A Protocol for Factor Identification

by

Kuntara Pukthuanthong and Richard Roll

Abstract

Several hundred factor candidates have been suggested in the literature. We propose a protocol for determining which factor candidates are related to risks and which candidates are related to mean returns. Factor candidates could be related to both risk and returns, to neither, or to one but not the other.

A characteristic cannot be a factor. Time variation in both risk premiums and covariances is a challenge, but manageable with recently developed statistical procedures. We illustrate those techniques and also propose a new instrumental variables method to resolve the errors-in-variables problem in estimating factor exposures (betas) for individual assets.

	Authors' Coordinates	
	Pukthuanthong	Roll
Address	Department of Finance University of Missouri Columbia, MO 65211	California Institute of Technology Pasadena, CA 91125
Voice	(619) 807-6124	(626) 395-3890
E-Mail	pukthuanthongk@missouri.edu	rroll@Caltech.edu

Acknowledgements

We thank 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 seminar participants at the 2012 Inquire UK conference, the 2013 Southwestern Finance Association annual meeting, the UCLA Brown Bag Seminar, the 2013 Rotman ICPM Seminar, the 2013 Australian Finance Conference and the 2014 Asian Finance Association Conference for very helpful and constructive suggestions. We owe a great debt of gratitude to Olivier Ledoit for generous and unstinting guidance. We are thankful for the financial support from Inquire UK, the Dauphine–Amundi Chair in Asset Management, and the Rotman International Centre for Pension Management (ICPM) at the University of Toronto

I. Multiple systematic common factors

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.

The situation is well illustrated by the fact that numerous candidates for factors have been proposed in a voluminous literature. Lewellen, Nagel and Shanken [2010] list several of the most prominent in their opening paragraph and note that although they explain some empirical regularities, they have "…little in common economically with each other" (p. 175.) Subrahmanyam [2010] surveys more than fifty characteristics that various papers contend are cross-sectionally related to mean returns. McLean and Pontiff (2013) examine 95 characteristics that were claimed in previous papers to explain returns cross-sectionally but find that predictability declines after publication. Lewellen (2014) finds strong predictive power of actual returns using 15 firm characteristics. Harvey, Liu, and Zhu [2014] uncover 316 factor candidates 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. Green, Hand, and Zhang [2013] identify 330 firm characteristics and Green, Hand, and Zhang [2014] 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.

There are few topics in finance that are more important; factors are the main principal determinants of investment performance and risk. Indeed, the comparative values of well-diversified portfolios are determined almost completely by their factor exposures. Whether investors know it or not, every diversified portfolio is absolutely in thrall to 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; e.g., across bond vs. equities vs. 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.[1]

The first empirical fact indicates the existence of at least one common underlying systematic influences, (or "risk drivers" or "factors") 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; otherwise diversified portfolios would be more correlated across asset classes, countries, and sectors.

This is the first study proposing a set of necessary and sufficient conditions to categorize factor candidates. Our protocol identifies not only factors associated with risk premiums, 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, since they offer potential profit opportunities.

Our main goal is to popularize a process that will be broadly acceptable to both scholars and practitioners. We also propose a novel way of applying instrumental variables to resolve the error-

[1] Cotter and Roll [2012] report that real estate investment trusts have rather low betas against the S&P 500.

in-variables problems in estimating risk premiums in Fama-MacBeth regression with individual assets.

Our protocol is motivated by a basic principle: factor movements should not be easily foreseeable; thus, a characteristic such as firm size, or anything else known in advance cannot be a factor. However, characteristics can be related to mean returns either because they align with factor loadings or because they represent arbitrage opportunities.

Our protocol is composed of six steps to check necessary conditions and one final step that examines a sufficient condition. A fundamental necessary condition for any factor candidate is that it must be related to the principal components of a (conditional) covariance matrix of returns. However, this necessary condition does not distinguish between pervasive priced factors, those with risk premiums, and non-pervasive, fully diversifiable factors, which are related to covariances of some individual assets but do not influence the aggregate portfolio. Our sufficient condition test provides this distinction.

Factor candidates that do not satisfy the fundamental necessary condition are not without interest, particularly to practical investors. If a factor candidate is reliably related to mean returns but is unrelated to the (conditional) covariance matrix, it represents a potential arbitrage opportunity, which is, of course, intensely alluring. We explain how to take advantage of such factor candidates, should any be found.

Finally, we note at the outset 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., (2014), both of which examine the validity of the "conditional" CAPM. For 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, equivalently, a conditional mean/variance efficient portfolio, is always available to conditionally explain the cross-section of asset returns with a single exposure coefficient.[2] But the time variation in the SDF (or in risk premiums) might be so complex that multiple factors become a practical way to explain unconditional returns over any finite sample, however short. Singleton (2006, chapter 11), presents an in-depth analysis of this situation. He notes that in some circumstances, when investors have incomplete information,

> …it may appear necessary to include additional risk factors…to explain the cross section of historical means, even though these factor are not priced by agents," (p. 297)

But from a practical perspective, either of an investor or a financial econometrician, finite samples are inevitable. Our aim is to popularize an identification procedure for risk and non-risk factors that is useful, though perhaps not theoretically pristine in the sense of being congruent with a SDF.

In addition to the protocol, we also propose a novel solution for the errors-in-variables problem in the Fama/MacBeth (1973) method for estimating risk premiums. This solution exploits instrumental variables that allows the use of individual assets in that method.

Our findings are that six of the previously-proposed factors we examine pass the necessary conditions but none except momentum is associated with risk premiums.

II. Factors versus Characteristics

The two most fundamental attributes of a factor are (1) it varies unpredictably in a time series sense and (2) its variations induce changes in asset prices. Many seemingly reasonable candidates for factors, such as macro-economic variables, are at least partly predictable over time, but relatively efficient asset prices should not be influenced by movements that can be easily predicted. This suggests that the unpredictable variation in a factor should be the main driver of changes in asset prices.[3]

---

[2] As emphasized by Cochrane (2001) and Singleton (2006, chapter 8)

[3] Essentially, we are abstracting from influences on low frequency changes in expected returns, such as time variation in risk preferences and we are also ruling out factor modeling based on irrationalities such as in Daniel, Hirshleifer and Subrahmanyam (2005).

A further implication is that any observable asset characteristic cannot itself be a factor because it is known in advance. There is some misunderstanding in the finance community about this point because several seemingly successful factor models have been built on the basis of firm characteristics; e.g., the size of a firm, the market/book ratio, the past momentum in prices, beta, modified duration, beta estimation risk,[4] or even (with tongue in cheek) the alphabetical ranking of a firm's name.[5]

Of course, firms sorted by a pre-known characteristic might happen to be differentially influenced by some unknown factor or factors. If such an empirical regularity is observed, then a portfolio can be constructed whose unpredictable variation proxies for the underlying unknown factor(s); but the quality of the proxy and the identity of the underlying factor(s) remain a mystery until additional information is uncovered that reveals the true nature of the connection.

Leading authorities agree on this point. For instance, Cochrane (2001) says

> The fact that betas (i.e., the true factor loadings or "exposures") are regression coefficients is crucially important...the betas cannot be asset-specific or firm-specific characteristics, such as the size of the firm, book to market ratio, or (to take an extreme example) the first letter of its ticker symbol, (p. 81.)

Cochrane points out the characteristics can be associated with expected returns (and hence with compensation for risk) but this association "...must be *explained* by some beta regression coefficient," (p. 81, emphasis in original.)

A good example of this phenomenon involves the book/market ratio. Firms sorted by book/market appear to have systematically diverse average returns, which prompted Fama and French (1992) to construct a portfolio that is long higher book/market stocks and short lower book/market stocks. The variation in this portfolio's value could be a proxy, albeit of unknown quality, for the movements in some underlying factor or factors; there has been much subsequent debate about its or their identities.

---

[4] See Graham (2012).
[5] See Ferson, et al., (1999).

However, Daniel and Titman (1997) note that book/market is a characteristic, not a factor. They show empirically that stocks with higher book/market ratios, but not higher loadings on the Fama/French portfolio, have higher returns on average.[6] They conclude that book/market was not related to riskiness. Perhaps so, but there is another possibility; the proxy represented by the Fama/French portfolio might not be all that closely related to the underlying but unknown factor that drives returns on higher book/market stocks. The Fama/French portfolio could simply be a poor proxy or not even related to the factor. Or the portfolio could be proxying for a rather weak factor and not be related to another stronger factor.

For practitioners, it is understandably tempting to find characteristics related to average returns because this could bring superior investment performance, provided that such characteristics are not proxies for factor risk loadings. Indeed, this is the foundation on which some firms provide advice to the investment community. For instance, MSCI uses the Barra Aegis approach to identify characteristics and use them to "optimize" portfolios.

