Characteristics *Are* Covariances: A Unified Model of Risk and Return*

Bryan Kelly
Yale University, NBER, and AQR Capital Management

Seth Pruitt
Arizona State University

Yinan Su
University of Chicago

February 15, 2018

Abstract

We propose a new modeling approach for the cross section of returns. Our method, Instrumented Principal Components Analysis (IPCA), allows for latent factors and time-varying loadings by introducing observable characteristics that instrument for the unobservable dynamic loadings. If the characteristics/expected return relationship is driven by compensation for exposure to latent risk factors, IPCA will identify the corresponding latent factors. If no such factors exist, IPCA infers that the characteristic effect is compensation without risk and allocates it to an "anomaly" intercept. Studying returns and characteristics at the stock-level, we find that four IPCA factors explain the cross section of average returns significantly more accurately than existing factor models and produce characteristic-associated anomaly intercepts that are small and statistically insignificant. Furthermore, among a large collection of characteristics explored in the literature, only eight are statistically significant in the IPCA specification and are responsible for nearly 100% of the model's accuracy.

Keywords: Cross section of returns, latent factors, anomaly, factor model, conditional betas, PCA, BARRA

*Corresponding author email: bryan.kelly@yale.edu. We thank Svetlana Bryzgalova (discussant), John Cochrane, Kent Daniel (discussant), Rob Engle, Wayne Ferson, Stefano Giglio, Lars Hansen, Andrew Karolyi, Serhiy Kozak, Toby Moskowitz, Andreas Neuhierl, Dacheng Xiu and audience participants at AQR, ASU, Chicago, Cornell, Duke, FRA, UMass Amherst, Minnesota, SoFiE, St. Louis Federal Reserve, and Yale for helpful comments. We are grateful to Andreas Neuhierl for generously sharing data with us. A previous draft was circulated under the name "Some Characteristics Are Risk Exposures, and the Rest Are Irrelevant." AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

One of our central themes is that if assets are priced rationally, variables that are related to average returns, such as size and book-to-market equity, must proxy for sensitivity to common (shared and thus undiversifiable) risk factors in returns.
Fama and French (1993)

We have a lot of questions to answer: First, which characteristics really provide independent information about average returns? Which are subsumed by others? Second, does each new anomaly variable also correspond to a new factor formed on those same anomalies? ... Third, how many of these new factors are really important? Cochrane (2011)

1 Introduction

The greatest collective endeavor of the asset pricing field in the past 40 years is the search for an empirical explanation of why different assets earn different average returns. The answer from equilibrium theory is clear—differences in expected returns reflect compensation for different degrees of risk. But the empirical answer has proven more complicated, as some of the largest differences in performance across assets continue to elude a reliable risk-based explanation.

This empirical search centers around return factor models, and arises from the Euler equation for investment returns. With only the assumption of "no arbitrage," a stochastic discount factor m_{t+1} exists and, for any excess return $r_{i,t}$, satisfies the equation

$$E_t[m_{t+1}r_{i,t+1}] = 0 \quad \Leftrightarrow \quad E_t[r_{i,t+1}] = \underbrace{\frac{Cov_t(m_{t+1}, r_{i,t+1})}{Var_t(m_{t+1})}}_{\beta'_{i,t}} \underbrace{\frac{Var_t(m_{t+1})}{E_t[m_{t+1}]}}_{\lambda_t}. \tag{1}$$

The loadings, $\beta_{i,t}$, are interpretable as exposures to systematic risk factors, and λ_t as the risk prices associated with those factors. More specifically, when m_{t+1} is linear in factors f_{t+1}, this maps to a factor model for excess returns of the form[1]

$$r_{i,t+1} = \alpha_{i,t} + \beta'_{i,t} f_{t+1} + \epsilon_{i,t+1} \tag{2}$$

where $E_t(\epsilon_{i,t+1}) = E_t[\epsilon_{i,t+1} f_{t+1}] = 0$, $E_t[f_{t+1}] = \lambda_t$, and, perhaps most importantly, $\alpha_{i,t} = 0$ for all i and t. The factor framework in (2) that follows from the asset pricing Euler equation

[1] Ross (1976), Hansen and Richard (1987).

(1) is the setting for most empirical analysis of expected returns across assets.

There are many obstacles to empirically analyzing equations (1) and (2), the most important being that the factors and loadings are unobservable.[2] There are two common approaches that researchers take.

The first pre-specifies factors based on previously established knowledge about the empirical behavior of average returns, treats these factors as fully observable by the econometrician, and then estimates betas and alphas via regression. This approach is exemplified by Fama and French (1993). A shortcoming of this approach is that it requires previous understanding of the cross section of average returns. But this is likely to be a partial understanding at best, and at worst is exactly the object of empirical interest.[3]

The second approach is to treat risk factors as latent and use factor analytic techniques, such as PCA, to simultaneously estimate the factors and betas from the panel of realized returns, a tactic pioneered by Chamberlain and Rothschild (1983) and Connor and Korajczyk (1986). This method uses a purely statistical criterion to derive factors, and has the advantage of requiring no ex ante knowledge of the structure of average returns. A shortcoming of this approach is that PCA is ill-suited for estimating conditional versions of equation (2) because it can only accommodate static loadings. Furthermore, PCA lacks the flexibility for a researcher to incorporate other data beyond returns to help identify a successful asset pricing model.

1.1 Our Methodology

In this paper, we use a new method called instrumented principal components analysis, or IPCA, that estimates market risk factors and loadings by exploiting beneficial aspects of both approaches while bypassing many of their shortcomings. IPCA allows factor loadings to partially depend on observable asset characteristics that serve as instrumental variables for the latent conditional loadings.[4]

---

[2] Even in theoretical models with well defined risk factors, such as the CAPM, the theoretical factor of interest is generally unobservable and must be approximated, as discussed by Roll (1977).

[3] Fama and French (1993) note that "Although size and book-to-market equity seem like ad hoc variables for explaining average stock returns, we have reason to expect that they proxy for common risk factors in returns. ... But the choice of factors, especially the size and book-to-market factors, is motivated by empirical experience. Without a theory that specifies the exact form of the state variables or common factors in returns, the choice of any particular version of the factors is somewhat arbitrary."

[4] Our terminology is inherited from generalized method of moments usage of "instrumental variables" for approximating conditioning information sets. See Hansen (1982), or Cochrane (2005) for a textbook treatment in the context of asset pricing.

The IPCA mapping between characteristics and loadings provides a formal statistical bridge between characteristics and expected returns, while at the same time remaining consistent with the equilibrium asset pricing principle that risk premia are solely determined by risk exposures. And, because instruments help consistently recover loadings, IPCA is then also able to consistently estimate the latent factors associated with those loadings. In this way, IPCA allows a factor model to incorporate the robust empirical fact that stock characteristics provide reliable conditioning information for expected returns. By including instruments, the researcher can leverage previous, but imperfect, knowledge about the structure of average returns in order to improve their estimates of factors and loadings, without the unrealistic presumption that the researcher can correctly specify the exact factors a priori.

Our central motivation in developing IPCA is to build a model and estimator that admits the possibility that characteristics line up with average returns because they proxy for loadings on common risk factors. Indeed, if the "characteristics/expected return" relationship is driven by compensation for exposure to latent risk factors, IPCA will identify the corresponding latent factors and betas. But, if no such factors exist, the characteristic effect will be ascribed to an intercept. This immediately yields an intuitive intercept test that discriminates whether a characteristic-based return phenomenon is consistent with a beta/expected return model, or if it is compensation without risk (a so-called "anomaly"). This test generalizes alpha-based tests such as Gibbons, Ross, and Shanken (1989, GRS). Rather than asking the GRS question "do some pre-specified factors explain the anomaly?," our IPCA test asks "Does there *exist* some set of common latent risk factors that explain the anomaly?" It also provides tests for the importance of particular groups of instruments while controlling for all others, analogous to regression-based t and F tests, and thus offers a means to address questions raised in the Cochrane (2011) quote above.

A standard protocol has emerged in the literature: When researchers propose a new characteristic that aligns with future asset returns, they build portfolios that exploit the characteristic's predictive power and test the alphas of these portfolios relative to some previously established pricing factors (such as those from Fama and French, 1993 or 2015). This protocol is unsatisfactory as it fails to fully account for the gamut of proposed characteristics in prior literature. Our method offers a different protocol that treats the multivariate nature of the problem. When a new anomaly characteristic is proposed, it can be included in an IPCA specification that also includes the long list of characteristics from past studies. Then, IPCA can estimate the proposed characteristic's marginal contribution to the model's factor loadings and, if need be, its anomaly intercepts, after controlling for other characteristics in a complete multivariate analysis.

Additional IPCA features make it ideally suited for state-of-the-art asset pricing analyses. One is its ability to jointly evaluate large numbers of characteristic predictors with minimal computational burden. It does this by building a dimension reduction directly into the model. The cross section of assets may be large and the number of potential predictors expansive, yet the model's insistence on a low-dimension factor structure imposes parameter parsimony. Our approach is designed to handle situations in which characteristics are highly correlated, noisy, or even spurious. The dimension reduction picks a few linear combinations of characteristics that are most informative about patterns in returns and discards the remaining uninformative high dimensionality. Another is that it can nest traditional, pre-specified factors within a more general IPCA specification. This makes it easy to test the incremental contribution of estimated latent IPCA factors while controlling for other factors from the literature.

1.2 Findings

Our analysis judges asset pricing models on two criteria. First, a successful factor model should excel in describing the common variation in realized returns. That is, it should accurately describe systematic risks. We measure this according to a factor model's total panel R^2. We define total R^2 as the fraction of variance in $r_{i,t}$ described by $\hat{\beta}'_{i,t-1}\hat{f}_t$, where $\hat{\beta}_{i,t-1}$ are estimated dynamic loadings and $\hat{f}_t$ are the model's estimated common risk factors. The total R^2 thus includes the explained variation due to contemporaneous factor realizations and dynamic factor exposures, aggregated over all assets and time periods.

Second, a successful asset pricing model should describe differences in average returns across assets. That is, it should accurately describe risk compensation. To assess this, we define a model's predictive R^2 as the explained variation in $r_{i,t}$ due to $\hat{\beta}'_{i,t-1}\hat{\lambda}$, which is the model-based conditional expected return on asset i given $t-1$ information, and where $\hat{\lambda}$ is the vector of estimated factor risk prices.[5]

Our empirical analysis uses data on returns and characteristics for over 12,000 stocks from 1962–2014. For example, in the specification with four factors and all stock-level intercepts restricted to zero, IPCA achieves a total R^2 for returns of 19.4%. As a benchmark, the matched sample total R^2 from the Fama-French five-factor model is 21.9%. Thus, IPCA is a competitive model for describing the variability and hence riskiness of stock returns.

[5]We discuss our definition of predictive R^2 in terms of the unconditional risk price estimate, $\hat{\lambda}$, rather than a conditional risk price estimate, in Section 4.

Perhaps more importantly, the factor loadings estimated from IPCA provide an excellent description of conditional expected stock returns. In the four-factor IPCA model, the estimated compensation for factor exposures ($\hat{\beta}'_{i,t}\hat{\lambda}$) delivers a predictive R^2 for returns of 1.8%. In the matched sample, the predictive R^2 from the Fama-French five-factor model is 0.3%. If we instead use standard PCA to estimate the latent four-factor specification, it delivers a 29.0% total R^2, but produces a *negative* predictive R^2 and thus has no explanatory power for differences in average returns across stocks. In summary, IPCA is the most successful model we analyze for jointly explaining realized variation in returns (i.e., systematic risks) and differences in average returns (i.e., risk compensation).

The model performance statistics cited above are based on in-sample estimation. If we instead use recursive out-of-sample estimation to calculate predictive R^2's for stock returns, we find that IPCA continues to outperform alternatives. The four-factor IPCA predictive R^2 is 0.7% per month out-of-sample, still more than doubling the Fama-French five-factor in-sample R^2.

By linking factor loadings to observable data, IPCA tremendously reduces the dimension of the parameter space compared to models with observable factors and even compared to standard PCA. To accommodate the more than 12,000 stocks in our sample, the Fama-French five-factor model requires estimation of 57,260 loading parameters. Four-factor PCA estimates 53,648 parameters including the time series of latent factors. Four-factor IPCA estimates only 2,688 (including each realization of the latent factors), or 95% fewer parameters than the pre-specified factor model or PCA, and incorporates dynamic loadings without relying on ad hoc rolling estimation. It does this by essentially redefining the identity of a stock in terms of its characteristics, rather than in terms of the stock identifier. Thus, once a stock's characteristics are known, only a small number of parameters (which are common to all assets) are required to map observed characteristic values into betas.

IPCA's success in explaining differences in average returns across stocks comes solely through its description of factor loadings—it restricts intercept coefficients to zero for all stocks. The question remains as to whether there are differences in average returns across stocks that align with characteristics and that are unexplained by exposures to IPCA factors.

By allowing intercepts to also depend on characteristics, IPCA provides a test for whether characteristics help explain expected returns above and beyond their role in factor loadings. Including alphas in the one-factor IPCA model improves its ability to explain average returns and rejects the null hypothesis of zero intercepts. Evidently, with a single factor, the specification of factor exposures is not rich enough to assimilate all of the return predictive content in stock characteristics. Thus, the excess predictability from characteristics spills into the intercept to an economically large and statistically significant extent.

However, when we consider specifications with $K \geq 2$, the improvement in model fit due to non-zero intercepts becomes small and statistically insignificant. The economic conclusion is that a risk structure with two or more dimensions coincides with information in stock characteristics in such a way that i) risk exposures are exceedingly well described by stock characteristics, and ii) the residual return predicability from characteristics, above and beyond that in factor loadings, falls to effectively zero, obviating the need to resort to "anomaly" intercepts.

The dual implication of IPCA's superior explanatory power for average stock returns is that IPCA factors are closer to being multivariate mean-variance efficient than factors in competing models. We show that the tangency portfolio of factors from the four-factor IPCA specification achieves an ex ante (i.e., out-of-sample) Sharpe ratio of 2.6, versus 1.3 for the five Fama-French factors. We also demonstrate that the performance of IPCA is very similar when we restrict the sample to large stocks and when modeling annual rather than monthly returns.

Lastly, IPCA offers a test for which characteristics are significantly associated with factor loadings (and, in turn, expected returns) while controlling for all other characteristics, in analogy to t-tests of explanatory variable significance in a regression model. In our main specification, we find that eight of the 36 firm characteristics in our sample are statistically significant at the 5% level, and five of these are significant at the 1% level. These include essentially three types of variables: size (e.g., market capitalization and total book assets), recent stock trends (e.g., short-term reversal and momentum), and market beta. Notably absent from the list of significant characteristics are measures of value such as book-to-market and earnings-to-price. If we re-estimate the model using the subset of eight significant regressors, we find that the model fit is nearly identical to the full 36-characteristic specification.

The fact that only a small subset of characteristics is necessary to explain variation in realized and expected stock returns shows that most characteristics are statistically irrelevant for understanding the cross section of returns, once they are evaluated in an appropriate multivariate context. Furthermore, that we cannot reject the null of zero alphas using only two or more IPCA risk factors leads us to conclude that the few characteristics that enter significantly do so because they help explain assets' exposures to systematic risks, *not* because they represent anomalous compensation without risk.

1.3 Literature

Our works builds on several literatures studying the behavior of stock returns. Calling this literature large is a gross understatement. Rather than attempting a thorough review, we briefly describe three primary strands of literature most closely related to our analysis and highlight a few exemplary contributions in each.

One branch of this literature analyzes latent factor models for returns, beginning with Ross's (1976) seminal APT. Empirical contributions to this literature, such as Chamberlain and Rothschild (1983) and Connor and Korajcyzk (1986, 1988), rely on principal component analysis of returns. Our primary innovation relative to this literature is to bring the wealth of characteristic information into latent factor models and, in doing so, make it possible to tractably analyze latent factor models with dynamic loadings.

Another strand of literature models factor loadings as functions of observables. Most closely related are models in which factor exposures are functions of firm characteristics, dating at least to Rosenberg (1974). In contrast to our contributions, Rosenberg's analysis is primarily theoretical, assumes that factors are observable, and does not provide a testing framework.[6] Ferson and Harvey (1991) allow for dynamic betas as asset-specific functions of macroeconomic variables. They differ from our analysis by relying on observable factors and focusing on macro rather than firm-specific instrumental variables. Daniel and Titman (1996) directly compare stock characteristics to factor loadings in terms of their ability to explain differences in average returns, an approach recently extended by Chordia, Goyal, and Shanken (2015). IPCA is unique in nesting competing characteristic and beta models of returns while simultaneously estimating the latent factors that most accurately coincide with characteristics as loadings, rather than relying on pre-specified factors.

