.011	0.114	**0.745**	−0.047	0.148	0.309	−0.280
Lag13_36	0.026	−0.343	−0.143	0.143	0.263	−0.307	0.175	0.388	−0.309
SizeLag1	0.145	**−0.642**	−0.092	0.175	0.366	−0.045	0.231	**0.639**	**−0.522**
Book/Mkt	−0.363	0.108	**0.643**	0.199	−0.006	0.078	−0.088	−0.142	−0.036
ProfRato	−0.032	**−0.508**	0.230	**0.611**	0.319	−0.034	0.248	**0.585**	**−0.593**
AsstGrth	0.390	0.023	−0.353	0.088	−0.059	−0.140	0.118	0.149	−0.067

This table provides the correlations between the hedge portfolios corresponding to the betas on the nine factors that pass necessary conditions and the hedge portfolios corresponding to the seven characteristics. These are hedge portfolios that are long the largest 30% of the values and short the smallest 30% for each factor beta and characteristic. The construction of the hedge portfolios is explained in Tables 7 and 8. See Table A1 for variable definitions. Correlations greater than or equal to 0.5 in absolute value are in boldface.

does not imply that rejected candidate is uninteresting, particularly to investors. Indeed, if such a rejected candidate is related to average returns on any set of assets, there is a potential profit opportunity. In principle, a diversified portfolio could be constructed to produce significant return with minimal risk. The second suggested stage entails testing whether factor candidates that satisfy the necessary conditions are pervasive and consequently have associated risk premiums or instead are unpriced in the cross-section.

One very important application of our protocol would be to study the relative importance of industry, country, and global factors. Intuitively, some factors might be pervasive globally but there is some doubt because many or perhaps most countries do not share fully integrated macroeconomic systems. This leaves room for country factors and, indeed, most previous studies of factors have been exclusively domestic. Finally, at an even lower level of aggregation, industry factors clearly have the ability to explain some individual firm covariances; but do they all carry no risk premiums or, instead, are at least some of them sufficiently pervasive to be genuine risk factors at either the country or global level?

Industry factors have been studied for a long time, from King (1966) through Moskowitz (2003). It seems to us that a very useful exercise would be to study industry factors globally. Following our suggested protocol, we would only need to assemble some international real asset returns, extract a time series of eigenvectors from their time-varying covariances, and check whether industry factors satisfy the necessary and sufficient conditions of Sections 4 and 5.

Table A.1
Variable definitions

Variables	Description
Rm-Rf	Market excess return factor (*Source*: Kenneth French's library)
SMB	Small minus big factor (*Source*: Kenneth French's library)
HML	High minus low book-to-market factor (*Source*: Kenneth French's library)
RMW	Robust minus weak operating profitability (*Source*: Kenneth French's library)
CMA	Conservative minus aggressive investment (*Source*: Kenneth French's library)
MOM	Momentum factor (*Source*: Kenneth French's library)
Rf	Three-month Treasury-bill rate (*Source*: Federal Reserve Bank of St. Louis)
LIQ	Traded factor of liquidity constructed from Pastor and Stambaugh's (2003) innovation series (*Source*: Lubos Pastor's Web site)
ST_Rev	Short-term reversal factor (*Source*: Kenneth French's library)
LT_Rev	Long-term reversal factor (*Source*: Kenneth French's library)
ΔDP	Traded factor of default risk premium where default risk premium is the yield spread between Moody's Baa and Aaa corporate bonds (*Source*: Federal Reserve Bank of St. Louis)
ΔIP	Traded factor of growth rate of industrial production, where industrial production is from the Federal Reserve Bank of St. Louis
ΔTS	Traded factor of term premium where term premium is the yield spread between the 10-year and the 1-year Treasury bonds (*Source*: Federal Reserve Bank of St. Louis)
UNEXPI	Traded factor of unexpected inflation where unexpected inflation at time t is the difference between inflation at time t and expected inflation at time t- 1. Both are available from the Federal Reserve Bank of St. Louis
ΔEI	Traded factor of change in expected inflation, where expected inflation is from the Federal Reserve Bank of St. Louis
RetLag1	Return in prior month (%) (*Source*: CRSP)
Lag2_12	Return in prior 2nd to 12th month (expressed as a percentage per month) (*Source*: CRSP)
Lag13_36	Return in prior 13th to 36th month (expressed as a percentage per month) (*Source*: CRSP)
SizeLag1	Natural log of size (market cap) lagged 1-month relative to the return. For instance, if the return is for February 1965, SizeLag1 is the log market cap at the end of January 1965. Price and number of shares outstanding are from CRSP
Book/Mkt	Book-to-market equity, the natural log of the ratio of the book value of equity to the market value of equity. Book equity is total assets (Compustat data item 6) for year t- 1, minus liabilities (181), plus balance sheet deferred taxes and investment tax credit (35) if available, minus preferred stock liquidating value (10) if available, or redemption value (56) if available, or carrying value (130). Market equity is price times shares outstanding at the end of December of t- 1 (*Source*: CRSP)
ProfRato	Profit ratio for June of year t is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity for the last fiscal year ending in t- 1
AsstGrth	Annual firm asset growth rate is calculated using the year-on-year percentage change in total assets (Compustat data item 6). The firm asset growth rate for year t is estimated as the percentage change in data item 6 from fiscal year ending in calendar year t- 2 to fiscal year ending in calendar year t- 1
Amihud's illiquidity ratio	The annual illiquidity ratio of stock i in year t measured as the average ratio of the daily absolute return to the (dollar) trading volume on that day divided by the number of days for which data are available for stock i in year t

References

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5:31–56.