Goyal (2012) remarks that "…it is especially easy to check for pricing of characteristics…" (p. 14), using the cross-sectional time-series method developed by Fama and MacBeth (1973). Goyal discusses various refinements including an important one devised by Brennan, Chordia and Subrahmanyam (1998) that efficiently mixes known factor betas, (presuming that some are known), with firm characteristics. Building on these ideas, Lewellen (2014) finds that firms sorted by known characteristics have predictable and reliably diverse future mean returns. The problem, of course, is that a "priced" characteristic is not necessarily associated with a genuine risk factor for the reason given by Cochrane above. It is not necessarily associated with a beta from a factor regression.

Potentially priced characteristics must seem alluring also to finance academics. Subrahmanyam (2010) after citing the finance literature more than fifty cross-sectional characteristics used to predict returns. But he warns,

[6] Daniel and Titman make a similar observation about firm size.

> The overall picture…remains murky, because more needs to be done to consider the correlational structure amongst the variables, use a comprehensive set of controls, and discern whether the results survive simple variations in methodology, (p. 27.)

It seems clear enough that a characteristic related to average return cannot simply be assumed to be related to a factor, without confirming information. The characteristic *might* proxy for a loading on an unknown risk factor. On the other hand, it *might* represent an investment opportunity and not be related to a factor at all.

The Holy Grail is to uncover the best confirming, or disconfirming, information. We will propose a procedure below but would at this point simply like to acknowledge that this quest is nothing new. A number of scholars have already pursued it; nonetheless, this is the first study proposing a procedure that provides necessary and sufficient conditions for a candidate factor to be a priced risk-factor.

One laudatory example is provided 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 correlated positively 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 though, that such an empirical finding is not a sufficient condition.

CC find empirically that several proposed factor candidates, including size, book/market, and a liquidity construct, satisfy the above necessary condition. Momentum[7] 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.

[7] A factor candidate originally proposed by Carhart (1997).

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, book/market and liquidity are in a plausible range. They find that the Sharpe ratios of size and book/market are "plausible" but the Sharpe ratio for liquidity is not. We are left in doubt. Although size, book/market and liquidity satisfy a necessary condition to be risk factors, CC's plausibility comparisons do not represent rigorous tests of sufficiency. Consequently, we must await further developments.

Since time variation in risk premiums is required for the CC necessity condition, it cannot provide evidence for factor candidates with stable risk premiums or with statistically small variation. Perhaps there are no such factors, and we just don't know at this juncture.

III. Factors and the (possibly conditional) covariance matrix

If the world were very simple, there would be one sure-fire method to extract good factor proxies. Linear combinations of factors would reveal themselves in the principal components analysis (PCA) of the covariance matrix of observed returns. In a really perfect world, there would be only a few principal components with large eigenvalues (many fewer than assets.) If this were true, factor proxies could be extracted from rather small subsets of assets.

To a first-order approximation, the covariance matrix is not affected by the cross-section of expected returns, so the temptation to identify return-related characteristics is absent. Instead, the main component is the time variation in asset prices, the very essence of factor influence.

A necessary condition for any candidate 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, book/market, 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. Book/market 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 book/market.

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 non-stationarity since there is no plausible reason why the number of factors or their relative importance should be constant through time[8];

c. It includes true risk drivers, pervasive non-diversifiable factors (or linear combinations thereof) along with diversifiable factors, perhaps such as industry factors, that are not associated with risk premiums.

Fortunately, there is a remedy, perhaps imperfect, for each of these conundrums. For (a), at least conceivably, the linear combinations extracted by PCA could be related to other candidate factors, such as macro-economic 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), time dependence could be introduced into the PCA method (see below.) For (c), a second stage method as in Fama and MacBeth (1973) or Roll and Ross (1980) could be employed to distinguish priced (presumably non-diversifiable) factors from others. Needless to say, none of these cures is without its own problems.[9]

IV. But what are the Underlying Factors?

What exactly are the salient features of factors, the underlying risk drivers? Cochrane (2001) says unequivocally, "The central and unfinished task of absolute asset pricing[10] is to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices." (p. xiv.) He

[8] Moreover, it seems that non-stationarity is an empirical fact. Moskowitz (2003) 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," (p. 436).

[9] 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, but merely 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.

[10] As opposed to relative asset pricing such as comparing an option price to the underlying stock price.

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.

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

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.

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 pre-specifies macro-economic variables that seem likely to be pervasive and then pre-whitens 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/French [1992], Carhart [1997].) Finally, a lesser known but simpler approach is to employ a handful of rather heterogeneous indexes or tradable portfolios.

Each of the above approaches has particular limitations. Purely statistical methods are theoretically sound but everything has to be stationary. Pre-specified macro-economic variables are the most theoretically solid but are observed with excruciatingly low frequency. Factor proxies suggested by asset pricing are weak theoretically and are not necessarily even related to risk. A group of heterogeneous diversified portfolios can have non-stationary compositions and be observed at high frequency - but heterogeneity must be sufficient to span all relevant and pervasive underlying risk drivers.

Heterogeneous portfolios work well for spanning global factors. Pukthuanthong and Roll (2009) went to a lot of trouble to extract ten global principal components. They employed the extracted global principal components as factor proxies and demonstrated a substantial increase in global market integration for many countries. Then, as a robustness check for their purely statistical procedures, they replaced the principal components with broad indexes from ten large countries and found virtually identical results. Country indexes are evidently sufficiently heterogeneous to span the same underlying macro perceptions as principal components.

Using a set group of portfolios is arguably the easiest and best approach to factor estimation if heterogeneity can be assured, which suggests that a well-chosen set of exchange traded funds (ETFs) might serve the purpose quite well. ETFs are often diversified portfolios or derivatives-based equivalents. As such, their returns must be driven mainly by underlying factors; i.e., by high-frequency changes in market perceptions of macro-economic conditions. Their idiosyncratic volatility should be relatively small. Moreover, they are generally liquid, transparent, and cheap to trade. Their variety across several asset classes suggests a healthy degree of heterogeneity but it remains to be seen if the available ETFs are heterogeneous enough to span the underlying factors. We examine this issue empirically below.

V. Putting it All Together: The Necessary Conditions.

Given the discussion above, we are ready to outline the first stage of our protocol for identifying factors. This stage involves a sequence of necessary conditions. It identifies factors that move asset prices but it does not distinguish between pervasive priced factors and diversifiable factors. That crucial information is postponed to a later stage. Here are the recommended steps for the necessary conditions:

First, collect a set of N equities for the factor candidates to explain. It needs not be all that large because the number of factors is presumably rather small. The test assets should belong to different industries and have enough heterogeneity so that the underlying risk premium associated factors can be detected.

Second, using simultaneous returns for the N real assets over some calendar interval, such as monthly observations for ten years, extract the real return eigenvectors corresponding to the L largest eigenvalues. 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 ninety percent.

Third, collect a set of K factor candidates that consist of two subsets, one subset is comprised of J ETFs, where J is a half-dozen or so, that are as heterogeneous as possible. This is a base group that possibly spans many of the underlying macroeconomic factors.[11] The second subset should consist of any K-J factor candidates of interest. These could be well known characteristics-based candidates, (size, book/market, momentum) or any of the 50 or so documented in Subrahmanyam (2010), or any new candidate as yet to be suggested.

Fourth, using the L eigenvectors from step #2 and the K factor candidates from step #3, calculate the time varying **conditional** real return covariance matrix, V_t (L+K x L+K). Our preferred method for calculating V_t is the innovative technique of Ledoit, Santa-Clara, and Wolf (2003), hereafter LSW. LSW estimates GARCH(1,1) coefficients for diagonal terms one at a time and off-diagonal coefficients pair-wise. It then minimizes a Frobenius norm to deliver modified coefficients that guarantee positive semi-definite conditional covariance matrices in each sample period. The conditional covariance matrices are time varying but typically somewhat persistent.

Fifth, from the conditional covariance matrix V_t, in each period t, break out a sub-matrix, the conditional cross-covariance matrix, which we denote C_t. It has K rows and L columns (i.e., K x

[11] An alternative would be to include a number of broad market indexes but it seems unlikely that these could be as heterogeneous as ETFs.

L); the entry in the i^{th} row and j^{th} column being the covariance between factor candidate i and eigenvector j. It will also be necessary to break out the conditional covariance sub-matrix of the factor candidates, $V_{f,t}$ (K x K) and the conditional covariance sub-matrix of the real eigenvectors, $V_{e,t}$ (L x L).

