CESifo WORKING PAPERS

8377
2020

Original Version: June 2020
This Version: October 2022

Are Characteristics Covariances? A Comment on Instrumented Principal Component Analysis

Christian Fieberg, Lars Hornuf, Gerrit Liedtke, Thorsten Poddig

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8377*

Are Characteristics Covariances? A Comment on Instrumented Principal Component Analysis

Abstract

We present analytical and simulation-based evidence that instrumented principal component analysis (IPCA) cannot reliably distinguish between whether covariances or characteristics explain asset returns because the question has to be answered jointly with the question of how many factors have to be modeled. IPCA finds a covariance-based explanation when estimating too many factors ("alpha-eating") and a characteristic-based explanation when estimating too few factors ("beta-eating"). Our results therefore call into question the empirical evidence recently obtained that stocks (Kelly et al., 2019), options (Büchner and Kelly, 2022), and bonds (Kelly et al., 2021) are explained by covariances.

JEL-Codes: C230, G120, G170.

Keywords: IPCA, covariances, characteristics, cross section of asset returns.

Christian Fieberg
University of Bremen
Max-von-Laue-Straße 1
Germany – 28359 Bremen
cfieberg@uni-bremen.de

Lars Hornuf
University of Bremen
Max-von-Laue-Straße 1
Germany – 28359 Bremen
hornuf@uni-bremen.de

Gerrit Liedtke
University of Bremen
Max-von-Laue-Straße 1
Germany – 28359 Bremen
gliedtke@uni-bremen.de

Thorsten Poddig
University of Bremen
Max-von-Laue-Straße 1
Germany – 28359 Bremen
poddig@uni-bremen.de

2022-10-05
The paper has benefited from helpful comments from Jerry Coakley, Amit Goyal, Michele Meoli, Simon Price, Frank Seifried, Marc Rieger, Silvio Vismara and participants at the Essex Finance Centre research seminar in Colchester, UK, the Quantitative Finance and Risk Analysis Seminar in Trier, Germany, the University of Bergamo Lunch Seminar, Italy, and the 3rd Frontiers of Factor Investing Conference in Lancaster, UK.

1 Introduction

In his presidential address, Cochrane (2011) summarizes the key findings from the past 40 years of asset pricing research and highlights two central challenges for future research. First, given the large number of potential predictors for the cross section of (stock) returns, those that provide independent information have to be identified. Due to the multidimensionality of the asset pricing puzzle (Green et al., 2017; Freyberger et al., 2020), new methods beyond portfolio sorts and cross-sectional regressions are required to tame the "zoo of anomalies". Second, assuming that a (new) predictor provides independent information about average returns, the question of what explains this relationship has to be answered.

Kelly et al. (2019) (KPS hereafter) seek to address the challenges raised by Cochrane (2011) by proposing instrumented principal component analysis (IPCA), a new factor modeling approach that finds latent (risk) factors, while estimating factor loadings as a function of a potentially large number of observed asset characteristics. This reduction in dimensionality from a large number of characteristics to a smaller number of factor loadings allows IPCA to process considerable information while being parsimoniously parameterized. Additionally, IPCA allows for time-variation in factor loadings due to dynamics in asset characteristics. KPS show that IPCA outperforms other competing factor models in explaining the covariation in stock returns and Windmüller (2021) extend this evidence to the international stock market. Büchner and Kelly (2022) and Goyal and Saretto (2022) find that IPCA also has superior performance in explaining the cross-section of option returns. Kelly et al. (2021) and Bianchi and Babiak (2022) extend this evidence to corporate bond and cryptocurrency returns, respectively. Beyond improved explanatory power for returns in several asset classes, KPS contend that IPCA can help to answer the pervasive question of whether cross-sectional asset return variation associated with asset characteristics is due to covariances with common risk factors or due to the characteristics themselves (Daniel and Titman, 1997; Davis et al., 2000; Daniel et al., 2001). Specifically, if characteristics proxy for systematic risk,[1] IPCA identifies the corresponding latent risk factors and the conditional factor loadings (i.e., betas). If there is a characteristic story above and beyond factor loadings, IPCA will additionally find a conditional

[1]According to KPS, the term "risk" refers to statistical covariation in asset returns and is, therefore, not restricted to shocks to economic fundamentals (e.g., shocks to productive technologies). Thus, the term "risk" also includes market-wide behavioral shocks (e.g., changes in investor preferences).

alpha, which represents a return unrelated to risk, i.e., the return is anomalous. KPS fit beta-only (i.e., only allowing for beta) and alpha-beta (i.e., allowing for alpha and beta) IPCA specifications to U.S. stock return data and find that a beta-only model with five factors explains the same amount of variation as the alpha-beta model. The authors present a hypothesis test for testing the null hypothesis that characteristics do not align with alpha and find that the alpha-beta model is rejected in favor of the beta-only model if the beta-only model includes at least five factors. Accordingly, they conclude that characteristics proxy for the covariances to common risk factors and that there are no "anomalous returns" in the sense that the characteristics themselves drive the returns. Similarly, Büchner and Kelly (2022) and Kelly et al. (2021) find that beta-only three- and five-factor models are sufficient for explaining the cross section of option and bond returns, respectively. These results suggest that not only stock returns but also option and bond returns are driven by covariances not characteristics.