A third literature models stock returns as a function of many characteristics at once. This literature has emerged only recently in response to an accumulation of research on predictive stock characteristics, and exploits more recently developed statistical techniques for high-dimensional predictive systems. Lewellen (2015) analyzes the joint predictive power of up to 15 characteristics in OLS regression. Light, Maslov, and Rytchkov (2016) and Freyberger, Neuhierl, and Weber (2017) consider larger collections of predictors than Lewellen and address concomitant statistical challenges using partial least squares and LASSO, respectively. These papers take a pure return forecasting approach and do not attempt to develop an asset pricing model or conduct asset pricing tests.

[6] Also see, for example, the documentation for BARRA's (1998) risk management and factor construction models.

Kozak et al. (forthcoming) show that a small number of principal components from 15 anomaly portfolios (from Novy-Marx and Velikov, 2015) are able to price those same portfolios with insignificant alphas. Our results illustrate that while that model works well for pricing their specific anomaly portfolios, the same model (and the PCA approach more generally) fails with other sets of test assets. We show that IPCA solves this problem by allowing assets' loadings to transition smoothly as its characteristics evolve. As a result, IPCA successfully prices individual stocks as well as anomaly or other stock portfolios. In follow on work, Kozak et al. (2017) use shrinkage estimation to select a subset of characteristic portfolios with good out-of-sample explanatory power for average returns. They directly model risk prices as functions of average portfolio returns then use shrinkage to identify a model with non-mechanical out-of-sample explanatory power for returns. Our findings differ in a few ways. First, our IPCA method selects pricing factors based on a factor variance criterion, then subsequently and separately tests whether loadings on these factors explain differences in average returns. Second, our approach derives formal tests and emphasizes statistical model comparison in a frequentist setting. Our tests i) differentiate whether a characteristic is better interpreted as a proxy for systematic risk exposure or as an anomaly alpha, ii) assess the incremental explanatory power of an individual characteristic against a (potentially high dimension) set of competing characteristics, and iii) compare latent factors against pre-specified alternative factors. The results of these tests conclude that stock characteristics are best interpreted as risk loadings, that most of the characteristics proposed in the literature contain no incremental explanatory power for returns, and that commonly studied pre-specified factors are inefficient in a mean-variance sense.

We describe the IPCA model and our estimation approach in Section 2. Section 3 develops asset pricing and model comparison tests in the IPCA setting. Section 4 reports our empirical findings and Section 5 concludes.

2 Model and Estimation

The general IPCA model specification for an excess return $r_{i,t+1}$ is

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} f_{t+1} + \epsilon_{i,t+1}, \tag{3}$$

$$\alpha_{i,t} = z'_{i,t}\Gamma_\alpha + \nu_{\alpha,i,t}, \quad \beta_{i,t} = z'_{i,t}\Gamma_\beta + \nu_{\beta,i,t}.$$

The system is comprised of N assets over T periods. The model allows for dynamic factor loadings, $\beta_{i,t}$, on a K-vector of latent factors, f_{t+1}. Loadings potentially depend on observable asset characteristics contained in the $L \times 1$ instrument vector $z_{i,t}$ (which includes a constant).

The specification of $\beta_{i,t}$ is central to our analysis and plays two roles. First, instrumenting the estimation of latent factor loadings with observable characteristics allows additional data to shape the factor model for returns. This differs from traditional latent factor techniques like PCA that estimate the factor structure solely from returns data. Anchoring the loadings to observable instruments can make the estimation more efficient and thereby improve model performance. This is true even if the instruments and true loadings are constant over time (see Fan, Liao, and Wang, 2016). Second, incorporating time-varying instruments makes it possible to estimate dynamic factor loadings, which is valuable when one seeks a model of conditional return behavior.

The matrix Γ_β defines the mapping between a potentially large number of characteristics and a small number of risk factor exposures. Estimation of Γ_β amounts to finding a few linear combinations of candidate characteristics that best describe the latent factor loading structure.[7] Our model emphasizes dimension reduction of the characteristic space. If there are many characteristics that provide noisy but informative signals about a stock's risk exposures, then aggregating characteristics into linear combinations isolates the signal and averages out the noise.[8] Any behavior of dynamic loadings that is orthogonal to the instruments falls into $\nu_{\beta,i,t}$. With this term, the model recognizes that firms' risk exposures are not perfectly recoverable from observable firm characteristics.

The Γ_β matrix also allows us to confront the challenge of migrating assets. Stocks evolve over time, moving for example from small to large, growth to value, high to low investment intensity, and so forth. Received wisdom in the asset pricing literature is that stock expected returns evolve along with these characteristics. But the very fact that the "identity" of the stock changes over time makes it difficult to model stock-level conditional expected returns using simple time series methods. The standard response to this problem is to dynamically form portfolios that hold the average characteristic value within the portfolio approximately constant. But if an adequate description of an asset's identity requires several characteristics, this portfolio approach becomes infeasible due to the proliferation of portfolios. IPCA

[7] The model imposes that $\beta_{i,t}$ is linear in instruments. Yet it accommodates non-linear associations between characteristics and exposures by allowing instruments to be non-linear transformations of raw characteristics. For example, one might consider including the first, second, and third power of a characteristic into the instrument vector to capture nonlinearity via a third-order Taylor expansion, or interactions between characteristics. Relatedly, $z_{i,t}$ can include asset-specific, time-invariant instruments.

[8] Dimension reduction among characteristics is a key differentiator between IPCA and BARRA's approach to factor modeling.

provides a natural and general solution: Parameterize betas as a function of the characteristics that determine a stock's expected return. In doing so, migration in the asset's identity is tracked through its betas, which are themselves defined by their characteristics in a way that is consistent among all stocks (Γ_β is a global mapping shared by all stocks). Thus, IPCA avoids the need for a researcher to perform an ad hoc preliminary dimension reduction that gathers test assets into portfolios. Instead, the model accommodates a high-dimensional system of assets (individual stocks) by estimating a dimension reduction that represents the identity of a stock in terms of its characteristics.[9]

We examine the null hypothesis that characteristics do *not* proxy for alpha: this corresponds to restricting Γ_α to zero in (3). The unrestricted version of (3) allows for non-zero Γ_α, representing the alternative hypothesis that conditional expected returns have intercepts that depend on stock characteristics. The structure of $\alpha_{i,t}$ is a linear combination of instruments mirroring the specification of $\beta_{i,t}$. IPCA estimates $\alpha_{i,t}$ by finding the linear combination of characteristics (with weights given by Γ_α) that best describes conditional expected returns while controlling for the role of characteristics in factor risk exposure. If characteristics align with average stock returns *differently* than they align with risk factor loadings, then IPCA will estimate a non-zero Γ_α, thus conceding anomalous compensation for holding stocks in excess of their systematic risk.

We focus on models in which the number of factors, K, is small, imposing a view that the empirical content of an asset pricing model is its parsimony in describing sources of systematic risk. At the same time, we allow the number of instruments, L, to be potentially large, as literally hundreds of characteristics have been put forward by the literature to explain average stock returns. And, because any individual characteristic is likely to be a noisy representation of true factor exposures, accommodating large L allows the model to average over characteristics in a way that reduces noise and more accurately reveals true exposures.

2.1 Restricted Model ($\Gamma_\alpha = 0$)

In this section we provide a conceptual overview of IPCA estimation. Our description here introduces two identifying assumptions and discusses their role in estimation. Kelly, Pruitt,

[9] This structure also makes it easy to calculate a firm's cost of capital as a function of observable characteristics. This avoids reliance on CAPM betas or other factor loadings that may be difficult to estimate with time series regression, or may rely on a badly misspecified model. Once Γ_β is recovered from a representative set of asset returns, it can be used to price other assets that may lack a long history of returns but at least have a snapshot of recent firm characteristics available.

and Su (2017) derive the IPCA estimator and prove that, together with the identifying assumptions, IPCA consistently estimates model parameters and latent factors as the number of assets and the time dimension simultaneously grow large, as long as factors and residuals satisfy weak regularity conditions (their Assumptions 2 and 3). We refer interested readers to that paper for technical details and present a practical summary here.

We first describe estimation of the restricted model in which $\Gamma_\alpha = \mathbf{0}_{L\times 1}$, ruling out the possibility that characteristics capture "anomalous" compensation without risk. This restriction maintains that characteristics explain expected returns only insofar as they proxy for systematic risk exposures. In this case, equation (3) becomes

$$r_{i,t+1} = z_{i,t}'\Gamma_\beta f_{t+1} + \epsilon^*_{i,t+1} \tag{4}$$

where $\epsilon^*_{i,t+1} = \epsilon_{i,t+1} + \nu_{\alpha,i,t} + \nu_{\beta,i,t} f_{t+1}$ is a composite error.[10]

We derive the estimator using the vector form of equation (4),

$$r_{t+1} = Z_t\Gamma_\beta f_{t+1} + \epsilon^*_{t+1},$$

where r_{t+1} is an $N\times 1$ vector of individual firm returns, Z_t is the $N\times L$ matrix that stacks the characteristics of each firm, and ϵ^*_{t+1} likewise stacks individual firm residuals. Our estimation objective is to minimize the sum of squared composite model errors:

$$\min_{\Gamma_\beta, F} \sum_{t=1}^{T-1} \left(r_{t+1} - Z_t\Gamma_\beta f_{t+1}\right)' \left(r_{t+1} - Z_t\Gamma_\beta f_{t+1}\right). \tag{5}$$

The values of f_{t+1} and Γ_β that minimize (5) satisfy the first-order conditions

$$\hat{f}_{t+1} = \left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' r_{t+1}, \quad \forall t \tag{6}$$

and

$$\text{vec}(\Gamma_\beta') = \left(\sum_{t=1}^{T} \left[Z_t \otimes f_t'\right]' \left[Z_t \otimes f_t'\right]\right)^{-1} \left(\sum_{t=1}^{T} \left[Z_t \otimes f_t'\right]' r_t\right). \tag{7}$$

Condition (6) shows that factor realizations are period-by-period cross section regression coefficients of r_{t+1} on the latent loading matrix β_t. Likewise, Γ_β, is the coefficient of returns

[10] That is, our data generating process has two sources of noise that affect estimation of factors and loadings. The first comes from returns being determined in large part by idiosyncratic firm-level shocks ($\epsilon_{i,t+1}$) and the second from the fact that characteristics do not perfectly reveal the true factor model parameters ($\nu_{\alpha,i,t}$ and $\nu_{\beta,i,t}$). There is no need to distinguish between these components of the residual in the remainder of our analysis.

regressed on the factors interacted with firm-specific characteristics. This system of first-order conditions has no closed-form solution and must be solved numerically. Fortunately, the numerical problem is solvable via alternating least squares in a matter of seconds even for high dimension systems. We describe our numerical method in Appendix A and discuss a special case that admits an exact analytical solution in Appendix C.2.

As common in latent factor models, IPCA requires one further assumption for estimator identification. Γ_β and f_{t+1} are unidentified in the sense that any set of solutions can be rotated into an equivalent solution $\Gamma_\beta R^{-1}$ and $R f_{t+1}$ for a non-singular K-dimensional rotation matrix R. To resolve this indeterminacy, we impose that $\Gamma_\beta' \Gamma_\beta = \mathbb{I}_K$, that the unconditional covariance matrix of f_t is diagonal with descending diagonal entries, and that the mean of f_t is non-negative. These assumptions place no economic restrictions on the model and solely serve to pin down a uniquely identified solution to the first-order conditions.[11]

2.1.1 A Managed Portfolio Interpretation of IPCA

To help develop intuition for IPCA factor and loading estimates, it is useful to compare with the estimator for a static factor model such as $r_t = \beta f_t + \epsilon_t$ (e.g., Connor and Korajczyk, 1988). The objective function in the static case is

$$\min_{\beta, F} \quad \sum_{t=1}^{T} \left(r_t - \beta f_{t+1}\right)' \left(r_t - \beta f_{t+1}\right)$$

and the first-order condition for f_{t+1} is $f_{t+1} = \left(\beta'\beta\right)^{-1} \beta' r_{t+1}$. Substituting this into the original objective yields a concentrated objective function for β of

$$\max_{\beta} \; \mathrm{tr}\left(\sum_t \left(\beta'\beta\right)^{-1} \beta' r_{t+1} r_{t+1}' \beta \right).$$

This objective maximizes a sum of so-called "Rayleigh quotients" that all have the same denominator, $\beta'\beta$. In this special case, the well-known PCA solution for β is given by the first K eigenvectors of the sample second moment matrix of returns, $\sum_t r_{t+1} r_{t+1}'$.

Naturally, dynamic betas complicate the optimization problem. Substituting the IPCA first-

[11] Our identification is a dynamic model counterpart to the standard PCA identification assumption for static models (see, for example, Assumption F1(a) in Stock and Watson, 2002a).

order condition (6) into the original objective yields

$$\max_{\Gamma_\beta} \quad \mathrm{tr}\left(\sum_{t=1}^{T-1}\left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1}\Gamma_\beta' Z_t' r_{t+1} r_{t+1}' Z_t \Gamma_\beta\right). \tag{8}$$

This concentrated IPCA objective is more challenging because the Rayleigh quotient denominators, $\Gamma_\beta' Z_t' Z_t \Gamma_\beta$, are different for each element of the sum and, as a result, there is no analogous eigenvector solution for Γ_β.

Nevertheless, the structure of our dynamic problem is closely reminiscent of the static problem. While the static PCA estimator applies the singular value decomposition to the panel of individual asset returns $r_{i,t}$, our derivation shows that the IPCA problem can be *approximately* solved by applying the singular value decomposition not to raw returns, but to returns interacted with instruments. Consider the $L \times 1$ vector defined as

$$x_{t+1} = Z_t' r_{t+1}. \tag{9}$$

This is the time $t+1$ realization of returns on a set of L managed portfolios. The l^{th} element of x_{t+1} is a weighted average of stock returns with weights determined by the value of l^{th} characteristic for each stock at time t. Stacking time series observations produces the $T \times L$ matrix $X = [x_1', ..., x_T']'$. Each column of X is a time series of returns on a characteristic-managed portfolio. If the first three characteristics are, say, size, value, and momentum, then the first three columns of X are time series of returns to portfolios managed on the basis of each of these.

If we were to approximate the Rayleigh quotient denominators with a constant (for example, replacing each $Z_t' Z_t$ by their time series average, $T^{-1}\sum_t Z_t' Z_t$), then the solution to (8) would be to set Γ_β equal to the first K eigenvectors of the sample second moment matrix of managed portfolio returns, $X'X = \sum_t x_t x_t'$.[12] Likewise, the estimates of f_{t+1} would be the first K principal components of the managed portfolio panel. This is a close approximation to the exact solution as long as $Z_t' Z_t$ is not too volatile. More importantly, because we desire to find the exact solution, this approximation provides an excellent starting guess to initialize the numerical optimization and quickly reach the exact optimum.

As one uses more and more characteristics to instrument for latent factor exposures, the number of characteristic-managed portfolios in X grows. Prior empirical work shows that there tends to be a high degree of common variation in anomaly portfolios (e.g. Kozak et al.,

[12] To be more precise, Γ_β is estimated by a rotation of x_t's second moment, where the rotation is a function of $T^{-1}\sum_t Z_t' Z_t$.

forthcoming). IPCA recognizes this and estimates factors and loadings by focusing on the common variation in X. It estimates factors as the K linear combinations of X's columns, or "portfolios of portfolios," that best explain covariation among the panel of managed portfolios. These factors differ from principal components of X in that the explained covariation is re-weighted across time and across assets to emphasize observations associated with the most informative instruments (this is the role of the $Z_t'Z_t$ denominator).

IPCA can also be viewed as a generalization of period-by-period cross section regressions as employed in Fama and MacBeth (1973) and Rosenberg (1974). When $K = L$ there is no dimension reduction—the estimates of f_{t+1} are the characteristic-managed portfolios themselves and are equal to the period-wise Fama-MacBeth regression coefficients (and $\Gamma_\beta = \mathbb{I}_L$). This unreduced specification is close in spirit to the BARRA model. But when $K < L$, IPCA's f_{t+1} estimate is a *constrained* Fama-MacBeth regression coefficient. The constrained regression not only estimates return loadings on lagged characteristics, but it must also choose a reduced-rank set of regressors—the $K < L$ combinations of characteristics that best fit the cross section regression.

The asset pricing literature has struggled with the question of which test assets are most appropriate for evaluating models (Lewellen, Nagel, and Shanken, 2010; Daniel and Titman, 2012). IPCA provides a resolution to this dilemma. On one hand, IPCA tests can be viewed as using the set of test assets with the finest possible resolution—the set of individual stocks. At the same time, the discussion regarding equation (8) shows that IPCA's tests can equivalently be viewed as using characteristic-managed portfolios, x_t, as the set of test assets, which have comparatively low dimension and average out a substantial degree of idiosyncratic stock risk.[13]

2.2 Unrestricted Model ($\Gamma_\alpha \neq \mathbf{0}$)

The unrestricted IPCA model allows for intercepts that are functions of the instruments, thereby admitting the possibility of "anomalies" in which expected returns depend on characteristics in a way that is not explained by exposure to systematic risk. Like the factor specification in (4), the unrestricted IPCA model assumes that intercepts are a linear com-

[13] A further convenience of the managed portfolio representation is that avoids issues with missing observations in stock-level data. In particular, when constructing managed portfolios in equation (9), we evaluate this inner product as a sum over elements of Z_t and r_{t+1} for which both terms in the cross-product are non-missing. Thus (9) is a slight abuse of notation.

bination of instruments with weights defined by the $L \times 1$ parameter vector Γ_α:

$$r_{i,t+1} = z_{i,t}'\Gamma_\alpha + z_{i,t}'\Gamma_\beta f_{t+1} + \epsilon^*_{i,t+1}. \tag{10}$$

Model (10) is unrestricted in that it admits mean returns that are not determined by factor exposures alone.