Sixth, Using the sub-matrices from step #5, compute canonical correlations between the factor candidates and the eigenvectors for each date t. This involves first finding two weighting column vectors, a_t and b_t, where a_t has K rows and b_t has L rows so the covariance between the weighted averages of factor candidates and eigenvectors is $a_t'C_tb_t$. Their correlation is

$$\rho = \frac{a_t'C_tb_t}{\sqrt{a_t'V_{f,t}a_tb_t'V_{e,t}b_t}}.$$

The correlation is maximized over all choices of a_t and b_t. It turns out that the maximum occurs when $a_t = V_{f,t}^{-1/2}g_t$ and g_t is the eigenvector corresponding to the maximum eigenvalue in the matrix $V_{f,t}^{-1/2}C_tV_{e,t}^{-1}C_t'V_{f,t}^{-1/2}$. The vector b_t is proportional to g_t. 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-Square 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. Also, by examining the relative sizes of the weightings in a_t, one can obtain an insight into which factor candidates, if any, are more related to real return covariances. Since a_t is estimated separately for each date t, it forms a time-series that can be used to test the long-term association of factor candidates and real returns.

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 V above.) But 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

[12] See Anderson (1984, ch. 12) or Johnson and Wichern (2007).

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. 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.

A rejected factor candidate, one that is not associated with the covariance matrix of real asset returns, is perhaps more interesting than the accepted candidates, at least to investors. If a rejected factor candidate is reliably associated with the mean returns on a diversified portfolio of assets, whether or not they are real assets, then an arbitrage is possible. To construct the arbitrage, factor loadings are required on the accepted factors and on the rejected factor; then a long/short position is engineered from the rejected factors loadings while a zero factor exposure is constructed against the accepted factors. This essentially eliminates priced risky factor exposure and produces a positive cash flow with little risk if the resulting idiosyncratic volatility is minimal.

The protocol just outlined bears a resemblance to the approach followed by Moskowitz (2003). He used a preliminary version of the LSW covariance estimation. He employed industry portfolios instead of heterogeneous ETFs and his main factor candidate of interest were the three Fama/French factors plus the Carhart momentum factor. His criteria for judging a factor, however, was based on a decomposition of the conditional covariance matrix, both the conditional in-sample (i.e., same period) matrix and the conditional look-ahead out-of-sample matrix. He found that principal components extracted in sample dominated any of the tested factor combinations but did poorly out of sample.

If one assumes, as Moskowitz does, that the estimated conditional LSW covariance matrix is the true matrix, then it's not immediately clear why an out-of-sample comparison is relevant. After

[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.)

all, the true covariance matrix of real returns could be dramatically different from one period to another. An arguably more pertinent question is whether factor candidates capture such variation. This is the same as asking whether the factor candidates span the same space as the eigenvectors; i.e., by the correlation between eigenvectors and factor candidates within each period.

VI. Putting is All Together: Sufficient Conditions.

Although factor candidates that do not satisfy the necessary conditions in section IV might be more interesting to investors, risk control and sound portfolio management requires that we test those factors that do pass muster in order to determine if they are associated with economically meaningful risk premiums. Factor candidates that are associated with the covariance matrix of real returns but do not entail risk premiums must, according to theory, be fully diversifiable.

In principle, this sufficiency stage is easy. We simply run a pooled cross-section/time series panel with real returns as dependent variables and factors that satisfy the necessary conditions 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.

A variant of the panel approach is venerable in finance; it was originated by Fama and MacBeth (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 problem because the betas are the explanatory variables in each FM cross-section. Moreover, the errors are certainly more egregious in beta estimates for individual stocks than they are in betas estimate for portfolios, which explains why FM used the latter. Fortunately, there is now a promising econometric cure: instrumental variables (IV) whose errors are not correlated with the errors in the beta estimates.

What would be good instruments for betas? It seems sensible to use betas estimated from other observations in the time series. For example, we could first estimate betas from a time series prior to the cross-sectional period, as in FM, and then estimate betas again but from the time series observations after the cross-sectional period; the latter would serve as instruments for the former (or vice versa.)

Alternatively, we could estimate betas and beta instruments from alternating observations and then run the IV cross-section on observations interspersed with, but not used in, estimating the betas and instruments. This approach takes time-varying betas into account.[14] To clarify, with N return observations over some calendar period, we could split the sample into three interspersed subsamples, estimating betas from observations 1, 4, 7, ….N-2, beta instruments from 2, 5, 8, …N-1, and then calculating the cross sectional IV regression using observations 3, 6, 9,…N. By using alternating observations over the same calendar period, the betas and their instruments should be well connected but their estimation errors should be virtually unrelated because both factor returns and real asset returns are not very serially dependent.

VII. Industry Factors, Domestic Factors, and Global Factors.

One very important application of the protocol suggested in the two preceding sections 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 are they diversifiable and 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 V and VI above.

---

[14] We are not assuming that PCs are constant and hence adopt the Ledoit et al approach of estimating time varying covariance matrices. The IV version of FM does not assume stationarity either but the developments after the original FM procedure did with rolling windows. We can also allow rolling windows and within each window use IV. In fact, our approach of using the alternative observations for estimating betas and IVs for betas takes account of non-stationarity better within each window. The original FM used sequential windows for estimating betas and doing the cross-sectional regressions.

VIII. Numerical Example: stocks, global ETFs, Fama/French and Momentum.

This section presents a numerical example of the suggested protocol using simultaneous monthly return observations over a ten-year period, 2002-2011 inclusive. The sample assets consist of three groups. The first consists of individual U.S. equities. We select firms with minimal debt according to COMPUSTAT to minimize non-linearities in the data that might be induced by embedded options. There are 1,100 minimal debt firms with complete monthly return records over the sample decade but there is no need to use all of them when checking the necessary conditions outlined above, so we randomly select 30 and then extract 10 principal components from their unconditional return covariance matrix. These firms are identified by name in Table 1. Since all of firms are chosen to have minimal debt, we do not pretend that they are a comprehensive sample over either industry or size. It appears that many of these sample firms are rather small, Microsoft being the glaring exception. Bank South Carolina Corp would seem to be an anomaly, a purported bank with no debt. In the notation of section V, N=30 and L=10.

Second, we collect a group of ETFs. The data problem here was that ETFs are a relatively recent phenomenon and there are not very many that have been listed continually for the ten years culminating in December 2011. Also, as mentioned earlier, it is important that they be relatively heterogeneous, so we discard a few continually-listed ETFs because their unconditional correlation exceeded .95 with one of those retained. Table 2 lists the seven remaining ETFs. Five of seven are internationally oriented, which is an advantage because we would ideally like to check factor candidates that are global.

One might question why we do not use ETFs as PCs in the first step. In order to do that, ETFs need to be heterogeneous enough to span the underlying factors, and have long series. To investigate this possibility, we examine ETFs that have data available since 2005; (most of the ETFs do not start trading until 2005.[15]) Next, we randomly select one hundred stocks that have the data from 2005 to 2014 and extract ten principal components from their returns. We regress the returns of the ETFs on the ten PCs thinking that truly heterogeneous ETFs would be related to

[15] There are only 28 ETFs with available data from 1999 and 146 that have the data from 2005 to the end of 2014.

many different PCs. However, of the 146 post-2005 ETFs examined, 139 have the their largest t-statistic on the first PC. Only one ETF has its largest t-statistic on the second PC and six have their largest on the third PC. This implies that the available ETFs are not that heterogeneous. Consequently, the ETFs are unlikely to span the true underlying factors.

Finally, from Ken French's web site, we collect the returns on two of the Fama/French (1992) factors, SMB and HML. In addition, we copy the returns on the Carhart (1997) momentum factor. We did not include the Fama/French overall market factor because it is highly correlated with the first principal component and, moreover, most people agree that it does not need to be tested as a factor candidate.

With the J=7 ETFs, the two Fama/French factors, the Carhart momentum factors, and the L=10 principal components, the conditional covariance matrix V_t is 20 X 20 in each month t. The unconditional covariance matrix over the 120 month sample is printed in Table 3 in the form of standard deviation along the diagonal and correlations off-diagonal. The first ten assets are the principal components, which are orthogonal to each other by construction. We normalize them to have variance unity. For the other ten assets, the factor candidates, standard deviations are in their natural units of percent per month.

As explained in section V above, the LSW procedure for estimating V_t proceeds in three stages. The first stage is to fit a GARCH(1,1) process to the own variance of each asset in V; essentially to estimate the 20 diagonal terms for each month t = 2,...120. For t=1, we arbitrarily assign the starting value of the GARCH process to the unconditional variance for each asset. GARCH is estimated by maximum likelihood and we assume a Gaussian likelihood function. The maximization is over a non-linear function with some complicated gradients and subject to constraints, for which we employ the SNOPT™ software.[16] Estimating GARCH for PCs is somewhat tricky numerically.[17]

[16] Obtained from Stanford Business Software, Inc.
[17] We will be happy to provide some insights into this estimation problem upon request.