Anderson, T. W. 1984. *An introduction to multivariate statistical analysis*, 2nd ed. New York: Wiley.

Arnott, R., J. Hsu, and P. Moore. 2005. Fundamental indexation. *Financial Analysts Journal* 61:83–99.

Bai, J. 2003. Inferential theory for factor models of large dimensions. *Econometrica* 71:135–71.

Barillas, F. and J. Shanken. 2018. Comparing asset pricing models. *Journal of Finance* 73:715–54.

Belloni, A., V. Chernozhukov, and C. Hansen. 2014. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies* 81:608–50.

Berk, J. B. 2000. Sorting out sorts. *Journal of Finance* 55:407–27.

Brennan, M. C., T. Chordia, and A. Subrahmanyam. 1998. Alternative factor specifications, security characteristics and the cross section of expected stock returns. *Journal of Financial Economics* 49:345–73.

Brown, S. 1989. The number of factors in security returns. *Journal of Finance* 44:1247–62.

Campbell, J. Y. and J. H. Cochrane. 2000. Explaining the poor performance of consumption-based asset pricing models. *Journal of Finance* 55:2863–78.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley. Forthcoming. An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics*.

Campbell, J. Y., M. Lettau, B. Malkiel, and Y. Xu. 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56:1–43.

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

Charoenrook, A., and J. Conrad. 2008. Identifying risk-based factors. Working paper, Vanderbilt University.

Chen, N., R. Roll, and S. A. Ross. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403.

Cochrane, J. H. 2001. *Asset pricing*. Princeton: Princeton University Press.

Connor, G. and R. A. Korajczyk. 1988. Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics* 21:255–89.

——— 1993. A test for the number of factors in an approximate factor model. *Journal of Finance* 48:1263–91

Cooper, I. and R. Priestley. 2011. Real investment and risk dynamics. *Journal of Financial Economics* 101:192–205.

Cotter, J., and R. Roll. 2015. A comparative anatomy of REITs and residential real estate indexes: Returns, risks and distributional characteristics. *Real Estate Economics* 43:209–40.

Daniel, K. and S. Titman. 1997. Evidence of the characteristics of cross sectional variation in Stock returns. *Journal of Finance* 52:1–33.

Engelberg, J., R. D. McLean, and J. Pontiff. Forthcoming. Anomalies and news. *Journal of Finance*.

Fama, E. and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47:427–66.

———. 1993 Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.

———. 2008. Dissecting anomalies. *Journal of Finance* 63:1653–78.

———. 2015. A Five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

———. 2018. Choosing factors. *Journal of Financial Economics* 128:234–52.

Fama, E. and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

Feng, G., S. Giglio, and D. Xiu. 2017. Taming the factor zoo. Working Paper, University of Chicago.

Green, J., J. Hand, and X. F. Zhang. 2013. The supraview of return predictive signals. *Review of Accounting Studies* 18:692–730.

———. 2017. The characteristics that provide independent information about average U.S. monthly stock returns. Review of Financial Studies 30:4389–436.

Harvey, C. R., and Y. Liu. 2016. Lucky factors. Working Paper, Duke University.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. …and the cross-section of expected returns. *Review of Financial Studies* 29:5–68.

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28:650–705.

Jegadeesh, N., J. Noh, K. Pukthuanthong, R. Roll, and J. Wang. Forthcoming. Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*.

Johnson, R. A., and D. W. Wichern. 2007. *Applied multivariate statistical analysis*, 6th ed. New York: Prentice-Hall.

Karolyi, G. A. 2016. The cross-section of expected returns: Where we stand today. *Review of Financial Studies* 29:2–4.

King, B. F. 1966. Market and industry factors in stock price behavior. *Journal of Business* 39:139–90.

Kozak, S., S. Nagel, and S. Santosh. 2018. Interpreting factor models. *Journal of Finance* 13:1183–223.

Lehmann, B., and D. Modest. 1988. The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics* 21:213–54.

Lewellen, J. 2015. The cross-section of expected stock returns. *Critical Finance Review* 4:1–44.

Lewellen, J., and S. Nagel. 2006. The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics* 82:289–314.

Lewellen, J., S. Nagel, and J. Shanken. 2010. A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96:175–94.

Linnainmaa, J., and M. R. Roberts. 2017. The history of the cross section of stock returns. forthcoming *Review of Financial Studies*.

Lo, A., and A. C. MacKinlay. 1990. When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3:175–205.

MacKinlay, A. C. 1995. Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics* 38:3–28.

McLean, R. D., and J. Pontiff. 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71:5–31.

Moskowitz, T. J. 2003. An analysis of covariance risk and pricing anomalies. *Review of Financial Studies* 16:417–57.

Onatski, A. 2009. Testing hypotheses about the number of factors in large factor models. *Econometrica* 77:1447–70.

Pastor, L., and R. F. Stambaugh. 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111:642–85.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Roll, R. 1977. A Critique of the asset pricing theory's test. *Journal of Financial Economics* 4:129–76.

Roll, R., and S. A. Ross. 1980. An empirical investigation of the arbitrage pricing theory. *Journal of Finance* 35:1073–103.

Ross, S. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13:341–60.

Singleton, K. J. 2006. *Empirical dynamic asset pricing: Model specification and econometric Assessment*. Princeton: Princeton University Press.

Subrahmanyam, A. 2010. The cross-section of expected stock returns: What have we learnt from the past twenty-five years of research? *European Financial Management* 16:27–42.

Zhang, L. 2005. The value premium. *Journal of Finance* 60:67–103.