Estimation proceeds nearly identically to Section 2.1. We rewrite (10) as $r_{i,t+1} = z_{i,t}'\tilde{\Gamma}\tilde{f}_{t+1} + \epsilon^*_{i,t+1}$, where $\tilde{\Gamma} \equiv [\Gamma_\alpha, \Gamma_\beta]$ and $\tilde{f}_{t+1} \equiv [1, f_{t+1}']'$. That is, the unrestricted model is mapped to the structure of (4) by simply augmenting the factor specification to include a constant.

The first-order condition for $\tilde{\Gamma}$ is the same as (7) except that $\tilde{f}_t$ replaces f_t. The f_{t+1} first-order condition changes slightly to

$$f_{t+1} = \left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' \left(r_{t+1} - Z_t\Gamma_\alpha\right), \quad \forall t. \tag{11}$$

This is a cross section regression of "returns in excess of alpha" on dynamic betas, and reflects the fact that the unrestricted estimator decides how to best allocate panel variation in returns to factor exposures versus anomaly intercepts.

3 Asset Pricing Tests

In this section we develop three hypothesis tests that are central to our empirical analysis. The first is designed to test the zero alpha condition that distinguishes the restricted and unrestricted IPCA models of Sections 2.1 and 2.2. The second tests whether observable factors (such as the Fama-French five-factor model) significantly improve the model's description of the panel of asset returns while controlling for IPCA factors. The third tests the incremental significance of an individual characteristic or set of characteristics while simultaneously controlling for all other characteristics.

3.1 Testing $\Gamma_\alpha = \mathbf{0}_{L\times 1}$

When a characteristic lines up with expected returns in the cross section, the unrestricted IPCA estimator in Section 2.2 decides how to split that association. Does the characteristic proxy for exposure to common risk factors? If so, IPCA will attribute the characteristic to beta via $\hat{\beta}_{i,t} = z_{i,t}'\hat{\Gamma}_\beta$, thus interpreting the characteristic/expected return relationship as compensation for bearing systematic risk. Or, does the characteristic capture anomalous differences in average returns that are unassociated with systematic risk? In this case, IPCA will be unable to find common factors for which characteristics serve as loadings, so it will attribute the characteristic to alpha via $\hat{\alpha}_{i,t} = z'_{i,t}\hat{\Gamma}_\alpha$.

In the restricted model, the association between characteristics and alphas is disallowed. If the data truly call for an anomaly alpha, then the restricted model is misspecified and will produce a poor fit compared to the unrestricted model that allows for alpha. The distance between unrestricted alpha estimates and zero summarizes the improvement in model fit from loosening the alpha restriction. If this distance is statistically large (i.e., relative to sampling variation), we can conclude that the true alphas are non-zero.

We propose a test of the zero alpha restriction that formalizes this logic. In the model equation

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} f_{t+1} + \epsilon_{i,t+1},$$

we are interested in testing the null hypothesis

$$H_0 : \Gamma_\alpha = \mathbf{0}_{L\times 1}$$

against the alternative hypothesis

$$H_1 : \Gamma_\alpha \neq \mathbf{0}_{L\times 1}.$$

Characteristics determine alphas in this model only if Γ_α is non-zero. The null therefore states that alphas are unassociated with characteristics in $z_{i,t}$. Because the hypothesis is formulated in terms of the common parameter, this is a joint statement about alphas of all assets in the system.

Note that $\Gamma_\alpha = \mathbf{0}_{L\times 1}$ does not rule out the existence of alphas entirely. From the model definition in equation (3), we see that $\alpha_{i,t}$ may differ from zero because $\nu_{\alpha,i,t}$ is non-zero. That is, the null allows for some mispricing, as long as mispricings are truly idiosyncratic and unassociated with characteristics in the instrument vector. Likewise, the alternative hypothesis is not concerned with alphas arising from the idiosyncratic $\nu_{\alpha,i,t}$ mispricings. Instead, it focuses on the more economically interesting mispricings that may arise as a regular function of observable characteristics.

In statistical terms, $\Gamma_\alpha \neq \mathbf{0}_{L\times 1}$ is a constrained alternative. This contrasts, for example, with the Gibbons, Ross, and Shanken (1989, GRS henceforth) test that studies the unconstrained alternative $\alpha_i = 0 \; \forall i$. In GRS, each α_i is estimated as an intercept in a time series regression. GRS alphas are therefore residuals, not a model. Our constrained alternative is itself a model that links stock characteristics to anomaly expected returns via a fixed mapping that is common to all firms. If we reject the null IPCA model, we do so in favor of a specific model for how alphas relate to characteristics. In this sense our asset pricing test is a frequentist counterpart to Barillas and Shanken's (forthcoming) Bayesian argument that it should take a model to beat a model. This has the pedagogical advantage that, if we reject H_0 in favor of H_1, we can further determine which elements of Γ_α (and thus which characteristics) are most responsible for the rejection. By isolating those characteristics that are a wedge between expected stock returns and exposures to aggregate risk factors, we can work toward an economic understanding of how the wedge emerges.[14]

We construct a Wald-type test statistic for the distance between the restricted and unrestricted models as the sum of squared elements in the estimated Γ_α vector,

$$W_\alpha = \hat{\Gamma}'_\alpha \hat{\Gamma}_\alpha.$$

Inference, which we conduct via bootstrap, proceeds in the following steps. First, we estimate the unrestricted model and retain the estimated parameters

$$\hat{\Gamma}_\alpha, \hat{\Gamma}_\beta, \text{ and } \{\hat{f}_t\}_{t=1}^T.$$

A convenient aspect of our model from a bootstrapping standpoint is that, because the objective function can be written in terms of managed portfolios x_t, we can resample portfolio residuals rather than resampling stock-level residuals. Following from the managed portfolio definition,

$$x_{t+1} = Z'_t r_{t+1} = (Z'_t Z_t)\Gamma_\alpha + (Z'_t Z_t)\Gamma_\beta f_{t+1} + Z'_t \epsilon^*_{i,t+1},$$

we define the $L \times 1$ vector of managed portfolio residuals as $d_{t+1} = Z'_t \epsilon^*_{i,t+1}$ and retain their fitted values $\{\hat{d}_t\}_{t=1}^T$.

[14]Our test also overcomes a difficult technical problem that the GRS test faces in large cross sections such as that of individual stocks. It is very difficult to test the hypothesis that all stock-level alphas are zero as it amounts to a joint test of N parameters where N potentially numbers in the tens of thousands. The IPCA specification reduces the joint alpha test to a $L \times 1$ parameter vector. Contrast this with recent advances in alphas tests such as Fan, Liao, and Wang (2016) who illustrate the complications of the large N testing problem and resort to a complicated thresholded covariance matrix estimator. Even with that advance, they look at only the stocks in the S&P 500, which is an order of magnitude smaller than the cross section we consider.

Next, for $b = 1, ..., 1000$, we generate the b^{th} bootstrap sample of returns as

$$\tilde{x}_t^b = (Z_t' Z_t)\hat{\Gamma}_\beta \hat{f}_t + \tilde{d}_t^b, \quad \tilde{d}_t^b = q_1^b \hat{d}_{q_2^b}. \tag{12}$$

The variable q_2^b is a random time index drawn uniformly from the set of all possible dates. In addition, we multiply each residual draw by a Student t random variable, q_1^b, that has unit variance and five degrees of freedom. Then, using this bootstrap sample, we re-estimate the unrestricted model and record the estimated test statistic $\tilde{W}_\alpha^b = \tilde{\Gamma}_\alpha^{b\prime}\tilde{\Gamma}_\alpha^b$. Finally, we draw inferences from the empirical null distribution by calculating a p-value as the fraction of bootstrapped $\tilde{W}_\alpha^b$ statistics that exceed the value of W_α from the actual data.

3.1.1 Comments on Bootstrap Procedure

The method described above is a "residual" bootstrap. It uses the model's structure to generate pseudo-samples under the null hypothesis that $\Gamma_\alpha = 0$. In particular, it fixes the explained variation in returns at their estimated common factor values under the null model, $Z_t\hat{\Gamma}_\beta\hat{f}_{t+1}$, and randomizes around the null model by sampling from the empirical distribution of residuals to preserve their properties in the simulated data. The bootstrap dataset x_t^b satisfies $\Gamma_\alpha = 0$ by construction because a non-zero $\hat{\Gamma}_\alpha$ is estimated as part of the unrestricted model but excluded from the bootstrap data. This approach produces an empirical distribution of $\tilde{W}_\alpha^b$ designed to quantify the amount of sampling variation in the test statistic under the null. In Appendix B, we report a variety of Monte Carlo experiments illustrating the accuracy of the test in terms of size (appropriate rejection rates under the null) and power (appropriate rejection rates under the alternative).

Premultiplying the residual draws by a random t variable is a technique known as the "wild" bootstrap. It is designed to improve the efficiency of bootstrap inference in heteroskedastic data such as stock returns (Goncalves and Killian, 2004). Appendix B also demonstrates the improvement in test performance, particularly power for nearby alternatives, from using a wild bootstrap.

Equation (8) illustrates why we bootstrap datasets of managed portfolios returns, x_t, rather than raw stock returns, r_t. The estimation objective ultimately takes as data the managed portfolio returns, $x_t = Z_t' r_t$, and estimates parameters from their covariance matrix. If we were to resample stock returns, the estimation procedure would anyway convert these into managed portfolios before estimating model parameters. It is thus more convenient to resample x_t directly. Bootstrapping managed portfolio returns comes with a number of practical advantages. It resamples in a lower dimension setting ($T \times L$) than stock returns ($T \times N$), which reduces computation cost. It also avoids issues with missing observations that exist in the stock panel, but not in the portfolio panel.

Our test enjoys the usual benefits of bootstrapping, such as reliability in finite samples and validity under weak assumptions on residual distributions. It is important to point out that our bootstrap tests are feasible *only* because of the fast alternating least squares estimator that we have derived for IPCA. Estimation of the model via brute force numerical optimization would not only make it very costly to use IPCA in large systems—it would immediately take bootstrapping off the table as a viable testing approach.

3.2 Testing Observable Factor Models Versus IPCA

Next, we extend the IPCA framework to nest commonly studied models with pre-specified, observable factors. The encompassing model is

$$r_{i,t+1} = \beta_{i,t} f_{t+1} + \delta_{i,t} g_{t+1} + \epsilon_{i,t+1}. \tag{13}$$

The $\beta_{i,t} f_{t+1}$ term is unchanged from its specification in (3). The new term is the portion of returns described by the $M \times 1$ vector of observable factors, g_{t+1}. Loadings on observable factors are allowed to be dynamic functions of the same conditioning information entering into the loadings on latent IPCA factors:

$$\delta_{i,t} = z'_{i,t} \Gamma_\delta + \nu_{\delta,i,t},$$

where Γ_δ is the $L \times M$ mapping from characteristics to loadings. The encompassing model imposes the zero alpha restriction so that we can evaluate the ability of competing models to price assets based on exposures to systematic risk (though it easy to incorporate $\alpha_{i,t}$ in in (13) if desired).

We estimate model (13) following the same general tack as Section 2.2. We rewrite (13) as $r_{i,t+1} = z'_{i,t} \tilde{\Gamma} \tilde{f}_{t+1} + \epsilon^*_{i,t+1}$, where $\tilde{\Gamma} \equiv [\Gamma_\beta, \Gamma_\delta]$ and $\tilde{f}_{t+1} \equiv [f'_{t+1}, g'_{t+1}]'$. That is, the model with nested observable factors is mapped to the structure of (4) by augmenting the factor specification to include the observable g_{t+1}. The first-order condition for $\tilde{\Gamma}$ the same as (7) except that $\tilde{f}_t$ replaces f_t. The f_{t+1} first-order condition changes slightly to

$$f_{t+1} = \left(\Gamma'_\beta Z'_t Z_t \Gamma_\beta\right)^{-1} \Gamma'_\beta Z'_t \left(r_{t+1} - Z_t \Gamma_\delta g_{t+1}\right), \quad \forall t.$$

This is a cross section regression of “returns in excess of observable factor exposures” on β_t, and reflects the fact that that nested specification decides how to best allocate panel variation in returns to latent IPCA factors versus observable pre-specified factors.

We construct a test of the incremental explanatory power of observable factors after controlling for the baseline IPCA specification. The hypotheses for this test are

$$H_0 : \Gamma_\delta = \mathbf{0}_{L\times M} \quad \text{vs.} \quad H_1 : \Gamma_\delta \neq \mathbf{0}_{L\times M}.$$

$\hat{\Gamma}_\delta$ denotes the estimated parameters corresponding to g_{t+1}, from which we construct the Wald-like test statistic

$$W_\delta = \text{vec}(\hat{\Gamma}_\delta)'\text{vec}(\hat{\Gamma}_\delta).$$

W_δ is a measure of the distance between model (13) and the IPCA model that excludes observable factors ($\Gamma_\delta = \mathbf{0}_{L\times M}$). If W_δ is large relative to sampling variation, we can conclude that g_{t+1} holds explanatory power for the panel of returns above and beyond the baseline IPCA factors.

Our sampling variation estimates, and thus p-values for W_δ, use the same residual wild bootstrap concept from Section 3.1. First, we construct residuals of managed portfolios, $\hat{d}_{t+1} = Z_t'\hat{\epsilon}^*_{i,t+1}$, from the estimated model. Then, for each iteration b, we resample portfolio returns imposing the null hypothesis. Next, from each bootstrap sample, we re-estimate Γ_δ and construct the associated test statistic $\tilde{W}^b_\delta$. Finally, we compute the p-value as the fraction of samples for which $\tilde{W}^b_\delta$ exceeds W_δ.

3.3 Testing Instrument Significance

The last test that we introduce evaluates the significance of an individual characteristic while simultaneously controlling for all other characteristics. We focus on the model in equation (4) with alpha fixed at zero and no observable factors. That is, we specifically investigate whether a given instrument significantly contributes to $\beta_{i,t}$.

To formulate the hypotheses, we partition the parameter matrix as

$$\Gamma_\beta = \left[\gamma_{\beta,1}\ ,\ \ldots\ ,\gamma_{\beta,L}\right]',$$

where $\gamma_{\beta,l}$ is a $K \times 1$ vector that maps characteristic l to loadings on the K factors. Let the

l^{th} element of $z_{i,t}$ be the characteristic in question. The hypotheses that we test are

$$H_0 : \Gamma_\beta = [\gamma_{\beta,1}\ ,\ \dots\ ,\gamma_{\beta,l-1}\ ,\ \mathbf{0}_{K\times 1}\ ,\gamma_{\beta,l+1}\ ,\ \dots\ ,\gamma_{\beta,L}]' \quad \text{vs.} \quad H_1 : \Gamma_\beta = [\gamma_{\beta,1}\ ,\ \dots\ ,\gamma_{\beta,L}]'.$$

The form of this null hypothesis comes from the fact that, for the l^{th} characteristic to have zero contribution to the model, it cannot impact *any* of the K factor loadings. Thus, the entire l^{th} row of Γ_β must be zero.

We estimate the alternative model that allows for a non-zero contribution from characteristic l, then we assess whether the distance between zero and the estimate of vector $\gamma_{\beta,l}$ is statistically large. Our Wald-type statistic in this case is

$$W_{\beta,l} = \hat{\gamma}'_{\beta,l}\hat{\gamma}_{\beta,l}.$$

Inference for this test is based on the same residual bootstrap concept described above. We define the estimated model parameters and managed portfolio residuals from the alternative model as

$$\{\hat{\gamma}_{\beta,l}\}_{l=1}^{L}, \{\hat{f}_t\}_{t=1}^{T}, \text{ and } \{\hat{d}_t\}_{t=1}^{T}.$$

Next, for $b = 1, ..., 1000$, we generate the b^{th} bootstrap sample of returns under the null hypothesis that the l^{th} characteristic has no effect on loadings. To do so, we construct the matrix

$$\tilde{\Gamma}_\beta = [\hat{\gamma}_{\beta,1}\ ,\ \dots\ ,\hat{\gamma}_{\beta,l-1}\ ,\ \mathbf{0}_{K\times 1}\ ,\hat{\gamma}_{\beta,l+1}\ ,\ \dots\ ,\hat{\gamma}_{\beta,L}]$$

and re-sample characteristic-managed portfolio returns as

$$\tilde{x}_t^b = Z_t\tilde{\Gamma}_\beta\hat{f}_t + \tilde{d}_t^b$$

with the same formulation of $\tilde{d}_t^b$ used in equation (12). Then, for each sample b, we re-estimate the alternative model and record the estimated test statistic $\tilde{W}^b_{\beta,l}$. Finally, calculate the test's p-value as the fraction of bootstrapped $\tilde{W}^b_{\beta,l}$ statistics that exceed $W_{\beta,l}$. This test can be extended to evaluate the joint significance of multiple characteristics $l_1, ..., l_J$ by modifying the test statistic to $W_{\beta,l_1,...,l_J} = \hat{\gamma}'_{\beta,l_1}\hat{\gamma}_{\beta,l_1} + ... + \hat{\gamma}'_{\beta,l_J}\hat{\gamma}_{\beta,l_J}$.