The LSW second stage fits a GARCH(1,1) to each pairwise covariance in the V matrix using the own variance estimates from the first stage. Again, this is a (Gaussian) maximum likelihood procedure, but with different gradients and constraints. Since V is 20 X 20 and symmetric, there are 190 separate covariance GARCH processes to estimate in this stage. We set the starting value of the GARCH covariance process at zero, but this induces some numerical difficulties in the first few iterations because the GARCH cross-product coefficient is pushed rapidly toward zero (incorrectly), so it has to be constrained at a small value slightly above zero (another trick for future implementers.) If this is done, the optimum is eventually reached after only a few more iterations.

The first two stages of the LSW procedure produce three 20 X 20 symmetric matrices, denoted A, B, and C in their paper, that contain in each corresponding i,j^{th} element the three GARCH parameters estimated for the 20 own variances and the 190 covariances. Since the covariance parameters are estimated one at time without regard to each other, the A, B, and C matrices need not be positive definite and usually are not. So a third stage is required, which entails finding three indefinite matrices that are as "close as possible" to A, B, and C. Closeness is measured by the Frobenius norm, the sum of squared differences between the elements of each of the A, B, and C matrices and the corresponding elements of their companion indefinite matrices. This again involves an optimization as outlined in the appendix of their paper.[18] The three indefinite matrices close to A, B, and C can then be used along with the squared returns and return cross-products, to produce a V_t for each month t that is positive definite and evolves over time as a GARCH multivariate process.

Consequently, for each month in the sample (except for the first month, when V_t has only the GARCH starting values and hence is not really an estimate), we obtain an estimated conditional covariance matrix. Since our original data sample was 2002-2011 inclusive, there are 119 of these conditional covariance matrices. As an example, the conditional covariance matrix for December 2006, roughly halfway through the sample, is reported in Table 4 (with standard deviations on the

[18] We are extremely indebted to Olivier Ledoit who not only showed us the Matlab code programmed by Ilya Sharapov for the procedure in the LSW appendix, but coached us through many missteps along the way in developing our own Fortran code, even checking our numerical results with his own code. Undoubtedly, we would still be working on this without Olivier's guidance.

diagonal and correlations off-diagonal.) This matrix can be compared to the unconditional matrix in Table 3. There appear to be rather minor differences. For example, nine of ten conditional PC standard deviations are slightly below 1.0, their unconditional values. There does not seem to be a systematic pattern in the conditional covariances.

The next and sixth step of our protocol is to calculate canonical correlations from the sub-matrices of V_t (see section V above.) Since we have ten PCs and ten factor candidates, there are ten 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, but the correlation itself can be either positive or negative. Table 5 reports, in the first column, the canonical correlations for the unconditional covariance matrix covering 2002-2011 monthly. The second column gives the time series means of the corresponding canonical correlations from the LSW conditional covariance matrices in each month from February 2002 through December 2011; (the first month is lost in the GARCH estimation.) For these means, the third column gives Newey-West T-statistics corrected for autocorrelation (with ten lags)[19] and heteroscedasticity.

As indicated by these results, the first and largest canonical correlation is dominant. Its mean conditional value is 0.84 with a T-statistic over 85. Although some of the smaller correlations seem moderately large unconditionally, only one exceeds .2 conditionally. There are some seemingly significant T-statistics for the other conditional means as well, but their order of magnitude is dwarfed by the first one, so we will focus on it from here on.

We are particularly interested in the vector a_t from section V, which weights the factor candidates, the J ETFs plus the two Fama/French factors and the Carhart momentum factor. For the reasons just mentioned, we report a_t only for the largest canonical correlation.[20] Since a_t is obtained from V_t for t=2,..,120, each element of a_t forms a time series whose mean can be checked for statistical significance. The same is true for the vector b_t, which weights the PCs. Results are in Table 6.

[19] We experiment with different lags and find that the t-statistics become flat around ten.
[20] The vectors for the second through the 10th canonical correlations are available from the authors upon request.

Turning first to b_t, the PC weightings, which are reported in the right panel of Table 6, we see that eight of the ten mean conditional weightings are statistically significant based on the Newey-West T-statistics. The mean weightings are also fairly close to the unconditional weightings except for those that are relatively small in absolute magnitude. The first principal component's weight is around 90% and it is clearly dominant. Since the PCs are mutually orthogonal unconditionally and have small intercorrelations conditionally, the variance of the first canonical variate depends on the squares of these weightings. In other words, the first PC induces more than eighty percent of the variance in the first canonical variate while the next most important PC contributes only around six percent. The eigenvector for the first PC has mostly positive elements, so it is a quasi-market index. The signs of the lower-ranking PCs weightings have little meaning because the PCs themselves have arbitrary signs.

Among the factor candidates, reported in the left panel of Table 6, five of the seven ETFs have significant mean conditional weightings. Only the first two are not significant. It is perhaps no surprise that the US ETFs (IYR and SPY) are significant since our sample is entirely from the US market, but there seem to be some global influences at work too because the Hong Kong, Japan, and Latin American ETFs are also significant.

The Fama/French SMB factor is also significant and positive with a T-statistic over seven. Thus, a portfolio constructed from firm size characteristics proxies for a factor that is truly related to the covariance matrix of returns. SMB passes the necessary conditions for a viable risk factor candidate. Remember too that first canonical correlation is very large (around +.9) and statistically significant (Table 5) and is dominated by the first PC (Table 6), so SMB is clearly related in a positive manner to underlying risk. Since SMB has usually been associated with a positive average returns, it is a strong candidate for a priced risk factor.

As regards HML and Momentum, both have negative weightings in the first canonical variate, unconditionally and conditionally, and they are statistically significant based on the T-statistic of their time series means. Since principal components have arbitrary signs, it seems that both of these factor candidates pass the necessary conditions of being related to the covariance matrix.

To get a bit more insight on this issue, we repeat this stage of the protocol without using any of the ETFs, but just SMB, HML, and Momentum by themselves as factor candidates. This allows us to check whether the ETFs are somehow interfering with a possible relation between SMB, HML, or Momentum and the underlying risk factors. The results reveal a drastic reduction of the largest canonical correlation, from conditional value of .840 in Table 5 with a T-statistic of 85, to a value of .449 with a T-statistic of 9.6. Clearly, the ETFs are very important in creating a large canonical correlation. The T-statistics of the SMB, HML, and Momentum weightings in the first canonical variate change to +10.2, +2.57 and -2.27, respectively. Thus, SMB becomes more significant and remains positive, HML changes sign and has reduced significance while Momentum retains its negative sign but has reduced significance.

The final step in the protocol is to check the sufficient conditions for a factor candidate that satisfies the necessary conditions. The results above suggest that the last five of the seven ETFs, the Fama/French SMB (size) and HML (book/market) factors, and Momentum pass muster and should be checked for sufficiency; i.e., for whether or not they are associated with statistically significant risk premiums.

For this illustrative example, we adopt a variant of Fama/MacBeth (1973) that differs in two significant respects: (1) it uses individual assets instead of portfolios[21] and (2) it employs instrumental variables to resolve the errors-in-variables problem inherent in estimating factor loadings (betas.)

Instrumental variables have their own particular difficulties, the most serious being that small sample properties are not amenable to analytic investigation. Statistical significance must rely on asymptotic arguments. Consequently, large samples are safer and more likely to produce reliable inferences, so we will use all 1,100 stocks at this stage. In addition, we will exploit the larger sample available with daily observations. There are 2517 daily return observations during 2002-2011 inclusive for the 1,100 stocks.

[21] Individual (real) assets are perhaps better than portfolios because averaging can wash out material economic features.

To validate this procedure, we perform a small-scale simulation that closely mimics the actual IV data analysis to come hereafter. Using a Gaussian random number generator, we first simulate 2517 returns for each of six factors.[22] By assumption and construction, the factor returns are independent and identically distributed, with standard deviations, respectively, of 0.4%, 0.3%, 0.3%, 0.2%, 0.2%, and 0.2% per day. Their assumed respective means in percent per day are, respectively, .04%, .03%, .02%, 0%, 0%, 0%. In other words, factors 1-3 have positive means (or risk premiums) while factors 4-6 have no risk premiums and hence represent diversifiable risk drivers.

For each of the six factors, the cross-sectional of betas (or risk exposures) is drawn from another independent Gaussian distribution. In this case, the mean beta on the first factor is set equal to 1.0, since this factor is supposed to mimic the market. The mean beta is zero for the other five factors, since these are supposed to mimic higher-order risk drivers (such as principal components) that cannot be moving all stocks in the same direction. However, the individual stock betas for higher-order are not zero since they are generated by a cross-sectional distribution with standard deviations all assumed to be 0.2.