4 Empirical Findings

4.1 Data

Our stock returns and characteristics data are from Freyberger, Neuhierl, and Weber (2017). The sample begins in July 1962, ends in May 2014, and includes 12,813 firms. For each firm we have 36 characteristics. They are market beta (`beta`), assets-to-market (`a2me`), total assets (`log_at`), sales-to-assets (`ato`), book-to-market (`beme`), cash-to-short-term-investment (`c`), capital turnover (`cto`), capital intensity (`d2a`), ratio of change in PP&E to change in total assets (`dpi2a`), earnings-to-price (`e2p`), fixed costs-to-sales (`fc2y`), cash flow-to-book (`free_cf`), idiosyncratic volatility (`idio_vol`), investment (`investment`), leverage (`lev`), log lagged market equity (`size`), lagged turnover (`lturnover`), net operating assets (`noa`), operating accruals (`oa`), operating leverage (`ol`), price-to-cost margin (`pcm`), profit margin (`pm`), gross profitability (`prof`), Tobin's Q (`q`), closeness to relative high price (`rel_high`), return on net operating assets (`rna`), return on assets (`roa`), return on equity (`roe`), momentum (`mom_12_2`), intermediate momentum (`mom_12_7`), short-term reversal (`mom_2_1`), long-term reversal (`mom_36_13`), sales-to-price (`s2p`), SG&A-to-sales (`sga2s`), bid-ask spread (`spread`), and unexplained volume (`suv`). We restrict attention to i, t observations for which all 36 characteristics are non-missing. For further details and summary statistics, see Freyberger, Neuhierl, and Weber (2017).

These characteristics vary both in the cross section and over time. The two dimensions potentially aid IPCA's estimation of the factor model in different ways. For example, the average level of a stock characteristic may be helpful for understanding a stock's unconditional factor loadings, while time variation around this mean may help understand the stock's conditional loadings, and the relevance of the two components for asset pricing may differ in magnitude. To allow for this possibility, we separate characteristics into their time series mean and their deviation around the mean. We denote the vector of characteristics on stock i at time t as $c_{i,t}$. The vector of IPCA instruments includes a constant, as well as means and deviations of each characteristic:

$$z_{i,t} = [1,\ \bar{c}_i',\ (c_{i,t} - \bar{c}_i)']', \quad \text{where} \quad \bar{c}_i \equiv \frac{1}{T} \sum_{t=1}^{T} c_{i,t}.$$

When we perform out-of-sample analyses, we replace the full sample mean $\bar{c}_i$ with the historical mean $\bar{c}_{i,t} \equiv \frac{1}{t} \sum_{\tau=1}^{t} c_{i,\tau}$. We cross-sectionally standardize instruments period-by-period. In particular, we calculate stocks' ranks for each characteristic, then divide ranks by the number of non-missing observations and subtract 0.5. This maps characteristics into the [-0.5,+0.5] interval and focuses on their ordering as opposed to magnitude. We use this standardization for its insensitivity to outliers, then show in robustness analyses that results are qualitatively the same without characteristic standardization.

4.2 The Asset Pricing Performance of IPCA

We estimate the K-factor IPCA model for various choices of K, and consider both restricted ($\Gamma_\alpha = \mathbf{0}$) and unrestricted versions of each specification. Two R^2 statistics measure model performance. The first we refer to as the "total R^2" and define it as

$$\text{Total } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - z'_{i,t}(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1})\right)^2}{\sum_{i,t} r^2_{i,t+1}}. \tag{14}$$

It represents the fraction of return variance explained by both the dynamic behavior of conditional loadings (and alphas in the unrestricted model), as well as by the contemporaneous factor realizations, aggregated over all assets and all time periods. The total R^2 summarizes how well the systematic factor risk in a given model specification describes the realized riskiness in the panel of individual stocks.

The second measure we refer to as the "predictive R^2" and define it as

$$\text{Predictive } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - z'_{i,t}(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{\lambda})\right)^2}{\sum_{i,t} r^2_{i,t+1}}. \tag{15}$$

It represents the fraction of realized return variation explained by the model's description of conditional expected returns. IPCA's return predictions are based on dynamics in factor loadings (and alphas in the unrestricted model). In theory, expected returns can also vary because risk prices vary. One limitation of IPCA is that, without further model structure, it cannot separately identify risk price dynamics. Hence, we hold estimated risk prices constant and predictive information enters return forecasts only through the instrumented loadings. When $\Gamma_\alpha = \mathbf{0}$ is imposed, the predictive R^2 summarizes the model's ability to describe risk compensation solely through exposure to systematic risk. For the unrestricted model, the predictive R^2 describes how well characteristics explain expected returns in any form—be it through loadings or through anomaly intercepts.

Panel A of Table I reports R^2's at the individual stock level for $K = 1, ..., 6$ factors. With

Table I
IPCA Model Performance

Note. Panel A and B report total and predictive R^2 in percent for the restricted ($\Gamma_\alpha = \mathbf{0}$) and unrestricted ($\Gamma_\alpha \neq \mathbf{0}$) IPCA model. These are calculated with respect to either individual stocks (Panel A) or characteristic-managed portfolios (Panel B). Panel C reports bootstrapped p-values for the test of $\Gamma_\alpha = \mathbf{0}$.

		K					
		1	2	3	4	5	6
		Panel A: Individual Stocks (r_t)					
Total R^2	$\Gamma_\alpha = \mathbf{0}$	15.1	17.0	18.5	19.4	19.9	20.3
	$\Gamma_\alpha \neq \mathbf{0}$	16.6	18.2	19.2	19.8	20.2	20.4
Pred. R^2	$\Gamma_\alpha = \mathbf{0}$	0.44	1.64	1.70	1.83	1.82	1.82
	$\Gamma_\alpha \neq \mathbf{0}$	1.89	1.87	1.87	1.86	1.86	1.85
		Panel B: Managed Portfolios (x_t)					
Total R^2	$\Gamma_\alpha = \mathbf{0}$	91.6	93.6	96.7	97.7	98.3	98.7
	$\Gamma_\alpha \neq \mathbf{0}$	92.6	96.2	97.6	98.2	98.6	98.8
Pred. R^2	$\Gamma_\alpha = \mathbf{0}$	2.06	2.84	2.75	2.97	2.98	2.97
	$\Gamma_\alpha \neq \mathbf{0}$	3.14	3.11	3.10	3.11	3.08	3.07
		Panel C: Asset Pricing Test					
W_α p-value		0.00	15.4	32.9	38.2	59.4	13.0

a single factor, the restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA model explains 15.1% of the total variation in stock returns. As a reference point, the total R^2 from the CAPM and Fama-French three-factor model is 11.9% and 18.9%, respectively, in our individual stock sample.

The predictive R^2 in the restricted one-factor IPCA model is 0.4%. This is for individual stocks and at the monthly frequency. To benchmark this magnitude, the predictive R^2 from the CAPM or the Fama-French three-factor model is 0.3% in a matched individual stock sample.

Allowing for $\Gamma_\alpha \neq \mathbf{0}$ increases the total R^2 by 1.5 percentage points to 16.6%, while the predictive R^2 rises dramatically to 1.9% per month for individual stocks. The unrestricted IPCA specification attributes predictive content from characteristics to either betas or alphas. The results show that with $K = 1$, IPCA can only capture about one-fifth of the return predictability embodied by characteristics while maintaining the $\Gamma_\alpha = \mathbf{0}$ constraint. This represents a failure of the restricted one-factor IPCA model to explain heterogeneity in conditional expected returns. This failure is statistically borne out by the hypothesis test of $\Gamma_\alpha = \mathbf{0}$ in Panel C, which rejects the null with a p-value below 0.01%.

When we allow for multiple IPCA factors, the gap between restricted and unrestricted models shrinks rapidly. At $K = 2$, the total R^2 for the restricted model is 17.0%, achieving more than 93% of the explanatory power of the unrestricted model. The predictive R^2 rises to 1.6%, capturing 88% of the characteristics' predictive content while imposing $\Gamma_\alpha = \mathbf{0}$. Our test fails to reject the null hypothesis that $\Gamma_\alpha = \mathbf{0}$ when $K = 2$ (p-value of 15.4%). For $K > 2$, the distance between the restricted and unrestricted models shrinks even further. At $K = 4$, the restricted and unrestricted models behave nearly identically.

The results of Table I show that IPCA explains essentially all of the heterogeneity in average stock returns associated with stock characteristics if at least two factors are included in the specification. It does so by identifying a set of factors and associated loadings such that stocks' expected returns align with their exposures to systematic risk—without resorting to alphas to explain the predictive role of characteristics. In other words, IPCA infers that characteristics are risk exposures, not anomalies.

Note that, because IPCA is estimated from a least squares criterion, it directly targets total R^2. Thus the risk factors that IPCA identifies are optimized to describe systematic risks among stocks. They are by no means specialized to explain average returns, however, as estimation does not directly target the predictive R^2. Because conditional expected returns are a small portion of total return variation (as evidenced by the 1.9% predictive R^2 in the unrestricted model), it is very well possible that a misspecified model could provide an excellent description of risk yet a poor description of risk compensation (in results below, the traditional PCA model serves as an example of this phenomenon). Evidently, this is not the case for IPCA, as its risk factors indirectly produce an accurate description of risk compensation across assets.

The asset pricing literature is accustomed to evaluating the performance of pricing factors in explaining the behavior of test portfolios, such as the 5×5 size and value-sorted portfolios of Fama and French (1993), as opposed to individual stocks. The behavior of portfolios can differ markedly from individual stocks because they average out a large fraction of idiosyncratic variation. As emphasized in Section 2.1.1, IPCA asset pricing tests can be at once interpreted as tests of stocks, or as tests of characteristic-managed portfolios, x_t. In this spirit, Panel B of Table I evaluates fit measures for managed portfolios.[15] The x_t vector includes returns on 73 portfolios, one correspond to each instrument (the means and time

[15]Fit measures for x_t are

$$\text{Total } R^2 = 1 - \frac{\sum_t \left(x_{t+1} - Z_t'Z_t(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1})\right)' \left(x_{t+1} - Z_t'Z_t(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1})\right)}{\sum_t x_{t+1}' x_{t+1}},$$

series deviations of 36 characteristics plus a constant).

With $K = 4$ factors, the total R^2's for the restricted and unrestricted models are 96.7% and 97.6%, respectively. The reduction in noise via portfolio formation also improves predictive R^2's to 3.0% and 3.1% for the restricted and unrestricted models, respectively. As in the stock-level case, the unrestricted IPCA specification performs almost identically to the unrestricted specification when $K \geq 2$.

4.3 Comparison with Existing Models

The results in Table I compare the performance of IPCA across specification choices for K and with or without the imposition of asset pricing restrictions. We now compare IPCA to leading alternative modeling approaches in the literature. The first includes models with pre-specified observable factors. We consider models with $K = 1, 3, 4, 5$, or 6 observable factors. The $K = 1$ model is the CAPM (using the CRSP value-weighted excess market return as the factor), $K = 3$ is the Fama-French (1993) three-factor model that includes the market, SMB and HML ("FF3" henceforth). The $K = 4$ model is the Carhart (1997, "FFC4") model that adds MOM to the FF3 model. $K = 5$ is the Fama-French (2015, "FF5") five-factor model that adds RMW and CMA to the FF3 factors. Finally, we consider a six-factor model ("FFC6") that includes MOM alongside the FF5 factors.

We report two implementations of observable factor models. The first is a traditional approach in which factor loadings are estimated asset-by-asset via time series regression. In this case, loadings are static and no characteristic instruments are used in the model. The second implementation places observable factor models on the same footing as our IPCA latent factor model by parameterizing loadings as a function of instruments, following the definition of $\delta_{i,t}$ in equation (13).

The last set of alternatives that we consider are static latent factor models estimated with PCA. In this approach, we consider one to six principal component factors from the panel of individual stock returns.[16]

and

$$\text{Predictive } R^2 = 1 - \frac{\sum_t \left(x_{t+1} - Z_t' Z_t(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{\lambda})\right)' \left(x_{t+1} - Z_t' Z_t(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{\lambda})\right)}{\sum_t x_{t+1}' x_{t+1}}.$$

[16] In calculating PCA, we must confront the fact that the panel of returns is unbalanced. We estimate PCA using the alternating least squares option in Matlab's `pca.m` function. As a practical matter, this means that PCA estimation for individual stocks bears a high computational cost. It also highlights the computational benefit of IPCA, which side steps the unbalanced panel problem by parameterizing betas with characteristics

Table II
IPCA Comparison With Other Factor Models

Note. The table reports total and predictive R^2 in percent and number of estimated parameters (N_p) for the restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA model (Panel A), for observable factor models with static loadings (Panel B), for observable factor models with instrumented dynamic loadings (Panel C), and for static latent factor models (Panel D). Observable factor model specifications are CAPM, FF3, FFC4, FF5, and FFC6 in the $K = 1, 3, 4, 5, 6$ columns, respectively.

Test Assets	Statistic	K				
		1	3	4	5	6
		Panel A: IPCA				
r_t	Total R^2	15.3	18.6	19.5	20.0	20.4
	Pred. R^2	0.46	1.64	1.77	1.76	1.76
	N_p	672	2016	2688	3360	4032
x_t	Total R^2	91.6	96.7	97.7	98.3	98.7
	Pred. R^2	2.06	2.75	2.97	2.98	2.97
	N_p	672	2016	2688	3360	4032
		Panel B: Observable Factors (no instruments)				
r_t	Total R^2	11.9	18.9	20.9	21.9	23.7
	Pred. R^2	0.31	0.29	0.28	0.29	0.23
	N_p	11452	34356	45808	57260	68712
x_t	Total R^2	34.3	51.0	56.4	55.3	60.4
	Pred. R^2	0.97	1.63	1.48	1.92	1.70
	N_p	73	219	292	365	438
		Panel C: Observable Factors (with instruments)				
r_t	Total R^2	10.7	14.7	15.7	15.3	16.2
	Pred. R^2	0.35	0.57	0.54	0.71	0.66
	N_p	73	219	292	365	438
x_t	Total R^2	65.8	84.4	86.5	85.7	87.7
	Pred. R^2	1.70	2.10	2.02	2.16	2.08
	N_p	73	219	292	365	438
		Panel D: Principal Components				
r_t	Total R^2	16.8	26.2	29.0	31.5	33.8
	Pred. R^2	< 0	< 0	< 0	< 0	< 0
	N_p	13412	40236	53648	67060	80472
x_t	Total R^2	87.9	93.9	95.3	96.4	97.1
	Pred. R^2	2.00	2.04	2.59	2.60	2.60
	N_p	672	2016	2688	3360	4032

and, as a result, is estimated from managed portfolios that can always be constructed to have no missing data.

We estimate all models in Table II restricting intercepts to zero. We do so by imposing $\Gamma_\alpha = \mathbf{0}$ in specifications relying on instrumented loadings, and by omitting a constant in the static (time series regression) implementation of observable factor models.

Table II reports the total and predictive R^2 as well as the number of estimated parameters (N_p) for each model.[17] For ease of comparison, Panel A re-states model fits for IPCA with $\Gamma_\alpha = \mathbf{0}$.[18]

Panel B reports fits for static (i.e., excluding instruments) observable factor models. In the analysis of individual stocks, observable factor models generally produce a slightly higher total R^2 than the IPCA specification using the same number of factors. For example, at $K = 5$, FF5 achieves a 1.9 percentage point improvement in R^2 relative to IPCA's fit of 20.0%. To accomplish this, however, observable factors rely on vastly more parameters than IPCA. The number of parameters in an observable factor model is equal to the number of loadings, or $N_p = NK$. For IPCA, the number of factors is the dimension of Γ_β plus the number of estimated factor realizations, or $N_p = LK + TK$. In our sample of 11,452 stocks with 72 instruments over 599 months, observable factor models therefore estimate 17 times ($\approx 11452/(72 + 599)$) as many parameters as IPCA. In short, IPCA provides a similar description of systematic risk in stock returns as leading observable factors while using almost 85% fewer parameters.

At the same time, IPCA provides a substantially more accurate description of stocks' risk compensation than observable factor models, as evidenced by the predictive R^2. Observable factor models' predictive power never rises beyond 0.3% for any specification, compared with a 1.8% predictive R^2 from IPCA with $K = 4$.

Among characteristic-managed portfolios x_t, the explanatory power from observable factor models' total R^2 suffers in comparison to IPCA. For $K \geq 3$, IPCA explains over 96% of total portfolio return variation, while FFC6 explains 60.4%. IPCA also achieves at least a 55%

[17] The R^2's for alternative models are defined analogously to those of IPCA. In particular, for individual stocks they are

$$\text{Total } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t} - \hat{\beta}_i \hat{f}_t\right)^2}{\sum_{i,t} r_{i,t}^2}, \text{Predictive } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t} - \hat{\beta}_i \hat{\lambda}\right)^2}{\sum_{i,t} r_{i,t}^2},$$

and are similarly adapted for managed portfolios x_t. In all cases, model fits are based on exactly matched samples with IPCA.

[18] These differ minutely from Table I results because, for the sake of comparability in this table, we drop stock-month observations with insufficient data for time series regression on observable factors. We require at least 60 non-missing months to compute observable factor betas, which filters out slightly more than a thousand stocks compared to the sample used in Table I, and leaves us with just under 1.4 million stock-month observations.

improvement in predictive R^2 (1.9% at best among observable factor models, versus 3.0% for IPCA). In sum, when test assets are managed portfolios, IPCA dominates in its ability to describe systematic risks as well as cross-sectional differences in average returns.

Panel C investigates the performance of observable factor models when loadings are allowed to vary over time in the same manner as IPCA. In this case, loadings are instrumented with the same characteristics used for the IPCA analysis, following the definition of $\delta_{i,t}$ in equation (13). Because the intercept and latent factor components in (13) are restricted to zero, the the estimation of Γ_δ reduces to a panel regression of $r_{i,t+1}$ on $z_{i,t} \otimes g_{t+1}$. For individual stocks, allowing for dynamic loadings decreases the total R^2 compared to the static model, but roughly doubles the predictive R^2 for $K \geq 2$. At the portfolio level, the dynamic loadings sharply increase total R^2, while at the same time improving the predictive R^2. This is an interesting result in its own right and, to the best of our knowledge, is new to the literature.

In Panel D we report the performance of models latent factors with static and uninstrumented loadings (estimated via PCA). At the individual stock level, PCA's total R^2 exceeds that of IPCA and observable factor models for all K. PCA, however, provides a dismal description of expected returns at the stock level; the predictive R^2 is negative for each K. On the other hand, when the model is re-estimated using data at the managed portfolio level, PCA fits in terms of both total and predictive R^2 are generally excellent and only exceeded by IPCA.

Table III formally tests whether the inclusion of observable factors improves over a given IPCA specification in the matched individual stock sample. The tests, described in Section 3.2, nest the various sets of observable factors studied in Table II (represented by rows) with different numbers of latent IPCA factors (represented by columns). Panels A and B show total and predictive R^2's for these joint models. For ease of comparison, we restate fits from the baseline IPCA specification in the rows showing zero observable factors. When $K = 1$, adding observable factors improves the model fit. The total R^2 rises from 15.1% for the IPCA-only model to 18.3% with the FFC6 factors. The predictive R^2 rises from 0.4% to 1.1%. Hypothesis test results show that RMW and CMA offer a statistically significant improvement over the $K = 1$ IPCA model (with p-values below 1%).

With more IPCA factors, however, observable factors become redundant. At $K = 2$, none of the FFC6 factors are statistically significant after controlling for IPCA factors. The predictive R^2 rises only 0.2% from 1.6% with IPCA alone to 1.8% with the addition of FFC6 factors, and the total R^2 rises from 17.0% to 19.4%. With three or more IPCA factors, the incremental explanatory power from observable factors is negligible.

Table III
IPCA Fits Including Observable Factors

Note. Panels A and B report total and predictive R^2 from IPCA specifications with various numbers of latent factors K (corresponding to columns) while also controlling for observable factors according to equation (13). Rows labeled 0, 1, 4, and 6 correspond to no observable factors or the CAPM, FFC4, or FFC6 factors, respectively. Panel C reports tests of the incremental explanatory power of each observable factor model with respect to the IPCA model.

Observ.	K					
Factors	1	2	3	4	5	6
			Panel A: Total R^2			
0	15.1	17.0	18.5	19.4	19.9	20.3
1	16.2	17.9	18.8	19.5	20.0	20.4
4	18.0	19.2	19.6	20.0	20.3	20.6
6	18.3	19.4	19.8	20.1	20.4	20.7
			Panel B: Predictive R^2			
0	0.44	1.64	1.70	1.83	1.82	1.82
1	0.43	1.71	1.83	1.83	1.82	1.82
4	0.91	1.81	1.82	1.82	1.83	1.81
6	1.08	1.82	1.82	1.83	1.81	1.81
			Panel C: Individual Significance Test p-value			
MKT–RF	29.1	33.8	94.4	79.1	71.0	61.8
SMB	29.8	50.2	81.5	64.3	52.4	61.0
HML	20.1	3.80	6.30	72.7	89.1	82.6
RMW	0.40	22.1	50.9	46.2	82.8	95.0
CMA	0.40	29.5	16.4	25.0	18.3	23.4
MOM	12.7	8.10	9.90	28.2	67.1	63.7

4.3.1 Discussion

Table II offers a synthesis of IPCA vis-à-vis existing cross-sectional pricing models. The spectrum of factor models can be classified on two dimensions. First, are factors latent or observable? Second, are loadings static or parameterized functions of dynamic instruments? Table II clearly differentiates the empirical role of each ingredient.

The empirical performance of traditional cross section pricing models is hampered in both model dimensions. First, they rely on pre-specified observable factors that are not directly optimized for describing asset price variation. Second, they rely on static loadings estimated via time series regression. The superior performance of IPCA shows that, by allowing the data to dictate factors that best describe common sources of risk, and by freeing loadings to vary through time as a function of observables, model fits are unambiguously improved.

Which of these dimensions is more important? Panel C shows that dynamic betas—particularly ones that are parameterized functions of observable stock-characteristics—are responsible for large improvements in predictive R^2. By tying loadings to characteristics, the no-arbitrage connection between loading and expected returns is substantially enhanced, even when factors are pre-specified outside the model. At the same time, this pulls down the total R^2. This is perhaps unsurprising, given the massive e xpansion in parameter count when one moves to regression-based beta estimates for each asset. It is likely that higher total R^2 in static models is merely an artifact of statistical overfit from over-parameterization.

Panel D shows that it is possible for a latent factor model to succeed in describing returns even when betas are static. However, this conclusion depends crucially on the choice of test assets—are they individual stocks or managed portfolios? The glowing performance of static IPCA among managed portfolios is closely in line with the findings of Kozak et al. (forthcoming). Yet at the stock level, static PCA leads to egregious model performance. This divergence suggests an inherent misspecification in the static latent factor model.

In contrast, IPCA is successful in describing returns for both individual stocks and for managed portfolios—and it does so using the exact same set of model parameters for both sets of assets. In short, the incorporation of both latent factors and dynamic betas are the key model enhancements that allow IPCA to achieve a unique level of success in describing the cross section of returns.

4.4 Out-of-sample Fits

Thus far, the performance of IPCA and alternatives has been based on in-sample estimates. That is, IPCA factors and loadings are estimated from the full panel of stock returns. Next, we analyze IPCA's out-of-sample fits.

To construct out-of-sample fit measures, we use recursive backward-looking estimation and track the post-estimation performance of estimated models. In particular, in every month $t \geq T/2$, we use all data through t to estimate the IPCA model and denote the resulting backward-looking parameter estimate as $\hat{\Gamma}_{\beta,t}$. Then, based on equation (6), we calculate the out-of-sample realized factor return at $t+1$ as $\hat{f}_{t+1,t} = \left(\hat{\Gamma}'_{\beta,t} Z'_t Z_t \hat{\Gamma}_{\beta,t}\right)^{-1} \hat{\Gamma}'_{\beta,t} Z'_t r_{t+1}$. That is, IPCA factor returns at $t+1$ may be calculated with individual stock weights $\left(\hat{\Gamma}'_{\beta,t} Z'_t Z_t \hat{\Gamma}_{\beta,t}\right)^{-1} \hat{\Gamma}'_{\beta,t} Z'_t$ that require no information beyond time t, just like the portfolio sorts used to construct observable factors.

The out-of-sample total R^2 compares r_{t+1} to $Z_t \hat{\Gamma}_{\beta,t} \hat{f}_{t+1,t}$ and x_{t+1} to $Z'_t Z_t \hat{\Gamma}_{\beta,t} \hat{f}_{t+1,t}$. The out-

Table IV
Out-of-sample Fits

Note. The table reports out-of-sample total and predictive R^2 in percent with recursive estimation scheme.

Test Assets	Statistic	K 1	2	3	4	5	6
r_t	Total R^2	14.6	16.0	16.9	17.6	18.1	18.5
	Pred. R^2	0.32	0.31	0.52	0.57	0.56	0.56
x_t	Total R^2	88.7	93.7	95.7	97.0	98.0	98.3
	Pred. R^2	2.44	2.39	2.66	2.88	2.83	2.83

of-sample predictive R^2 is defined analogously, replacing $\hat{f}_{t+1,t}$ with the factor mean through t, denoted $\hat{\lambda}_t$.

Table IV reports out-of-sample R^2 statistics. The main conclusion from the table is that the strong performance of IPCA is not merely an in-sample phenomenon driven by statistical overfit. IPCA delivers nearly the same out-of-sample total R^2 that it achieves in-sample. And while the predictive R^2 is somewhat reduced, it remains economically large and outperforms even the in-sample predictions from observable factors, both for individual stocks and managed portfolios.

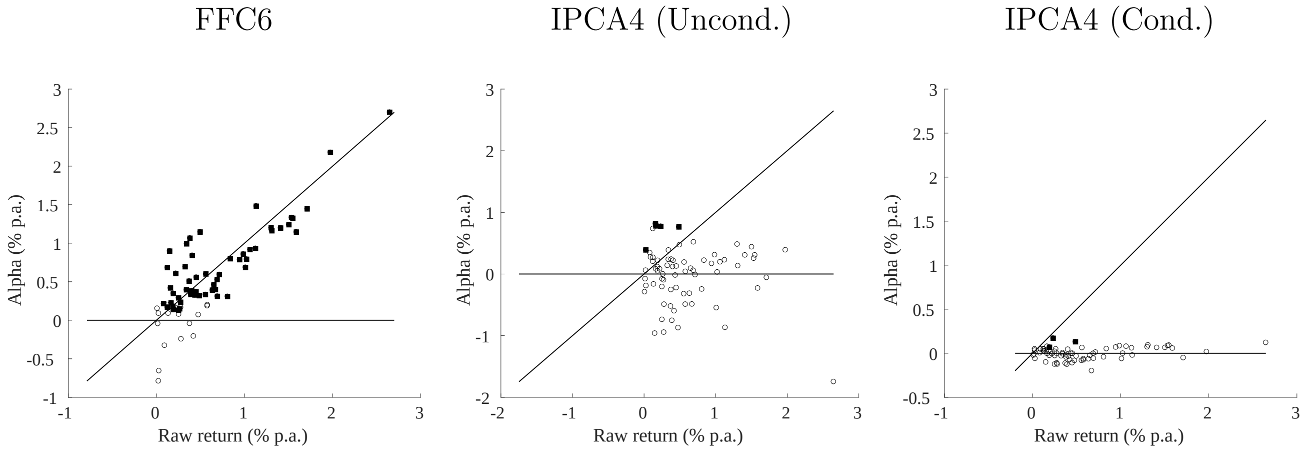
4.5 Unconditional Mean-Variance Efficiency

Zero intercepts in a factor pricing model are equivalent to multivariate mean-variance efficiency of the factors. This fact is the basis for the Gibbons, Ross, and Shanken (1989) alpha test and bears a close association with our W_α test. The evidence thus far indicates that IPCA loadings predict asset returns more accurately than competing factor models both in-sample and out-of-sample, suggesting that IPCA factors achieve higher multivariate efficiency than competitors. In this section, we directly investigate factor efficiency.

An important, if subtle, fact to bear in mind is that ours is a conditional asset pricing model—factor loadings are parameterized functions of conditioning instruments. A conditional asset pricing model is attractive because it often maps directly to decisions of investors, who seek period-by-period to hold assets with high conditional expected returns relative to their conditional risk. At the same time, conditional models are more econometrically challenging than unconditional models because they typically require the researcher to estimate a dynamic model and take a stand on investors' conditioning information. These complica-

Figure 1: Alphas of Characteristic-Managed Portfolios

Note. The left and middle panels report unconditional alphas for characteristic-managed portfolios (x_t), relative to FFC6 factors and four IPCA factors, respectively, estimated from time series regression. The right panel reports the time series averages of conditional alphas in the baseline four-factor IPCA model. Alphas are plotted against portfolios' raw average excess returns. Alphas with t-statistics in excess of 2.0 are shown with filled squares, while insignificant alphas are shown with unfilled circles.



tions lead the factor model literature to focus predominantly on unconditional estimation and testing.[19]

Our findings of small and insignificant Γ_α suggests that estimated IPCA factors are multivariate mean-variance efficient, *conditionally.* This result does not necessarily imply *unconditional* efficiency of our factors, and therefore our results are not directly comparable to the existing analysis that revolves around unconditional models. To bridge this gap, we conduct two sets of analyses that assess the unconditional efficiency of IPCA factors and thereby link our analysis with the large empirical literature on unconditional factor models.

4.5.1 Does IPCA "Price" Anomalies Unconditionally?

First, we investigate whether IPCA factors and observable factors accurately "price" anomaly portfolios unconditionally. To do so, we estimate unconditional alphas in a full-sample time series regression of portfolio returns onto each set of factors.

The anomaly portfolios that we study are the 73 characteristic-managed portfolios, x_t, discussed in Section 4.2. For comparability, we re-sign portfolios to have positive means and scale them to 10% annualized volatility. The mean and median annualized Sharpe ratios

[19]For an excellent treatment of conditional versus unconditional asset pricing models and mean-variance efficiency, see chapter 8 of Cochrane (2005).

among these portfolios are 0.50 and 0.74, respectively, with 17 portfolios having a Sharpe ratio in excess of 1.50.[20]

Figure 1 reports unconditional alpha estimates for each portfolio. The left-hand figure plots portfolios' anomaly alphas from the FFC6 model against their raw average excess returns, overlaying the 45-degree line. Alphas with t-statistics in excess of 2.0 are depicted with filled squares, while insignificant alphas are shown with unfilled circles. With only a few exceptions, characteristic-managed portfolios appear anomalous with respect to the FFC6 model. Their alphas are mostly statistically significant and clustered around the 45-degree line, indicating that their average returns are essentially unexplained by observable factors.

The middle figure shows unconditional alphas for the same portfolios with respect to the four-factor IPCA model.[21] Specifically, we first estimate the four IPCA factors from the baseline conditional specification, then we regress portfolios on these estimated factors in full sample time series regressions (i.e., with static betas) to recover unconditional alphas. In this case, alphas are clustered around the zero line. Only four of the 73 alphas are significantly greater than zero, and even these are small in magnitude compared to the distribution of average anomaly returns. Lastly, for comparison, the right-hand figure reports time series averages of conditional alphas from the four-factor IPCA specification (i.e, averages of period-by-period residuals from the main conditional IPCA model). Figure 1 supports the conclusion that, not only are IPCA factors close to conditionally multivariate mean-variance efficient, they appear to be unconditionally efficient as well.

4.5.2 Factor Tangency Portfolios

Second, we analyze out-of-sample unconditional Sharpe ratios for IPCA factors.[22] We report univariate annualized Sharpe ratios to describe unconditional efficiency of individual factors, and we report the ex ante unconditional tangency portfolio Sharpe ratio for a group of factors to describe multivariate efficiency. We calculate out-of-sample factor returns following the

[20]The performance of x_t portfolios is broadly comparable to anomaly portfolios studied in the literature. For example, within the dataset of 41 anomaly portfolios posted by Novy-Marx in conjunction with Novy-Marx and Velikov (2015), the mean and median Sharpe ratios during the same time period are 0.47 and 0.57, respectively, with one having a Sharpe ratio over 1.50.

[21]We report the ($\Gamma_\alpha = \mathbf{0}$) $K = 4$ specification rather than including plots for all specifications due to space constraints. This analysis helps assess IPCA factor efficiency, and $K = 4$ is the point at which both the in-sample and out-of-sample predictive R^2 peak for the $\Gamma_\alpha = \mathbf{0}$ model. Other values of $K \geq 1$ produce similar results.

[22]Kozak, Nagel, and Santosh (forthcoming) emphasize that higher-order principal components of "anomaly" portfolios tend to suffer from in-sample overfit and generate unreasonably high in-sample Sharpe ratios. In light of this, we focus our analysis on out-of-sample IPCA factor returns.

Table V
Out-of-Sample Factor Portfolio Sharpe Ratios

Note. The table reports out-of-sample annualized Sharpe ratios for individual factors ("univariate") and for the mean-variance efficient portfolio of factors in each model ("tangency").