Finally, the idiosyncratic shock, or simulated residual, for each simulated stock is drawn from a normal distribution with a mean and cross-sectional spread fixed so that the time series explained variance (R-square) averages to approximately that of the actual data, around 21%. This dictates drawing the cross-section of residual standard deviations from a lognormal Gaussian with a mean of 0.8 and a (cross-sectional) standard deviation of 0.2 for the underlying normal distribution.

It should be noted that the sample R-square is influenced by all of the betas, even those for factors associated with a zero cross-sectional mean. This is due to the fact that the variance of an individual stock's returns depends on all of its squared betas in addition to the variances of all the factors and of the idiosyncratic term.

[22] We use six instead of eight because, as we shall see later, two of the factor candidates must be excluded to avoid the weak instrument problem.

The magnitudes of risk premiums, factor volatilities, betas, idiosyncratic shocks, and R-squares are chosen to be realistic and close to actual equity return parameters. The simulation parameters are summarized in Panel A of Table 7.

Given the simulated factor values, "true" betas, and idiosyncratic shocks, it is straightforward to generate 2517 daily returns for 1100 simulated stocks. Then, our suggested instrumental variables procedure is implemented as follows: the time series observations are divided into three subsets; using only alternate observations 1, 4, 7,..., 2515, time series multiple regressions for each individual stock produce estimated slope coefficients (betas) on the six factors. Then, observations 2, 5, 8,..., 2516, are employed to estimate a second set of betas, which serve as the instruments for the first set. Finally, for each period 3, 6, 9,..., 2517, we run a cross-sectional instrumental variables regression where the dependent variable observations are the 1,100 individual stock simulated returns in that particular period and the explanatory variables are the estimated betas, whose instruments are the betas estimated from alternate observations.

This procedure produces one set of expected return estimates (risk premiums) each period. If returns are independent over time, estimation errors should also be unrelated across periods; hence, any errors in beta are unrelated to the errors in expected returns because their estimation is derived from entirely different observations. Moreover, the errors in beta are unrelated to the errors in the beta instruments, which is a desirable and indeed an essential feature of an instrumental variables.

We also calculate permutations of the above alternating observations procedure wherein the cross-sectional regressions are run again with observations from alternate thirds of the sample. In other words, the cross-sectional regressions are run with observations 1, 4, … and also 2, 5, …; in these cases, betas are estimated with one of the other thirds of the sample and beta instruments with the complementary third. When all these calculations are done, there are cross-sectional risk premium estimates available for every daily observation from 1 to 2517. These are more or less independently distributed because returns are, so they are readily employable to estimate sample risk premiums along with associated standard errors.

The above simulation uses generated factors to generate hypothetical individual stock returns and also to estimate time series regression (beta) coefficients, but it does not assume direct knowledge of the factors' risk premiums, volatilities, or true betas. Consequently, the estimated risk premiums from the final step of the simulation procedure can be legitimately compared against the true risk premiums, which are known from the generating process but not used in the estimation procedure.

Panel B of Table 7 presents the most important simulation results, the estimated risk premiums from the IV procedure along with their associated t-statistics, where the underlying standard errors are computed using the Newey/West autocorrelation correction with two lags. Recall that errors in variables would bias estimated risk premiums toward zero if OLS betas were used in the cross-sectional regressions. We know this because the factor returns are independent of each other. However, as shown in the second column of Panel B, Table 7, the estimated coefficients are close to the true risk premiums (listed in Panel A.) Moreover, the first three are significantly positive, as they should be, while the last three, whose true risk premiums are zero, are not significant.

These encouraging results support a conclusion that our IV procedure, which is in itself a methodological innovation in the asset pricing literature, is sensible and produces risk premiums that have the correct signs and magnitudes, at least when there are a lot of cross-sectional and time series observations.

Panel C of Table 7 reports the correlations for betas and beta instruments in the first two sub-samples. (The correlations between the other subsamples are qualitatively the same.) These numbers provide a gauge of how closely instruments must be to their corresponding betas in order to achieve reasonable large sample results. The correlations here are encouraging. Recall again that estimation errors in betas and beta instruments are unrelated by construction.

In a minor supplement to our small-scale simulation, we also examine the possible confounding influence of stochastic variation in the betas of individual stocks. In this modification, we allow each of the six betas (for every simulated stock) to follow an independent AR(1) process with an autocorrelation parameter of .8 and with a standard deviation of .1 (one half of the cross-sectional standard deviation of betas; see Panel A of Table 7.)[23] The long-term mean of each beta is set equal to the same values as the constant means in the previous simulations, which have the same cross-sectional means and cross-sectional variation as indicated again in the Table 7, Panel A.

Then the IV estimation procedure was implemented without taking any account of the fact that the true betas change significantly over time. This implies, of course, that the procedure is inefficient; stochastic regression in the first stage of the IV procedure would have been superior. Yet, as shown in Panel D of Table 7, the estimated risk premiums are still not far from their true values. The first three factors still have significant risk premiums as they should. The only slightly disturbing aspect of these results is that the factors without true risk premiums, factors 4, 5, and 6, now display substantially larger estimated premiums and the one for factor 4 is marginally significant. Unknown non-stationary betas can evidently compromise IV test power, at least to a modest extent, yet overall the procedure works reasonably well.

Turning now to the actual data, which, corresponding to the just-reported simulations, include 1,100 stocks and 2517 daily return observations. Exactly the same IV procedure is followed; (indeed, the very same Fortran programs are used, just to make sure there are no data processing mistakes.) As with the simulated data, the actual data sample is partitioned into three sub-samples by alternating observations, betas are estimated in each sub-sample using the eight factor candidate found earlier to pass the necessary conditions, and cross-sectional regressions are computed every third day using betas and beta instruments estimated from the other sub-samples.

Table 8 reports summary statistics for representative time series regressions with alternating observations; Panel A using observations 1, 4, 7,…to estimate the first sub-sample betas and Panel B using observations 2, 5, 8…to estimate the betas again, which will serve as instruments for the first sub-set. These regressions include only 1059 of the original 1100; we noticed that many low-priced and/or thinly traded stocks had enormous daily return kurtosis and we exclude 41 of them whose excess kurtosis exceeds 100.

[23] This permits quite a large amount of change over time in the "true" simulated betas.

The explanatory power (adjusted R-square) on average is about 21%, which is roughly what one would expect with daily returns. Notice that the excess kurtoses in the estimated betas for the first four ETFs are relatively large as is that of the intercept. Nonetheless, the t-statistics indicate reasonable significance for all of the coefficients except the beta on IYR.[24]

Table 9, Panel A gives another hint that the actual data might be more complex than the simulated data. This Panel gives the simple correlations between betas from alternating sub-samples. Since one set of sub-sample betas will serve as instruments for betas from its companion sub-sample, these correlations should be respectable; ideally, they would be quite high as they are for the simulated data (Cf. Panel C of Table 7.) But for the first two ETFs, EWH and EWJ, the correlations are very disappointing and well below a reasonable level such as 0.5. This is strong evidence of a "weak instrument" problem, which is well-known to potentially produce nonsensical results. Consequently, we are eliminating these two ETFs from consideration and will not use them in the next stage, the instrumental variable cross-sectional regressions.

The other six factor candidates, (ILF, IYR, SPY, SMB, HML and MOM,) have inter-subsample beta correlations mostly below those observed for the simulated data, but they are at least in the .5 to .7 range. We thus retain them in the next stage, but with some trepidation.

Panel B of Table 9 reports the risk premium estimates for the four retained factor candidates, using the same procedure as explained earlier for the simulated data and reported in Panel B of Table 7. To say the least, the results are underwhelming. Although all the risk premiums except that for IYR are positive and in a reasonable range, only Momentum is statistically significant.

There are a number of possible explanations to consider for this puzzling result. First, our instrumental variable procedure may be flawed. It works very well for simulated data, but the real data may be so bizarre that the Gaussian simulated data are highly unrepresentative. Note, however, that the IV method does not require the data to be normally distributed. The IV estimates are unbiased and are asymptotically normal under fairly unrestrictive assumptions. Second, we might have weak instruments, even though their estimation errors are independent of the

[24] However, these t-statistics are likely overstated because the regressions are not cross-sectionally independent.

explanatory variables and they have at least decent correlations. Moreover, in the final stage we have eliminated factor candidates that seem to have rather weak instruments. Third, even though these factors may be associated with risk premiums over an extended period, the ex post reward for risk in the 2002-2011 decade might have been minimal. Fourth, although these factors are indeed associated with the covariance matrix of returns, they are diversifiable and therefore do not carry risk premiums.

We doubt the validity of the first two explanations because the IV procedure seems to work well with simulated data and because we have expunged weak instruments. As for the third and fourth explanations, it can only be settled one way or the other by expanding the data sample over a longer period.