	K					
	1	2	3	4	5	6
			Panel A: IPCA			
Univariate	0.48	0.17	0.69	1.25	1.31	0.79
Tangency	0.48	0.50	2.22	2.55	3.54	3.60
			Panel B: Observable Factors			
Univariate	0.46	0.33	0.41	0.46	0.62	0.51
Tangency	0.46	0.51	0.78	1.01	1.29	1.37

same recursive estimation approach from Section 4.3. The tangency portfolio return for a set of factors is also constructed on a purely out-of-sample basis by using the mean and covariance matrix of estimated factors through t and tracking the post-formation $t+1$ return.[23]

Out-of-sample IPCA Sharpe ratios are shown in Panel A of Table V. The K^{th} column reports the univariate Sharpe ratio for factor K as well as the tangency Sharpe ratio based on factors 1 through K. For comparison, we report Sharpe ratios of observable factor models in Panel B.[24] The first IPCA factor produces a Sharpe ratio of 0.48, versus 0.46 for the market over the same out-of-sample period.[25] The fourth IPCA factor has an individual out-of-sample Sharpe ratio of 1.25, and boosts the Sharpe ratio for the four-factor tangency portfolio to 2.55. Adding additional factors increases the tangency Sharpe ratio further, reaching as high as 3.60 for $K = 6$. The out-of-sample Sharpe ratios of IPCA factors exceed those of observable factor models such as the FFC6 model, which itself reaches an impressive tangency Sharpe ratio of 1.37.

The high unconditional Sharpe ratio statistics in Panel A of Table V provide a succinct summary of the fact that IPCA captures extensive comovement among assets while success-

[23] We scale the tangency weights each period by targeting 1% monthly portfolio volatility based on historical estimates.

[24] A difference between IPCA and observable factors is that observable factors are pre-constructed on an out-of-sample basis. We construct observable factor tangency portfolios using historical mean and covariance estimates, following the same approach as for IPCA.

[25] The second half of our sample, which corresponds to our out-of-sample evaluation period, was an especially good period for the market in terms of Sharpe ratio. In the full post-1964 sample, the market Sharpe ratio is 0.37.

Table VI
IPCA Performance for Large versus Small Stocks

Note. Panel A and B report in-sample and out-of-sample total and predictive R^2 for subsamples of large and small stocks. We evaluate fits within each subsample using the same parameters (estimated from the unified sample of all stocks). All estimates use the restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA specification.

		K					
		1	2	3	4	5	6
		Panel A: Large Stocks					
In-Sample	Total R^2	24.1	25.9	28.7	30.0	30.8	31.6
	Pred. R^2	0.88	1.43	1.59	1.70	1.72	1.71
Out-of-Sample	Total R^2	22.4	25.3	26.4	27.1	28.1	28.8
	Pred. R^2	0.84	0.65	0.71	0.70	0.59	0.57
		Panel B: Small Stocks					
In-Sample	Total R^2	12.8	14.6	15.9	16.6	17.1	17.3
	Pred. R^2	0.33	1.69	1.73	1.86	1.85	1.85
Out-of-Sample	Total R^2	12.0	13.1	14.0	14.6	15.0	15.2
	Pred. R^2	0.21	0.23	0.52	0.58	0.59	0.60

fully aligning their factor loadings with differences in average returns. These results are not a statement about implementability of the factor tangency portfolio as a trading strategy. The analysis of Table V is designed to describe the mean-variance efficiency of IPCA factors irrespective of practical frictions such as trading costs, and in doing so is consistent with prior literature on testing factor models, all of which test models with returns gross of transaction costs. From a trading perspective, the tangency portfolio that we estimate has high turnover, implying high implementation costs. This is unsurprising given that, as we show in Section 4.8, the IPCA model is driven to a significant extent by fast-moving characteristics like momentum and short-term reversal. Table V raises an interesting question for follow-on research: How can IPCA be used for developing practical strategies to exploit the mean-variance tradeoff that it identifies? And, relatedly, how may IPCA be adapted to incorporate net-of-costs returns into its notion of factor efficiency?

4.6 Large Versus Small Stocks

Returns of large and small stocks tend to exhibit different behavior in terms of their covariances, liquidity, and expected returns. In particular, the fact that small stocks are much more volatile than large stocks raises a question of whether the high explained variation

Table VII
IPCA Cross-Validation for Large Versus Small Stocks

Note. The table reports total and predictive R^2 for large and small stock subsamples using parameters estimated separately in each subsample. Rows correspond to the sample from which parameters are estimated and columns represent the sample in which fits are evaluated. In particular, when row and column labels differ, we are using fits in one sample (e.g., small stocks) to cross-validate the reliability of parameters estimated in the other sample (e.g., large stocks). All estimates use the four-factor restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA specification.

	Fit Sample			
Estimation	Total R^2		Predictive R^2	
Sample	Large	Small	Large	Small
Large	32.1	11.4	1.72	1.55
Small	24.3	16.9	1.49	1.97

from the IPCA model occurs through an especially good description of small stock behavior at the expense of large stocks. To better understand the role of large and small stocks in IPCA fits, Table VI breaks out model R^2's for each group. These results are *not* based on separate model re-estimation for the two groups, which would mechanically allow IPCA to fit both subsamples but with potentially different parameters. Instead, these fits hold the model parameters fixed at their estimates from the unified sample, and we recalculate corresponding R^2's among each subsample.

We define the "large" group as the 1,000 stocks with the highest market capitalization each month, and "small" as all remaining stocks. Overall, the performance of IPCA is broadly similar for large and small stocks, and similar to our earlier results for the unified sample. If anything, IPCA offers an especially accurate description of large stock variation, with total R^2's exceeding 25% both in-sample and out-of-sample.

Next, to investigate the stability of model estimates for the two groups, we re-estimate the $K = 4$ IPCA model for large and small stocks separately. When IPCA is estimated using large stocks alone, the fits are strikingly similar to those from the unified sample. The 30.0% total R^2 among large stocks from Table VI increases slightly to 32.1% when the model is re-optimized on the large stock subsample, and the 1.7% predictive R^2 from the unified estimation is unchanged when estimated from large stocks only. The same pattern holds for small stocks. The total R^2 from unified sample estimates and small-only estimates are 16.6% and 16.9%, respectively, and the predictive R^2's are 1.9% and 2.0%.

Perhaps more impressively, when we cross-validate the fits by computing R^2's in one group using f_t and Γ_β parameters estimated from the *other* group, fits suffer only mildly. When

Table VIII
Out-of-sample Tangency Sharpe Ratios, Large Versus Small Stocks

Note. The table repeats the analysis of Table V for large and small stocks using parameters estimated separately in each subsample.

	K					
	1	2	3	4	5	6
Large	0.53	0.43	1.24	2.64	2.95	2.87
Small	0.43	0.56	1.85	2.05	2.90	3.23

we use large stock parameter estimates to fit small stocks, the small stock total R^2 drops to 11.4%, and the predictive R^2 drops to 1.6%. Likewise, when small stock parameters are used to fit large stocks, the total R^2 is 24.3% and the predictive R^2 is 1.5%. These cross-validated R^2's represent a further out-of-sample evaluation of the IPCA model: None of the information in small stock data enters directly into the large stock estimates, and vice versa (there are only indirect spillovers due to cross-correlation between large and small stocks). Like our other out-of-sample tests, the large/small sample split demonstrates remarkable stability of IPCA model fits.

Finally, Table VIII reports Sharpe ratios for out-of-sample tangency portfolios constructed from IPCA factors estimated separately from the large and small stock subsamples. Large stock Sharpe ratios remain high, though are somewhat muted compared to the unified sample. In both size categories, the overall Sharpe ratio patterns are quantitatively similar to estimates from the unified sample, suggesting that small stocks are not the sole drivers of baseline IPCA tangency results in Table V.

4.7 Annual Returns

Our analysis has thus far focused on the one-month return horizon. As an extension and robustness assessment, we re-analyze the IPCA model using annual stock returns. For many investors, the annual frequency represents a more realistic investment horizon. And, because our empirical analysis works with data at the individual stock level, it is possible that monthly return patterns identified by IPCA in part capture short lived fluctuations among illiquid individual stocks, and these effects are less influential at the annual frequency.

The basic structure of this analysis is unchanged from above with the exception that the left-hand-side return is aggregated over months $t+1$ through $t+12$. Time t characteristics now describe conditional loadings on annual factor returns from month $t+1$ through $t+12$.

Table IX
Annual Returns

Note. The table repeats the analysis of Table I using annual rather than monthly returns.

Test Assets	Statistic	K = 1	2	3	4	5	6
		Panel A: In-sample					
r_t	Total R^2	21.3	25.5	27.5	28.4	29.0	29.4
	Pred. R^2	9.20	11.0	11.0	10.9	10.9	10.9
x_t	Total R^2	83.6	92.9	96.5	97.9	98.6	98.9
	Pred. R^2	17.3	20.7	20.8	20.9	20.9	20.8
		Panel B: Out-of-sample					
r_t	Total R^2	15.1	17.6	18.5	19.2	19.7	20.0
	Pred. R^2	2.85	3.16	3.44	3.33	3.14	3.09
x_t	Total R^2	89.8	95.1	97.1	98.2	98.7	99.0
	Pred. R^2	18.6	19.3	20.0	20.2	18.3	18.4
		Panel C: Asset Pricing Test					
	W_α p-value	0.60	70.3	49.4	86.9	77.4	47.2

Panel A of IX reports the in-sample fits from re-estimating the restricted ($\Gamma_\alpha = \mathbf{0}$) model with annual returns. For individual stocks and $K = 4$, the total R^2 rises from 19.5% monthly to 28.4% with annual returns, and the predictive R^2 rises from 1.8% to 10.9%, respectively. Among characteristic-managed portfolios, the total R^2 rises from 97.7% to 97.9% and the predictive R^2 rises from 3.0% to 20.9%. These improvements are only partially due to re-estimating the model. Even if we hold parameter values fixed from the monthly data estimates of Table I, the total R^2 is 17.9% and 64.2% for r_t and x_t, respectively, and the predictive R^2's are 8.2% and 19.9%, respectively.

Panel B reports out-of-sample fits from the annual model. Out-of-sample fits somewhat attenuate at the individual stock level. The out-of-sample total R^2 is 19.2% and the predictive R^2 is 3.3%. In contrast, portfolio-level R^2's are essentially unchanged from their in-sample values. Panel C assesses improvement in fit for the ($\Gamma_\alpha \neq \mathbf{0}$) via the W_α statistic of Section 3.1. We cannot reject the null hypothesis that $\Gamma_\alpha = \mathbf{0}$ as long as $K > 1$.[26]

The primary conclusion from Table IX is that IPCA continues to provide an excellent de-

[26] Our analysis of annual returns uses overlapping monthly data. This overlap is accounted for via block bootstrap in our p-value calculation.

Table X
Individual Characteristic Contribution

Note. The table reports the contribution of each individual characteristic to overall model fit, defined as the reduction in total R^2 from setting all Γ_β elements pertaining to that characteristic to zero (in the restricted IPCA specification with $K = 4$). $**$ and $*$ denote that a variable significantly improves the model at the 1% and 5% levels, respectively.

size	2.18	**	a2me	0.13		e2p	0.03
mom_12_2	1.71	**	s2p	0.12		c	0.03
mom_1_0	0.83	**	pcm	0.10		d2a	0.02
mom_12_7	0.69	*	fc2y	0.08		dpi2a	0.02
beta	0.65	**	roe	0.08		q	0.02
rel_high	0.61	**	roa	0.08		free_cf	0.02
ol	0.49		sga2m	0.07		rna	0.01
at	0.44	*	suv	0.07		investment	0.01
cto	0.39		mom_36_13	0.06	*	prof	0.01
idio_vol	0.21		pm	0.05		lev	0.01
lturnover	0.16		beme	0.03		oa	0.01
spread_mean	0.14		ato	0.03		noa	0.01

scription of risk and compensation among annual returns. Its conditional factor loadings successfully explain cross-sectional differences in expected returns with alphas that are small and insignificant. This suggests that our monthly findings are unlikely to be dominated by illiquidity or other sources of very short-lived predictability.

4.8 Which Characteristics Matter?

Our statistical framework allows us to address questions about the incremental contribution of characteristics to help address Cochrane (2011)'s quotation in our preface. We test the statistical significance of an individual characteristic while simultaneously controlling for all other characteristics. Each characteristic enters the beta specification through two rows of the Γ_β matrix: one row corresponding to the average level of the characteristic and the other to deviations around this level. A characteristic is irrelevant to the asset pricing model if all Γ_β elements in these two rows are zero. Our tests of characteristic l's significance are based on the W_β statistic, described in Section 3.3, that measures the distance of these two Γ_β rows from zero.

Table X reports the contribution to overall model fit due to each characteristic. We define this contribution as the reduction in total R^2 from setting all Γ_β elements pertaining to that

Table XI
IPCA Fits Excluding Insignificant Instruments

Note. IPCA percentage R^2 at the individual stock level including only the eight characteristics from Table X that are significant at the 5% level.

		K					
		1	2	3	4	5	6
Total R^2	$\Gamma_\alpha = \mathbf{0}$	14.8	16.5	18.1	18.9	19.4	19.6
	$\Gamma_\alpha \neq \mathbf{0}$	16.1	17.7	18.7	19.2	19.5	19.6
Pred. R^2	$\Gamma_\alpha = \mathbf{0}$	0.41	0.62	1.50	1.63	1.63	1.62
	$\Gamma_\alpha \neq \mathbf{0}$	1.67	1.66	1.66	1.66	1.65	1.65

characteristic to zero while holding the remaining model estimates fixed. $**$ and $*$ denote that a variable significantly contributes to the model at the 1% and 5% levels, respectively.

Of the 36 characteristics in our sample, only five are significant at the 1% level: market beta, short-term reversal, size, momentum (the 12_2 version), and trailing 52-week high. Three more (long-term reversal, 12_7 momentum, and total assets), are significant at the 5% level. Two characteristics stand out in the magnitude of their model contribution. These are size and momentum (12_2), with each contributing roughly 2% to the restricted model's total R^2.

The insignificance of so many characteristics begs the question of whether the small subset of significant characteristics produces a factor model with similar explanatory power to the 36 characteristics dataset. Table XI repeats the IPCA model fit analysis of Table I but instead uses only the eight characteristics that are significant at the 5% level. The eight-characteristic model performs very similarly to the model with all 36 characteristics. For example, with $K = 4$, the total R^2 in the $\Gamma_\alpha = \mathbf{0}$ model is 18.9% with eight characteristics versus 19.4% with 36 characteristics, and the predictive R^2 is 1.6% versus 1.8%.

The results of Table I show that characteristics align with expected returns through their association with risk exposures rather than alphas. Additionally, Table XI suggests that the success of IPCA is obtainable using only a few characteristics, with the others being statistically irrelevant for the model's fit.

We gauge stability of this set of selected characteristics in Table XII by conducting the same characteristic tests in a number of alternative settings. First, holding the specification fixed (restricted IPCA with $K = 4$), we re-estimate the model separately for large and small stocks

Table XII
Characteristic Significance Comparison

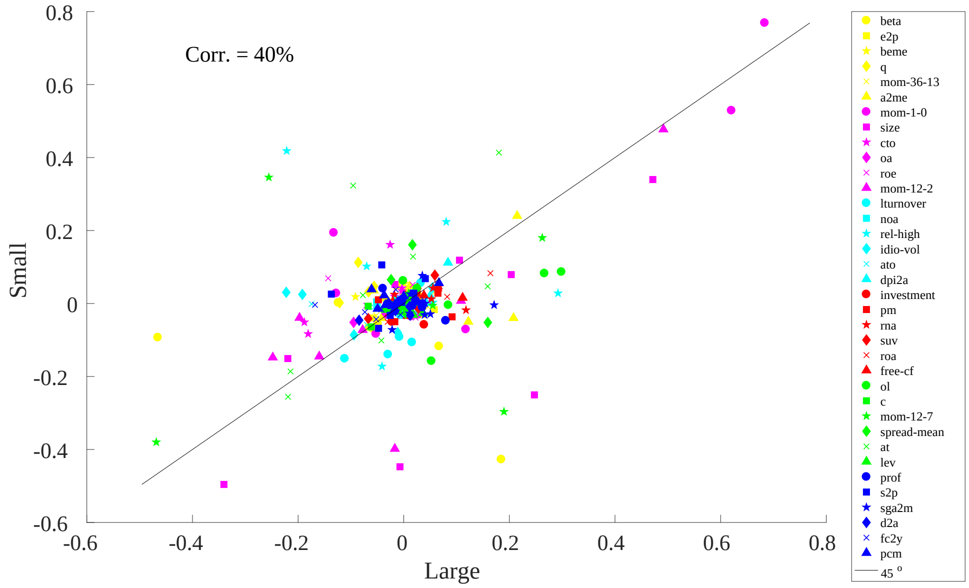
Note. Significance levels for individual characteristic contribution to overall model fit in various subsamples and model specifications. “Baseline” refers to the restricted IPCA specification with $K = 4$ using the full sample of monthly returns. Also reported are results from the large and small stock subsamples, using annual returns, and from the $K = 3$ and $K = 5$ specifications. $**$ and $*$ denote variable significance at the 1% and 5% levels, respectively.

	Baseline	Large	Small	Annual	$K = 3$	$K = 5$
size	**	*	**	**	*	**
mom_12_2	**	*	**	**	*	**
mom_1_0	**	**	**	**	**	**
mom_12_7	*	*	*	*	*	
beta	**	**	**	**	**	**
rel_high	**	**	**	**	**	**
ol		*		**	*	
at	*	*	*			**
cto				*		
idio_vol						
lturnover			*			**
spread_mean						
a2me						
s2p						
pcm						
fc2y						
roe						
roa				*		
sga2m						
suv				*		
mom_36_13	*				*	
pm				*		
beme						
ato						
e2p						
c		*				
d2a				*		
dpi2a				*		
q						
free_cf		*				
rna						
investment				*		
prof						
lev						
oa						
noa						

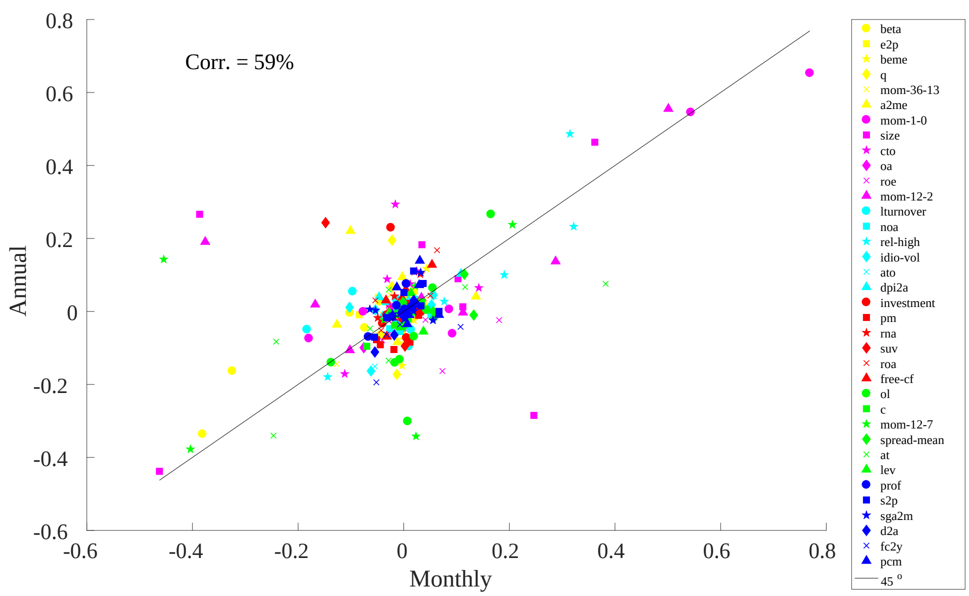
Figure 2: Parameter Stability

Note. Panel A plots Γ_β parameter estimates element-wise from the small stock sample against those from the large stock sample. Panel B likewise compares estimates from annual and monthly returns. All estimates use the four-factor restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA specification.

Panel A: Large Versus Small Stocks



Panel B: Monthly Versus Annual Returns



and report characteristic significance in each subsample. The large and small stock results disagree on only four of the 36 characteristics in terms of statistical statistical significance at the 5% level. Of the eight characteristics that are significant in the unified sample, seven are also significant in both the large and small stock samples. Annual return tests agree on six of the eight significant characteristics from the baseline monthly tests. Finally, when we alter the specification from the $K = 4$ baseline to either $K = 3$ or $K = 5$, the tests again agree on six of the eight significant characteristics.

Next, in Figure 2, Panel A plots Γ_β parameter estimates element-wise from the small stock sample against those from the large stock sample. Each characteristic appears in eight elements of Γ_β, as it enters into the specification of each loading twice (once via $\bar{c}_i$ and once via $c_{i,t} - \bar{c}_i$). The correlation of Γ_β elements across large and small stock estimates is 40%. Panel B plots Γ_β estimates from monthly returns to those from annual returns. In this case, the correlation in parameter estimates between the two return frequencies is 59%. In both panels, estimates cluster around the 45-degree line, and no characteristic shows a systematic difference in its parameter values between the two samples.

In summary, Tables XI and XII and Figure 2 demonstrate a high degree of stability in characteristics' influence on factor loadings across subsamples, data frequencies, and model specifications.

4.9 Other Robustness

Appendix C reports a number of additional empirical analyses. We summarize them here and refer interested readers to the appendix for further detail.

Our definition of $z_{i,t}$ splits each characteristic into two instruments: its mean and its time series deviation from the mean. In Appendix C.1, we analyze the relative contribution of average characteristic levels ($\bar{c}_i$) versus time series fluctuations around those levels ($c_{i,t} - \bar{c}_i$). We find that, for the purposes of describing asset riskiness (total R^2), the unconditional levels of characteristics $\bar{c}_i$ are nearly as informative as our main specification and there is little incremental contribution from $c_{i,t} - \bar{c}_i$. In contrast, predictive R^2 results show that the pure temporal variation in characteristics ($c_{i,t} - \bar{c}_i$) is the most important component for describing risk compensation. Overall, Appendix C.1 indicate that the data significantly prefer a dynamic factor model over a static one, and the variation in characteristics across assets and over time contribute differently (and both significantly) to the specification of betas.

In our main analysis, we use rank-transformed characteristics to limit the undue impact of outliers on our estimation. In Appendix C.2, we demonstrate the robustness of our findings to alternative characteristic transformations. The first normalizes each characteristic by its cross-sectional standard deviation period-by-period. This removes fluctuations in the scale of characteristics over time, while preserving differences in their relative magnitudes across stocks, and is thus a less invasive transformation than ranking. We also report results using raw characteristics with no functional transformation or outlier mitigation. The resulting fits in terms of total R^2 are nearly identical to our main analysis, while the predictive R^2 is slightly higher than in our main analysis.

As a third standardization approach, we cross-sectionally orthonormalize characteristics each period, so that $Z_t'Z_t = \mathbb{I}_L$ for all t. This standardization has a special property that the IPCA model becomes directly calculable via singular value decomposition with no need for numerical optimization, just like PCA. Appendix C.2 derives this property and shows that, with orthonormal characteristics, model fits are qualitatively similar but slightly weaker than our main IPCA results.

In Appendix C.3, we investigate the similarity between the IPCA factors and other commonly studied factors in the literature, including the FFC6 factors and 15 anomaly portfolios returns studied by Novy-Marx and Velikov (2015). We report pairwise and multiple correlations among IPCA factors and factors from prior literature. The excess market has the highest multiple correlation (93%) with the six IPCA factors, followed by idiosyncratic volatility factor (87%) followed by failure probability (84%). Among the FFC6 factors, the momentum has the highest IPCA multiple correlation (78%). We also calculate multiple correlations in the other direction, i.e. regression each IPCA factor on all 21 portfolios. The leading (highest variance) IPCA factor has a surprisingly low multiple correlation of 68% with previously studied factors, indicating that a substantial fraction of it's variation (more than 50% in terms of R^2) is unspanned by the 21 factor portfolios that we study. The remaining IPCA factor have multiple correlations with the 21 factors ranging from 75% to 89%.

The absence of book-to-market from the list of significant characteristics is surprising given its prominence in the empirical asset pricing literature. A possible explanation is that our sample includes financial stocks along with non-financials, and book-to-market ratios may be incomparable across these two groups. Appendix C.4 assesses the robustness of our results to excluding financial stocks (SIC codes 6000–6999). We find that financial stocks are not responsible for our findings regarding book-to-market. Γ_β estimates from the sample of all stocks are 95% correlated with those from the non-financial stock sample. Total and predictive R^2's are nearly identical when restricting our analysis to non-financials, and book-to-market remains statistically insignificant in the non-financials sample.

5 Conclusion

Our primary conclusions are three-fold. First, by estimating latent factors as opposed to relying on pre-specified observable factors, we find a low dimension factor model that is successfully describes riskiness of stock returns (by explaining realized return variation) and risk compensation (by explaining cross section differences in average returns). We show that there are no significant anomaly intercepts associated with a large collection stock characteristics, and instead show that the differences in average returns across stocks align with differences in exposures to a few common factors.

Second, our factor model outperforms leading observable factor models, such as the Fama-French five-factor model, in delivering small pricing errors. This is true in-sample and out-of-sample. Our factors also achieve a higher level of out-of-sample mean-variance efficiency than alternative models.

Third, only a small subset of the stock characteristics in our sample are responsible for IPCA's empirical success. 80% of the characteristics in our sample are statistically irrelevant for describing returns. Our tests conclude that the 20% of characteristics that significantly contribute to our model do so by better identifying dynamic latent factor loadings, and show no statistical evidence of generating anomaly alphas.

The key to isolating a successful factor model is incorporating information from stock characteristics into the estimation of factor loadings. In our asset pricing model, risk loadings are depend on observable asset characteristics. We propose a new method, instrumental principal components analysis (IPCA), which treats characteristics as instrumental variables for estimating dynamic loadings on latent factors. The estimator is as easy to work with as standard PCA while allowing the researcher to bring information beyond just returns into estimation of factors and betas.

We introduce a set of statistical asset pricing tests that offer a new research protocol for evaluating hypotheses about patterns in asset returns. When researchers encounter a new anomaly characteristic, they should evaluate its significance in a multivariate setting against the large body of previously studied characteristics via IPCA. In doing so, the researcher can draw inferences regarding the incremental explanatory power of the candidate characteristic after controlling for the wide gamut of previously proposed predictors. And, if the characteristic does contribute significantly to the return model, a researcher can then test whether it contributes as a risk factor loading or as an anomaly alpha. Thus, the researcher need no longer ask the narrow question, "Is my proposed characteristic/return association explained by a specific set of pre-specified factors," and instead can ask "Does there exist *any* set of factors that explains the observed characteristic/return pattern?"

Finally, our model has an exciting practical benefit. It allows investors and managers to easily assess a firm's cost of capital without relying on the obviously misspecified CAPM beta or other factor loadings that may be infeasible to estimate with time series regression. Instead, IPCA prescribes a simple cost of capital calculation as a function of the asset's observable characteristics and estimated model parameters. The essence of IPCA is to describe the riskiness and commensurate expected return of an asset by viewing it as an evolving collection of its defining characteristics. It estimates a set of universal parameters, Γ_β, that map characteristics into factor loadings and thus into expected returns. These parameters do not depend on time or the asset in question, so once they are estimated from a group of representative assets, they can be used to extrapolate conditional expected returns for other assets whose characteristics are available even if a long history of returns is not.

References

Asness, C. S., A. Frazzini, and L. H. Pedersen (2014): "Quality Minus Junk," Discussion paper, AQR.

Barra (1998): "United States Equity Version 3 (E3)," *White Paper*.

Carhart, M. (1997): "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52(1), 57–82.

Chamberlain, G., and M. Rothschild (1983): "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51(5), 1281–304.

Chordia, T., A. Goyal, and J. Shanken (2015): "Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics," Working papers, Emory Goizueta and Swiss Finance Institute.

Cochrane, J. (2005): *Asset Pricing*. Princeton University Press, second edn.

Connor, G., and R. A. Korajczyk (1986): "Performance measurement with the arbitrage pricing theory: A new framework for analysis," *Journal of financial economics*, 15(3), 373–394.

Connor, G., and R. A. Korajczyk (1988): "Risk and return in an equilibrium APT : Application of a new test methodology," *Journal of Financial Economics*, 21(2), 255–289.

Daniel, K., and S. Titman (1997): "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns," *Journal of Finance*, 52(1), 1–33.

Daniel, K., S. Titman, et al. (2012): "Testing factor-model explanations of market anomalies," *Critical Finance Review*, 1(1), 103–139.

Fama, E., and K. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33(1), 3–56.

Fama, E., and J. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81(3), 607–36.

Fama, E. F., and K. R. French (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116(1), 1–22.

Fan, J., Y. Liao, and W. Wang (2016): "Projected principal component analysis in factor models," *Ann. Statist.*, 44(1), 219–254.

Ferson, W., and C. Harvey (1991): "The Variation of Economic Risk Premiums," *Journal of Political Economy*, 99(2), 385–415.

Freyberger, J., A. Neuhierl, and M. Weber (2017): "Dissecting Characteristics Nonparametrically," Working papers, University of Wisconsin and Notre Dame and Chicago Booth.

Gibbons, M. R., S. Ross, and J. Shanken (1989): "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57(5), 1121–52.

Goncalves, S., and L. Kilian (2004): "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 123(1), 89–120.

Goyal, A. (2012): "Empirical cross-sectional asset pricing: a survey," *Financial Markets and Portfolio Managament*, 26, 3–38.

Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica: Journal of the Econometric Society*, pp. 1029–1054.

Hansen, L. P., and S. F. Richard (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55(3), 587–613.

Kelly, B. T., S. Pruitt, and Y. Su (2017): "Instrumented Principal Component Analysis," Working papers, Chicago Booth and ASU WP Carey.

Kozak, S., S. Nagel, and S. Santosh (2017): "Shrinking the Cross-Section," Discussion paper, Michigan Ross and Chicago Booth and Maryland Smith.

——— (forthcoming): "Interpreting Factor Models," *Journal of Finance*.

Lewellen, J. (2015): "The Cross-section of Expected Stock Returns," *Critical Finance Review*, 4, 1–44.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A skeptical appraisal of asset pricing tests," *Journal of Financial economics*, 96(2), 175–194.

Light, N., D. Maslov, and O. Rytchkov (2017): "Aggregation of Information About the Cross Section of Stock Returns: A Latent Variable Approach," *Review of Financial Studies*, 30(4), 1339–1381.

Roll, R. (1977): "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory," *Journal of Financial Economics*, 4(2), 129–176.

ROSENBERG, B. (1974): "Extra-Market Components of Covariance in Security Returns," *Journal of Financial and Quantitative Analysis*, 9(2), 263–274.

ROSS, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

STOCK, J. H., AND M. W. WATSON (2002a): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97(460), 1167–1179.

——— (2002b): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 20(2), 147–62.

Internet Appendix

A Estimation via Alternating Least Squares

This appendix describes our ALS approach to optimizing IPCA objective function (5). To initialize the algorithm, we choose a starting guess for Γ_β as the left eigenvectors corresponding to the leading K eigenvalues of the characteristic-managed portfolio second moment matrix, $\sum_t x_t x_t'$. As described in Section 2, this initial guess, which amounts to the static loading matrix estimate from applying PCA to the x_t dataset, is a close approximation to the exact solution of (5) as long as $Z_t' Z_t$ is not too volatile. Initializing the optimization with this guess ensures that the algorithm converges very quickly (typically within 10 iterations, taking roughly 0.25 seconds on a standard desktop computer).

Given the initial guess for Γ_β, we evaluate the least squares regression corresponding to first-order condition (6) for all t. Then, given the resulting solutions f_t's, we evaluate the least squares regression corresponding to first-order condition (7). We iterate between evaluations of (6) and (7) until convergence, defined as the point at which the maximum absolute change in any element of Γ_β or f_t (for all t) is smaller than 10^{-6}.

B Bootstrap Monte Carlo Analysis

This appendix investigates the finite sample behavior of our test for the null hypothesis that $\Gamma_\alpha = \mathbf{0}$. Our analysis focuses on assessing the size of the test (rejection rates under the null) and power of the test (rejection rates under the alternative).

We simulate data according to the following general model:

$$r_t = Z_{t-1}\Gamma_\alpha \kappa + (Z_{t-1}\Gamma_\beta + \nu_t)F_t + \eta_t.$$

We use cross sections sizes (N) of 500, 5000, or 25000 assets with monthly observation count (T) of 100 or 1000. We set number of factors (K) to 3 or 5 factors, and consider instrument set sizes (L) of 10 or 20.

We draw factor realizations f_t as a $K \times 1$ vector of standard normals. We draw Γ_β as a random orthonormal matrix. To do this, we draw an $L \times K$ matrix of normals, M, and set $\Gamma_\beta = M M_C^{-1}$ for M_C the Cholesky of $M'M$. We draw Z_t as an $N \times L$ matrix of standard

Table A.1
Size and Power of $\Gamma_\alpha = 0$ Test

Note. Simulated rejection probabilities of a 5%-level test. Rejection rates in the $\kappa = 0$ column describe the size of the test. Rejection rates in the $\kappa = \kappa_1, \kappa_2$, and κ_3 columns describe the power of the test.