Some insight about the third explanation, however, can be gleaned by simply examining the average returns of the factors over the 2002-2011 sample decade; these are reported in Panel C of Table 9. An excess mean return for each factor would be the numbers in the second column less a riskless rate. For instance, SPY's mean annual return was about 0.02087 X 250 or 7.175% per annum, which seems to be more or less representative of a long-term expected level. This mean was not, however, significantly different from zero during the decade. Indeed, only ILF has a mean return significant at a customary level.

Note that the reported excess kurtosis is very large for all the factor returns; this is an indication of non-normality and compromises any significance inferences about the means. As mentioned earlier, however, a large kurtosis is not necessarily a serious problem for our IV procedure.

As a final check, we extract ten principal components (PCs) from the daily covariance matrix of the returns of our 1100 stocks and repeat the IV procedure using the PCs rather than the factor candidates. It turns out that the alternating observation betas suffer from the weak instruments problem for six of the ten PCs. For the others, only one PC displays a statistically significant risk premium while a second PC is marginally significant (results not reported.) Hence, one might be entitled to conclude that the 2002-2011 decade had risk drivers that were only weakly associated with rewards.

IX. Summary and Conclusions.

Our goal in this paper is to suggest a protocol for sorting 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 sound and acceptable. Ours is just a first attempt.

We begin with an empirical observation: asset returns reveal an underlying factor structure because diversification is not all that powerful. Moreover, weak correlations across diversified portfolios in different asset classes and/or countries suggest that there must be multiple factors.

An underlying factor cannot have movements that are easily predictable because asset prices adjust in advance. One implication is that a characteristic cannot be a factor. This rules out firm-specific attributes such as size, dividend yield, book/market and so on. Such characteristics can be related to factor loadings or exposures, but they cannot be factors per se because they are known. Over 300 factors and characteristics have been claimed by the extant literature to explain expected returns. These characteristics might be related to risk exposures (i.e., to "betas" on unknown risk drivers) but they might also be symptomatic of arbitrage opportunities. Our protocol would ascertain their true nature.

Our suggested protocol has two stages. The first stage provides a sequence of six 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 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 diversifiable even though they affect some real assets but not all of them. We propose a new variant of the Fama/MacBeth (1973) tests using individual (real) assets and instrumental variables to overcome the errors-in-variables problem induced by the estimation of factor exposures. Our simple empirical example of the protocol suggests that very large cross-sectional and time series samples are required to correctly identify factors. Although simulations support our proposed procedure in this last stage, there is no evidence of risk premiums associated with any of the example factors we tried except momentum, even though six of them survive the necessary conditions of our protocol.

References

Anderson, T. W., 1984, An Introduction to Multivariate Statistical Analysis, 2nd ed., (New York: Wiley.)

Berger, Dave, Kuntara Pukthuanthong, and Richard Roll, 2014, On Valuing Human Capital And Relating it to Macro-Economic Conditions, working paper, UCLA.

Brennan, Michael C., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics and the cross section of expected stock returns, *Journal of Financial Economics* 49, 345–373.

Brown, S., 1989, The number of factors in security returns, *Journal of Finance* 44, 1247-1262.

Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley, 2014, An Interteporal CAPM with Stochastic Volatility, working paper, Department of Economics, Harvard University, April.

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

Charoenrook, Anchada and Jennifer Conrad, 2008, Identifying Risk-Based Factors, working paper, The Owen School at Vanderbilt University.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic Forces and the Stock Market, *Journal of Business*, 59, 383-403.

Cochrane, John H., 2001, *Asset Pricing*, Princeton, Princeton University Press.

Connor, Gregory, and Robert A. Korajczyk, 1988, Risk and Return in an Equilibrium APT: Application of a New Test Methodology, *Journal of Financial Economics* 21, 255-289.

Cotter, John, and Richard Roll, 2012, A Comparative Anatomy of REITs and Residential Real Estate Indexes: Returns, Risks and Distributional Characteristics, working paper, UCLA.

Daniel, Kent, David Hirshleifer and Avanidhar Subrahmanyam, 2005, Investor Psychology and Tests of Factor Pricing Models, working paper, UCLA.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the Characteristics of Cross Sectional Variation in Stock Returns, *Journal of Finance* 56, 1-33.

Fama, Eugene, and Kenneth French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427-466.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607-636.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 1999, The alpha factor asset pricing model: A parable, *Journal of Financial Markets* 2, 49-68.

Goyal, Amit, 2012, Empirical cross-sectional asset pricing: a survey, *Financial Markets and Portfolio Management* 26(1), 3−38.

Graham, Nathaniel, 2012, Pricing beta: Parameter uncertainty and factor models, working paper, University of Kentucky.

Green, Jeremiah, John Hand, and X. Frank Zhang, 2013, The supraview of return predictive signals, Review of Accounting Studies 18, 692-730.

Green, Jeremiah, John Hand, and X. Frank Zhang, 2014, The remarkable multidimensionality in the cross-section of expected U.S. stock returns, working paper, Pennsylvania State University.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2013, …and the Cross-Section of Expected Returns," Forthcoming *Review of Financial Studies*.

Johnson, Richard A., and Dean W. Wichern, 2007, Applied Multivariate Statistical Analysis, 6th ed., (New York: Prentice-Hall.)

King, Benjamin F., 1966, Market and Industry Factors in Stock Price Behavior, *Journal of Business* 39, 139-190.

Ledoit, Olivier, Pedro Santa-Clara, and Michael Wolf, 2003, Flexible Multivariate Garch Modeling with an Application to International Stock Markets, *The Review of Economics and Statistics* 85, 735-747.

Lewellen, Jonathan, 2014, The cross-section of expected stock returns, forthcoming Critical Finance Review.

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

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175-194.

McLean, David, Jeffrey Pontiff, 2014, Does academic research destroy stock return predictability?, working paper, Boston College and University of Alberta.

Moskowitz, Tobias J., 2003, An Analysis of Covariance Risk and Pricing Anomalies, *The Review of Financial Studies* 16, 417-457.

Petersen, Mitchell, A., 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* 22, 435-480.

Pukthuanthong, Kuntara, and Richard Roll, 2009, Global Market Integration: An Alternative Measure and Its Application, *Journal of Financial Economics* 94, 214-232.

Roll, Richard and Stephen A. Ross, 1980, An Empirical Investigation of the Arbitrage Pricing Theory, *Journal of Finance* 35, 1073-1103.

Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2011, An Empirical Investigation of Volume in Equity-Contingent Claims, working paper, UCLA.

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*, Princeton University Press.

Subrahmanyam, Avanidhar, 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.

Table 1
Sample stocks

Out of 1,100 firms, thirty are selected randomly for analysis. The random selection simply uses the first thirty stocks listed by Permno. Ten principal components are extracted from the unconditional covariance matrix of these stocks, whose CRSP Permnos, CUSIPs, Tickers, and Names are shown here. The sample period is January 2002 to December 2011 inclusive.

Permno	CUSIP	Ticker	Company Name
10026	46603210	JJSF	J & J Snack Foods Corp
10044	77467810	RMCF	Rocky Mountain Chocolate Fac Inc
10100	27135100	ALRN	American Learning Corp
10107	59491810	MSFT	Microsoft Corp
10138	74144T10	TROW	T Rowe Price Group Inc
10163	98385710	XRIT	X Rite Inc
10200	75991610	RGEN	Repligen Corp
10239	57755200	BWINB	Baldwin & Lyons Inc
10258	15117B10	CLDX	Celldex Therapeutics Inc
10259	82656510	SIGM	Sigma Designs Inc
10272	87910110	TKLC	Tekelec
10299	53567810	LLTC	Linear Technology Corp
10302	23280610	CY	Cypress Semiconductor Corp
10318	57665200	BCPC	Balchem Corp
10355	23391210	DJCO	Daily Journal Corp
10363	00163U10	AMAG	A M A G Pharmaceuticals Inc
10382	46224100	ASTE	Astec Industries Inc
10395	63890410	NAVG	Navigators Group Inc
10397	95075510	WERN	Werner Enterprises Inc
10463	76091110	REFR	Research Frontiers Inc
10530	58958410	VIVO	Meridian Bioscience Inc
10547	18482P10	CLFD	Clearfield Inc
10550	74265M20	PDEX	Pro Dex Inc Colo
10644	88337510	TGX	Theragenics Corp
10645	51179510	LAKE	Lakeland Industries Inc
10656	44461000	ACET	Aceto Corp
10781	65066100	BKSC	Bank South Carolina Corp
10812	42234710	HTLD	Heartland Express Inc
10838	23282830	CYTR	Cytrx Corp
10853	76017410	RENT	Rentrak Corp

Table 2

Exchange Traded Funds

Permno, ticker, company name, and names of the ETFs in our sample are presented below. These ETFs are heterogeneous and have unconditional correlations less than 0.95 with each other. The sample period is from January 2002 to December 2011.