				Wild				Non-wild			
N	T	K	L	0	κ_1	κ_2	κ_3	0	κ_1	κ_2	κ_3
500	100	3	10	3.4	5.4	32.0	75.0	4.4	5.4	11.3	70.0
500	100	5	20	5.6	6.6	30.2	81.4	5.0	4.4	7.4	68.4
5000	100	3	10	4.6	6.2	27.6	93.4	4.0	5.8	6.6	93.4
5000	100	5	20	4.2	5.2	15.8	82.4	3.2	2.8	5.8	86.8
5000	1000	3	10	4.2	18.0	24.8	75.0	4.6	11.2	21.0	71.8
5000	1000	5	20	3.4	19.8	50.6	90.4	3.4	15.0	32.6	87.8
25000	100	3	10	2.6	7.2	16.0	44.8	2.2	5.6	14.8	38.0
25000	100	5	20	3.6	7.4	15.6	98.6	2.2	6.8	14.6	44.2
25000	1000	3	10	4.4	66.2	78.4	95.8	3.0	54.6	80.6	99.2
25000	1000	5	20	5.0	77.0	93.2	99.8	4.6	76.4	95.2	100.0
Average				4.1	21.9	38.4	83.7	3.7	18.8	29.0	76.0

normals. We allow betas to possess a purely unobservable component, ν_t, which is an $N \times K$ array of independent normals with the variance parameter chosen such that 50% of the variation of $\beta_{t-1} \equiv Z_{t-1}\Gamma_\beta + \nu_t$ is due to ν_t.

Idiosyncratic returns, η_t, are drawn as a $N \times 1$ array of normals with time-varying volatility. In particular, the conditional variance η_t is log-normally distributed with its own variance chosen such that on average 15% of return variance is systematic and 85% is idiosyncratic. By incorporating heteroskedasticity akin to that in the data, we can assess the usefulness of the wild bootstrap in conducting inference.

Γ_α is a random normal vector that is orthogonal to Γ_β. In particular, we draw a $L \times 1$ vector of normals, Y, and set $\Gamma_\alpha = (I_L - \Gamma_\beta(\Gamma_\beta'\Gamma_\beta)^{-1}\Gamma_\beta')Y\kappa$, where κ controls the distance from the null. When $\kappa = 0$ the model embodies the null hypothesis that $\Gamma_\alpha = 0$. For $\kappa > 0$, data is generated under the alternative. We consider three values of $\kappa > 0$ such that when $\kappa = \kappa_x$, $Z_{t-1}\Gamma_\alpha\kappa_x$ drives about $x \times 0.25\%$ of the variation in returns.

In every simulated data set, we calculate the W_α statistic and perform the bootstrap procedure of Section 3.1 to arrive at a p-value. We also consider a variation on our test that uses a standard rather than wild bootstrap. We conduct 1000 simulations with 500 bootstraps for each simulation. Table A.1 reports the results.

In all cases, the test maintains appropriate size when data are generated under the null. Rejection rates based on a 5% significance level never fall below 2.2% or rise above 5.6%. The power of the test to reject the null when data are generated under the alternative also behaves well. Rejection rates increase steadily with κ and with sample size. Finally, we find that the wild bootstrap indeed improves inference, both in terms of size and power, compared to the standard bootstrap.

C Additional Empirical Results

C.1 Static or Dynamic Loadings?

Our definition of $z_{i,t}$ splits each characteristic into two instruments: its mean and its time series deviation from the mean. We next analyze the relative contribution of these two components to the return factor model. In addition to our main specification in which $z_{i,t} = (\bar{c}_i', c_{i,t}' - \bar{c}_i')'$, we consider three nested variations of the IPCA model. The first sets instruments equal the total characteristic value, $z_{i,t} = c_{i,t}$. This is equivalent to imposing that the Γ_β coefficients corresponding to $\bar{c}_i$ and $c_{i,t} - \bar{c}_i$ are equal, and helps answer the question, "Do the level and variation in characteristics contribute different information to factor loadings?" If this specification performs as well as our main specification, then level and variation enter $\beta_{i,t}$ equally and there is no need to split them.

The second specification sets instruments equal to the mean characteristic value, $z_{i,t} = \bar{c}_i$. This is equivalent to setting Γ_β coefficients on to $c_{i,t} - \bar{c}_i$ equal to zero. By testing this against our main specification we address the question, "Is a static factor model sufficient for describing returns?," in which case the benefits of characteristics shown in our preceding results arise from their ability to better differentiate loadings across assets rather than over time. The third specification asks the complementary question, "Is time series variation in characteristics the primary contributor to IPCA success?" In this case we set instruments equal to the deviation value only, $z_{i,t} = c_{i,t} - \bar{c}_i$, and fix coefficients on $\bar{c}_i$ to zero.

Because these comparisons can be formulated as restrictions on rows of Γ_β, the model comparison test in Section 3.3 can be used to conduct formal statistical inference for whether our main split-characteristic specification significantly improves over each of the three nested

Table A.2
Static Versus Dynamic Loadings

Note. Percentage R^2 from IPCA specifications based on either total characteristics ("c"), average characteristic levels ("$\bar{c}$"), time series deviations from average levels ("$c-\bar{c}$"), or our baseline specification in which levels and deviations are allowed to enter with difference coefficients ("$\bar{c}, c-\bar{c}$").

	Total R^2				Predictive R^2			
K	c	$\bar{c}$	$c-\bar{c}$	$\bar{c}$, $c-\bar{c}$	c	$\bar{c}$	$c-\bar{c}$	$\bar{c}$, $c-\bar{c}$
1	14.8	14.4	13.3	15.1	0.35	0.28	0.52	0.44
2	16.4	16.1	15.0	17.0	0.34	0.44	1.04	1.64
3	17.4	16.8	15.9	18.5	0.42	0.60	1.25	1.70
4	18.0	17.4	16.4	19.4	0.42	0.88	1.25	1.83
5	18.6	17.8	16.7	19.9	0.69	0.88	1.24	1.82
6	18.9	18.0	17.0	20.3	0.68	0.89	1.25	1.82

variations.[27]

Table A.2 reports model fits for each specification. In terms of total R^2, the nested sub-specification that is closest to our main specification for $K = 4$ is that using the total characteristic, $z_{i,t} = c_{i,t}$. It produces a total R^2 of 18.0%, versus 19.4% in the main model. However, for the purposes of describing asset riskiness, even the unconditional levels of characteristics $\bar{c}_i$ are nearly as informative as our main specification.

In contrast, the predictive R^2 results show that temporal variation around the mean is the most important characteristic component for describing risk compensation. Conditional expected return estimates from time series variation alone describe 1.3% of total return variance, versus 1.8% from the main split-characteristic specification and 0.9% for the $\bar{c}_i$ specification. Lastly, the statistical test rejects all three nested variations in favor of the more general split-characteristic specification with p-values below 1%.[28] This is true for all K. In summary, the data significantly prefer a dynamic factor model over a static one, and the variation in characteristics across assets and over time contribute differently (and both significantly) to the specification of betas.

[27]Goyal (2012) notes that "In practice, one almost always employs firm characteristics that vary over time. There are relatively few analytical results in the literature for the case dealing with time-varying characteristics." The tests we describe in this section provide a means of investigating the role of time-varying characteristics in a fully formulated statistical setting.

[28]We omit percentage p-values from the table because as they all are zero to two decimal places.

Table A.3
Other Characteristic Transformations

Note. This table repeats the analysis of Panel A in Table I using characteristics that are cross-sectionally volatility-scaled each period (Panel A), untransformed (Panel B), or cross-sectionally orthonormalized each period (Panel C).

	K					
	1	2	3	4	5	6
	Panel A: Volatility-scaled					
Total R^2	15.2	17.6	19.4	20.2	20.8	21.1
Pred. R^2	0.53	2.10	2.17	2.26	2.27	2.26
	Panel B: Raw					
Total R^2	15.6	17.9	19.5	20.4	21.0	21.4
Pred. R^2	0.58	2.22	2.29	2.27	2.32	2.17
	Panel C: Orthonormal					
Total R^2	14.6	16.4	17.9	18.8	19.3	19.6
Pred. R^2	0.46	0.85	1.52	1.59	1.60	1.60

C.2 Alternative Characteristic Transformations

In our main analysis, we use rank-transformed characteristics to limit the undue impact of outliers on our estimation. We now demonstrate the robustness of our findings to alternative characteristic transformations. The first normalizes each characteristic by its cross-sectional standard deviation period-by-period. This removes fluctuations in the scale of characteristics over time, while preserving differences in their relative magnitudes across stocks, and is thus a less invasive transformation than ranking. We also report results using raw characteristics with no functional transformation or outlier mitigation. The IPCA fits for these approaches are shown in Panel A and B of Table A.3, respectively. The fits in terms of total R^2 are nearly identical to our main analysis, while the predictive R^2 is slightly lower with ranked characteristics (roughly 1.9%, versus 2.1%–2.3% for volatility-scaled and raw characteristics when $K > 1$).

C.2.1 Orthonormal Characteristics: An Exact Analytical Estimator

As a third standardization approach, we cross-sectionally orthonormalize characteristics each period, so that $Z_t'Z_t = \mathbb{I}_L$ for all t. A convenient feature of this normalization is that the IPCA estimator of Γ_β becomes directly calculable via singular value decomposition with no need for numerical optimization.

This result follows from the concentrated IPCA objective described in equation (8) of Section 2. When instruments are orthonormal period-by-period, (8) reduces to

$$\max_{\Gamma_\beta} \quad \mathrm{tr}\left(\sum_{t=1}^{T-1} \left(\Gamma_\beta'\Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' r_{t+1} r_{t+1}' Z_t \Gamma_\beta\right). \tag{16}$$

That is, the objective function collapses to a sum of homogeneous Rayleigh quotients. As a result, the K leading eigenvectors of $\sum_t Z_t' r_{t+1} r_{t+1}' Z_t$ satisfy the maximization problem and thus estimate Γ_β, a solution well known from the PCA literature.

More specifically, we derive the algebraic solution for Γ_β via the following eigenvalue decomposition:

$$USU' = \sum_t Z_t' r_{t+1} r_{t+1}' Z_t,$$

The IPCA estimator of Γ_β is

$$\hat{\Gamma}_\beta = U_K$$

where the columns of U are arranged in decreasing eigenvalue order and U_K denotes the first K columns of U. The factor estimates are likewise analytical, and simplify to

$$\hat{f}_{t+1} = \hat{\Gamma}_\beta'(Z_t' r_{t+1}).$$

Once can directly impose characteristic orthonormality in the data. In particular, given some matrix of "raw" instruments $\tilde{Z}_t$, we construct orthonormal instruments Z_t using the Gram-Schmidt process. This uses regression to sequentially orthogonalize instruments in the cross section each period, then cross-sectionally volatility-standardizes the residuals. This orthogonalization is not invariant to the ordering of characteristics. As a result, our tests of individual characteristics can be influenced by the order of the instruments. We choose an instrument ordering on economic grounds. In particular, characteristics are arranged within $z_{i,t}$ according to the date that the proposed characteristic effect was published. Thus market beta is ordered first, size second, and so forth.

We report IPCA fits for orthonormalized instruments in Panel C of Table A.3. Model fits are qualitatively similar, though slightly weaker, compared to our main IPCA results, generate total R^2's of nearly 20%. Predictive R^2's reach 1.6% and handily outperform observable factor models in this dimension.

Table A.4
Correlation with Observable Factors

Note. Pairwise and multiple correlations in percent. The first six columns report pairwise correlations between each of the six IPCA factors (listed in the corresponding column) and each of the previously proposed portfolios (rows). The last column reports the absolute multiple correlation ($\sqrt{R^2}$) from a regression of each portfolio on the six IPCA factors. The last column reports the absolute multiple correlation of each IPCA factor with the 21 portfolios.

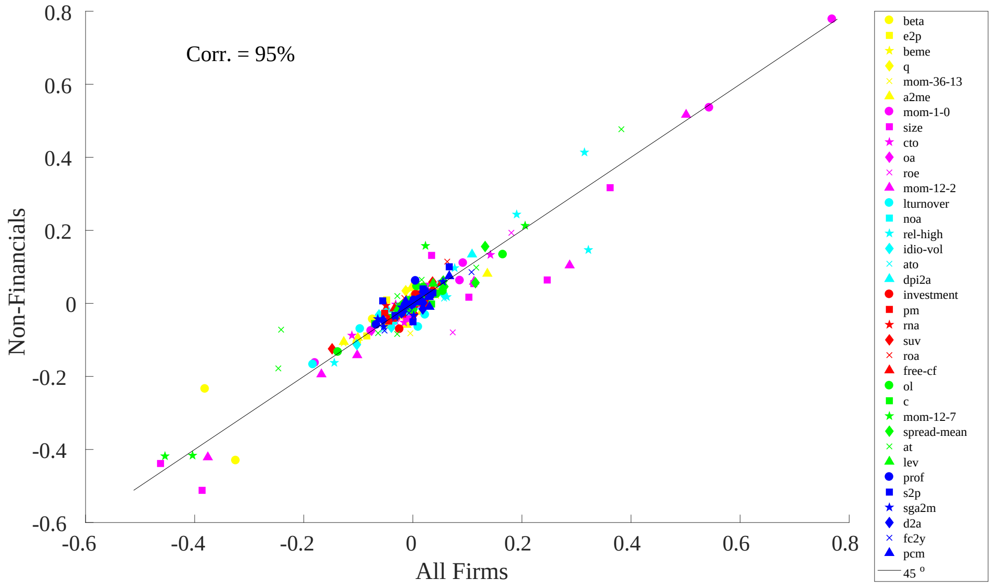
Portolio	f_1	f_2	f_3	f_4	f_5	f_6	$\sqrt{R^2}$
Mkt-RF	24.3	−2.8	−39.0	59.4	−59.8	2.9	93.4
SMB	43.6	−47.2	−43.4	17.8	−29.2	−7.0	74.7
HML	−6.7	2.1	46.9	−13.5	−4.6	−50.9	70.9
RMW	−23.0	45.3	16.4	3.9	17.5	−21.7	56.9
CMA	−2.2	−0.4	33.7	−27.2	8.1	−37.9	58.5
MOM	2.3	34.1	−54.8	−42.2	15.2	1.9	77.8
Net Issuance	−19.2	22.4	29.5	−13.7	13.6	−25.7	49.6
Ret. Book Equity	−29.8	57.2	4.7	−18.7	24.9	−9.9	66.3
Failure Prob.	−34.9	63.4	−7.0	−38.3	40.3	−2.5	83.8
ValMomProf	15.4	2.7	−38.7	−25.9	22.6	−12.0	53.8
ValMom	3.8	13.5	−20.0	−38.2	3.6	−24.9	51.6
Idios. Volatility	−39.8	53.3	35.3	−28.6	42.1	−26.5	87.0
Momentum	−3.3	38.6	−45.9	−39.9	17.8	0.5	73.4
PEAD SUE	−8.8	36.2	−30.7	−26.0	5.2	15.1	54.5
PEAD CAR	−8.9	21.4	−21.0	−25.3	5.4	8.7	39.2
Ind. Mom.	−16.4	8.0	−6.2	−25.7	9.7	−0.4	33.5
Ind. Rel. Reversal	20.4	−24.2	14.9	39.9	−27.4	8.6	58.0
High Freq. Combo	2.6	−5.9	−6.7	3.4	−14.2	3.2	18.3
Short-run Reversal	18.9	−19.6	11.1	38.0	−20.8	6.1	50.8
Seasonality	8.4	1.5	−4.8	3.4	−6.1	5.4	13.5
IRR (Low vol.)	11.2	−16.8	0.0	23.1	−23.4	0.3	37.1
$\sqrt{R^2}$ All Ptfs.	55.1	75.4	84.4	76.5	72.2	68.9	

C.3 Comparison with Factors in Prior Literature

Table A.4 reports pairwise and multiple correlations between the estimated factors in our restricted $K = 6$ IPCA specification and 21 portfolios from prior literature, including the FFC6 factors and the 15 anomaly portfolios from Novy-Marx and Velikov (2015). The first six columns report pairwise correlations between factors listed in the corresponding column and row. The last column reports the absolute multiple correlation ($\sqrt{R^2}$) from a regression of each portfolios on the six IPCA factors. The last column reports the absolute multiple correlation of each IPCA factor with the 21 portfolios.

Figure A.1: Parameter Stability

Note. The figure plots Γ_β parameter estimates element-wise from the all stock sample against those from the sample excluding financials (SIC codes 6000–6999) in the four-factor restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA specification.



C.4 Exclusion of Financial Stocks

This appendix assesses the robustness of our results to excluding financial stocks, defined as those having SIC codes 6000–6999. Figure A.1 compares Γ_β estimates from the sample including all stocks (financial and non-financials) to the sample excluding financials.

Given the insignificance of book-to-market ratios in our main analysis, and given the stark differences in book-to-market for financial and non-financial firms, we are particularly interested in how the role of book-to-market ratios for IPCA factor loadings in the non-financial sample might differ from that for the financial sample. When we conduct significance tests in the non-financials sample, book-to-market remains statistically insignificant. In the four-factor restricted ($\Gamma_\alpha = \mathbf{0}$) IPCA specification, the non-financial total and predictive R^2's are 19.6% and 1.8%, essentially unchanged from the R^2's in our baseline sample that includes financials.