Permno	Ticker	Company	ETF name
89187	EPP	iShares	MSCI Pacific ex-Japan
88405	IEV	iShares	S&P Europe 350 Index
83215	EWH	iShares	MSCI Hong Kong Index
83225	EWJ	iShares	MSCI Japan Index
88290	ILF	iShares	S&P Latin America 40 Index
88294	IYR	iShares	Dow Jones US Real Estate
84398	SPY	SPDR	S&P 500

Table 3

Unconditional covariance matrix of All Assets

The unconditional covariance matrix with dimension 20 X 20 is reported below for ten principal components extracted from firms shown in Table 1, seven ETFs shown in Table 2, the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor (MOM). The sample period spans 120 months from January 2002 through December 2011. For clarity, the matrix is reported with standard deviations along the diagonal and correlations off diagonal. The principal components are orthogonal to each other by construction and are normalized to have unit variance. For the other ten assets, the factor candidates, standard deviations are shown in natural units of percent per month.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	EPP	IEV	EWH	EWJ	ILF	IYR	SPY	SMB	HML	MOM
PC1	1.00																			
PC2	0.00	1.00																		
PC3	0.00	0.00	1.00																	
PC4	0.00	0.00	0.00	1.00																
PC5	0.00	0.00	0.00	0.00	1.00															
PC6	0.00	0.00	0.00	0.00	0.00	1.00														
PC7	0.00	0.00	0.00	0.00	0.00	0.00	1.00													
PC8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00												
PC9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00											
PC10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00										
EPP	0.58	0.00	0.16	-0.14	0.01	0.02	0.22	0.12	0.08	0.06	6.67									
IEV	0.61	-0.08	0.18	-0.15	0.00	0.06	0.27	0.14	0.10	0.05	0.89	6.14								
EWH	0.53	-0.04	0.13	-0.18	-0.04	0.01	0.21	0.05	0.04	0.05	0.86	0.77	6.54							
EWJ	0.44	-0.02	0.23	-0.05	-0.05	0.04	0.12	0.04	0.08	0.07	0.67	0.67	0.55	5.18						
ILF	0.56	0.07	0.21	-0.22	0.03	-0.08	0.26	0.02	0.15	0.03	0.80	0.74	0.77	0.55	7.68					
IYR	0.48	-0.09	0.31	-0.13	-0.04	0.16	0.21	0.01	0.12	0.05	0.70	0.70	0.58	0.60	0.57	7.50				
SPY	0.71	0.00	0.21	-0.18	-0.05	0.04	0.25	0.16	0.12	0.03	0.85	0.90	0.74	0.63	0.75	0.74	4.62			
SMB	0.56	-0.06	0.24	-0.01	0.08	0.14	0.21	-0.11	-0.02	0.09	0.29	0.27	0.29	0.30	0.40	0.43	0.32	2.49		
HML	0.01	-0.06	0.22	0.16	0.08	0.13	0.22	-0.02	0.16	0.07	0.26	0.26	0.04	0.35	0.14	0.48	0.23	0.21	2.51	
MOM	-0.36	0.05	-0.15	0.06	0.06	0.04	-0.06	0.08	-0.02	0.02	-0.28	-0.39	-0.31	-0.27	-0.37	-0.41	-0.44	-0.08	-0.16	5.42

Table 4

Conditional covariance matrix for an example month, December 2006

This table presents the estimated conditional covariance matrix for December 2006, a month roughly halfway through the overall 2002-2011 sample. Conditional covariance matrices are constructed each month using the flexible multivariate procedure of Ledoit, Santa Clara, and Wolf (2003). Each conditional covariance matrix has dimension 20 X 20 and portrays ten principal components extracted from firms shown in Table 1, seven ETFs shown in Table 2, the Fama-French (1992) SMB and HML factor, and the Carhart (1997) momentum factor (MOM). For clarity, the example monthly matrix below is displayed with standard deviations (%) along the diagonal and correlations off diagonal. The first ten assets are the principal components and are unconditionally orthogonal to each other by construction and scaled to have unit unconditional variance. Standard deviations for the other ten assets, the factor candidates, are in natural units of percent per month. The corresponding unconditional matrix is reported in Table 3.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	EPP	IEV	EWH	EWJ	ILF	IYR	SPY	SMB	HML	MOM
PC1	0.89																			
PC2	-0.09	0.91																		
PC3	0.18	0.00	0.90																	
PC4	0.10	-0.03	-0.03	0.97																
PC5	0.01	-0.01	0.06	-0.08	0.92															
PC6	-0.02	-0.04	-0.01	-0.02	-0.09	0.96														
PC7	0.07	-0.06	0.05	-0.12	0.08	0.06	0.85													
PC8	0.12	-0.04	0.00	0.02	-0.13	0.08	0.10	0.92												
PC9	-0.02	-0.06	0.07	-0.05	-0.07	0.04	0.18	-0.02	0.91											
PC10	0.21	-0.05	0.03	-0.01	-0.05	0.00	0.03	0.04	0.06	1.00										
EPP	0.61	-0.01	0.27	-0.11	0.03	-0.02	0.25	0.06	0.13	0.16	5.42									
IEV	0.61	-0.09	0.31	-0.09	-0.10	-0.01	0.28	0.14	0.20	0.25	0.82	3.66								
EWH	0.61	-0.03	0.23	-0.18	0.42	0.02	0.19	0.06	0.09	0.12	0.70	0.51	6.02							
EWJ	0.52	0.06	0.29	-0.07	-0.07	-0.04	0.10	0.03	0.13	0.06	0.50	0.48	0.50	3.51						
ILF	0.60	0.00	0.35	-0.22	-0.04	-0.12	0.29	-0.04	0.18	0.16	0.68	0.68	0.54	0.44	6.39					
IYR	0.64	-0.14	0.31	-0.05	-0.07	0.08	0.19	0.06	0.14	0.04	0.65	0.61	0.55	0.50	0.54	6.00				
SPY	0.61	-0.09	0.29	-0.14	-0.13	0.00	0.29	0.05	0.18	0.21	0.74	0.85	0.47	0.44	0.66	0.69	2.69			
SMB	0.34	-0.09	0.37	0.00	0.09	0.18	0.31	-0.16	0.02	0.12	0.25	0.30	0.27	0.39	0.43	0.40	0.36	2.39		
HML	-0.04	-0.05	0.14	0.33	-0.08	0.14	0.13	0.02	0.01	-0.06	0.12	0.12	0.01	0.27	-0.01	0.17	-0.01	0.01	1.71	
MOM	-0.12	0.08	-0.05	0.12	0.09	0.21	0.14	-0.03	-0.02	0.17	0.13	-0.07	0.16	-0.03	-0.08	-0.11	-0.08	0.17	0.06	2.23

Table 5

Canonical Correlations

Canonical Correlations between Factor Candidates, the seven ETFs listed in Table 2 plus the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor (MOM), versus ten principal components computed from the random sample of thirty stocks listed in Table 1. The first column reports the ten canonical correlations computed from the unconditional covariance matrix. The second column reports the time series means of canonical correlations from the conditional covariance matrices computed with the flexible multivariate procedure of Ledoit, Santa Clara, and Wolf (2003). The canonical correlations are sorted in descending order by their estimated squares. The right-most column gives Newey-West T-statistics (ten lags) for the mean correlations in the second column.

Unconditional	Mean Conditional	Newey-West T-Statistic
0.921	0.840	85.68
-0.517	-0.199	-2.74
0.442	0.271	5.10
0.350	0.073	2.05
-0.317	0.104	3.62
-0.232	0.060	2.32
0.130	0.041	2.17
-0.096	0.018	1.42
0.082	0.009	1.56
-0.016	-0.003	-0.89

Table 6

Weightings for First Canonical Variates

Weightings are reported below for canonical variates that yield the largest canonical correlation. The first canonical variate (left panel) is a weighted average of factor candidates, the seven ETFs listed in Table 2 plus the Fama-French (1992) SMB and HML factors, and the Carhart (1997) momentum factor (MOM). The second canonical variate (right panel) is a weighted average of ten principal components computed from the random sample of thirty stocks listed in Table 1. The unconditional weightings are from the covariance matrix computed over 120 months, January 2002 through December 2011 inclusive. The time series mean conditional weightings are from conditional covariance matrices computed with the flexible multivariate procedure of Ledoit, Santa Clara, and Wolf (2003) in each month, February 2002 through December 2011. Newey-West T-statistics (ten lags) are for the mean conditional weighting.

	Unconditional Weighting	Mean Conditional Weighting	Newey-West T-Statistic		Unconditional Weighting	Mean Conditional Weighting	Newey-West T-Statistic
	Factor Candidates				Principal Components		
EPP	-0.0113	-0.0079	-1.592	PC1	0.8919	0.9073	36.76
IEV	0.0275	0.0005	0.065	PC2	-0.0292	0.0268	5.346
EWH	-0.0100	0.0914	16.12	PC3	0.2592	0.2109	12.69
EWJ	-0.0057	0.0236	5.121	PC4	-0.1663	-0.2518	-11.55
ILF	-0.0014	0.0232	6.495	PC5	-0.0080	0.1957	8.460
IYR	-0.0126	0.0195	4.980	PC6	0.0861	0.0209	1.147
SPY	0.1670	0.0382	4.868	PC7	0.2960	0.1470	10.43
SMB	0.2044	0.0781	7.015	PC8	0.0793	-0.0363	-4.824
HML	-0.0479	-0.0724	-5.273	PC9	0.0663	0.1249	12.49
MOM	-0.0157	-0.0484	-4.479	PC10	0.0553	-0.0102	-1.850

Table 7

Simulation of the Instrumental Variable Procedure for Estimating Factor Risk Premiums

Corresponding to the actual data used here, Gaussian-distributed returns are simulated for six factors over 2517 days. Betas are generated from a cross-sectional Gaussian distribution for 1100 hypothetical stocks. The cross-section of idiosyncratic volatility is generated from a lognormal distribution. Panel A presents the parameters used in the simulation. Using these parameters, 2517 daily returns are generated with independent idiosyncratic shocks for 1100 stocks. Next, our suggested instrumental variable (IV) procedure is implemented with the simulated individual stock returns. The procedure divides the sample into three alternating observation sub-samples. First, betas for each stock are estimated from time series regressions using daily observations 1, 4, 7,.., 2515, which we denote the first sub-sample. Then, beta instruments are estimated from time series regressions using days 2, 5, 8,..., 2516, the second sub-sample. Finally, cross-sectional regressions are estimated for each day, 3, 6, 9,..., 2517, the third sub-sample. Two other permutations are also employed: (1) sub-sample one for estimating betas, sub-sample three for estimating beta instruments, and sub-sample two for the cross-sectional regression and (2) sub-sample two for estimating betas, sub-sample three for estimating beta instruments, and sub-sample one for the cross-sectional regression. Since the returns are roughly independent across time, the three sets of cross-sectional coefficients form a time series whose means in percent per day are reported Panel B along with Newey/West t-statistics with two lags. For reference, Panel C reports the cross-correlations between estimated betas and beta instruments, using the first two sub-samples as an example.

Panel A.
Simulation Parameters

Factor	True Risk Premium (%/day)	Standard Deviation (%/day)	Betas: Cross-Sectional Mean	Betas: Standard Deviation
1	.04%	.4	1.0	.2
2	.03%	.3	Zero	.2
3	.02%	.3	Zero	.2
4	Zero	.2	Zero	.2
5	Zero	.2	Zero	.2
6	Zero	.2	Zero	.2

Idiosyncratic Shock
lognormal cross-sectional distribution

Mean of logs: 0.8	Standard deviation: 0.2

Panel B.
Estimated Risk Premiums

Factor	Mean IV Coefficient	T-Statistic
1	.0496%	5.86
2	.0275%	4.21
3	.0189%	2.86
4	.0011%	.234
5	-.0007%	-.139
6	-.0002%	-.223

Panel C.
Correlations between Betas and Their Instruments, sub-samples one and two

1	.891
2	.816
3	.827
4	.674
5	.704
6	.685

Panel D.
Estimated Risk Premiums
with Stochastic Betas

Factor	Mean IV Coefficient	T-Statistic
1	.0463%	5.50
2	.0234%	3.60
3	.0183%	2.81
4	.0098%	1.92
5	-.0026%	-.518
6	-.0071%	-1.41

Table 8

Summary Statistics for Betas and Beta Instruments
Estimated from Alternate Daily Observations

With returns on 1,100 non-debt equities over ten years, 2002-2011, as dependent variables and, as explanatory variables, five ETFs, the Fama/French SMB size and HML book/market factors and Momentum (all of which are significant in Table 6), multiple time series regressions estimate the slope coefficients (betas) using daily observations numbered 1, 4, 7,…, 2515 (sub-sample #1.) Beta instruments are estimated from observations 2, 5, 8,…, 2516 (sub-sample #2.) Summary statistics calculated across the 1,100 sets of estimates are reported here.

	Mean	Sigma	T-Stat	Skewness	Kurtosis	Maximum	Minimum
Regressions using sub-sample #1, observations 1, 4, 7…							
EWH	0.016	0.142	3.752	0.426	3.033	1.017	-0.513
EWJ	0.009	0.149	2.057	-0.039	3.705	1.072	-0.770
ILF	0.086	0.183	15.278	1.163	2.022	0.937	-0.422
IYR	0.005	0.210	0.804	0.801	4.837	1.333	-0.827
SPY	0.698	0.450	50.473	-0.151	-0.037	2.122	-0.803
SMB	0.827	0.540	49.822	0.139	-0.277	2.571	-0.859
HML	0.080	0.483	5.381	0.099	1.486	2.864	-1.468
MOM	-0.104	0.259	-13.062	-0.239	1.014	0.798	-1.125
Intercept	0.048	0.121	13.060	0.984	3.308	0.785	-0.324
Adj. R^2	0.207	0.159	42.483	0.578	-0.061	0.798	-0.006
Regressions using sub-sample #2, observations 2, 5, 8…							
EWH	-0.011	0.151	-2.279	-0.760	5.610	0.704	-0.985
EWJ	0.029	0.159	5.879	0.472	2.146	0.814	-0.586
ILF	0.066	0.177	12.148	1.047	2.342	0.879	-0.542
IYR	-0.007	0.219	-1.070	0.915	5.193	1.347	-0.842
SPY	0.763	0.440	56.385	-0.070	0.170	2.161	-1.098
SMB	0.797	0.571	45.388	0.122	-0.323	2.805	-1.150
HML	0.055	0.493	3.644	0.388	3.021	3.735	-1.493
MOM	-0.085	0.248	-11.174	-0.410	0.622	0.566	-1.099
Intercept	0.056	0.113	16.118	0.762	3.560	0.774	-0.404
Adj. R^2	0.219	0.162	44.053	0.499	-0.188	0.824	-0.005

Table 9

Estimating Risk Premiums for Factors that Satisfy the Necessary Conditions

Our suggested instrumental variable (IV) procedure is implemented with 2517 daily returns over the 2002-2011 decade for 1100 non-debt stocks and eight factor candidates that survive our protocol of necessary conditions. The procedure divides the sample into three alternating observation sub-samples. First, betas for each stock are estimated from time series regressions using daily observations 1, 4, 7,.., 2515, which we denote the first sub-sample. Then, beta instruments are estimated from time series regressions using days 2, 5, 8,..., 2516, the second sub-sample. Finally, cross-sectional regressions are estimated for each day, 3, 6, 9,..., 2517, the third sub-sample. Two other permutations are also employed: (1) sub-sample one for estimating betas, sub-sample three for estimating beta instruments, and sub-sample two for the cross-sectional regression and (2) sub-sample two for estimating betas, sub-sample three for estimating beta instruments, and sub-sample one for the cross-sectional regression. Panel A reports correlation coefficients between betas and beta instruments for the three sub-sample permutations. The first two ETFs listed display weak instruments and are therefore dropped from the cross-sectional IV regressions. The three sets of cross-sectional coefficients form a time series whose means in percent per day are reported Panel B along with Newey/West t-statistics with two lags. These are the estimated risk premium. Panel C reports statistics pertaining to the daily factor returns over the 2002-2011 decade for the four surviving factor candidates.

Panel A.
Correlations between Betas and Their Instruments

Factor	Sub-samples		
	1 & 2	1&3	2&3
EWH	0.111	0.078	0.173
EWJ	0.143	0.092	0.114
ILF	0.607	0.531	0.482
IYR	0.628	0.617	0.602
SPY	0.702	0.698	0.701
SMB	0.765	0.761	0.781
HML	0.631	0.586	0.565
MOM	0.510	0.518	0.478

Panel B.
Estimated Risk Premiums

Factor	Mean IV Coefficient (%/day)	T-Statistic
ILF	0.0586	0.758
IYR	-0.0083	-0.148
SPY	0.0089	0.215
SMB	0.0160	0.864
HML	0.0053	0.241
MOM	0.1544	2.924

Panel C.
Factor Returns, 2002-2011

Factor	Mean (%/day)	T-Statistic	Excess Kurtosis
ILF	0.09337	2.055	13.702
IYR	0.05709	1.308	12.625
SPY	0.02087	0.757	11.519
SMB	0.01482	1.244	3.886
HML	0.1160	0.987	7.165
MOM	0.0059	0.279	9.213