Although KPS demonstrate that IPCA's central features, i.e., its low parametrization, utilized information in the characteristics, and the time-variation in factor loadings, lead to its superior performance compared to existing factor models, they do not present evidence that IPCA can truly differentiate between covariances and characteristics. Therefore, the aim of this paper is to test for the validity of their findings in two ways. First, we present analytical evidence that, irrespective of whether asset returns are truly driven by covariances or characteristics, IPCA can always identify an indistinguishable model that fits the other explanation. Second, we simulate asset return data according to three different data-generating processes (DGPs) that correspond to (1) the covariance story, (2) the characteristic story, or (3) a mixture of both by creating data according to a multifactor and/or a characteristic model. For all simulated datasets, we fit three IPCA versions, one allowing for beta only, one allowing for alpha only, and a third that allows for both alpha and beta. We then apply the hypothesis test proposed by KPS to assess whether IPCA can reliably identify the true model.

Our results suggest that the question of whether covariances or characteristics explain asset returns cannot be answered independently of the question of how many factors must be modeled. If IPCA estimates more factors than there are in the true model, we observe an "alpha-eating" effect, that is, factor loadings subsume the return variation that is actually attributable to an anomaly intercept. The "alpha-eating" effect biases the results such that the hypothesis test proposed by

KPS fails to reject the null hypothesis of a zero anomaly alpha, resulting in distorted conclusions when trying to distinguish between the "covariances vs. characteristics" story. Conversely, if IPCA estimates too few factors, a "beta-eating" effect occurs. That is, alpha subsumes some explanatory power of the omitted factors. Even if IPCA estimates the true but in reality unknown number of factors, the alpha- or beta-eating effect causes a correct model to be indistinguishable from an incorrect model. Thus, given all of these results, we find that IPCA cannot reliably distinguish between covariances and characteristics because either the "alpha-eating" or the "beta-eating" effect will distort the results. By the nature of the test specification proposed in KPS, one will always find that covariances explain returns if a sufficient number of factors is included. Therefore, IPCA cannot provide an answer to the "characteristics vs. covariances" debate.

This paper is structured as follows. In Section 2, we introduce the general IPCA model and decompose the model to analytically derive the "beta-eating" and "alpha-eating" effects. Section 3 presents our simulation study. Section 4 summarizes and concludes the paper.

2 A model for the cross section of asset returns

The central IPCA specification presented in KPS assumes that asset (excess) returns are generated according to a linear multifactor model as described in equation (1):

$$r_{i,t+1} = \alpha_{i,t} + \boldsymbol{\beta}_{i,t}\boldsymbol{f}_{t+1} + \epsilon_{i,t+1} \tag{1}$$

with $r_{i,t+1}$ being the (excess) return of asset i at time $t+1$, $\boldsymbol{f}_{t+1}$ being a $K \times 1$ vector of contemporaneous (latent) factor realizations, with K denoting the number of factors, $\boldsymbol{\beta}_{i,t}$ being a corresponding $1 \times K$ vector of asset-specific (conditional) factor loadings (i.e., betas), $\alpha_{i,t}$ is an asset-specific (conditional) intercept (i.e., alpha), and $\epsilon_{i,t+1}$ is the residual return of asset i at time $t+1$. It is assumed that there is an $L \times 1$ vector of asset i's characteristics $\boldsymbol{z}_{i,t}$ (with L denoting the number of asset characteristics) at time t that is informative about either the alpha, the betas, or both. Accordingly, the alpha and the betas are conditional on observed characteristics, as described in equations (2) and (3), respectively:

$$\alpha_{i,t} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\boldsymbol{\alpha}} + \nu_{\alpha,i,t} \tag{2}$$

and

$$\boldsymbol{\beta}_{i,t} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\boldsymbol{\beta}} + \boldsymbol{\nu}_{\beta,i,t} \tag{3}$$

with $\boldsymbol{\Gamma}_{\boldsymbol{\alpha}}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\beta}}$ denoting an $L \times 1$ vector and an $L \times K$ matrix of characteristic coefficients for the alpha and the betas, respectively. The terms $\nu_{\alpha,i,t}$ and $\boldsymbol{\nu}_{\beta,i,t}$ describe errors in the alpha or beta estimates that are uncorrelated with the asset characteristics. KPS note that if all coefficients of $\boldsymbol{\Gamma}_{\boldsymbol{\alpha}}$ are zero, this does not imply that there are no alphas in general but that alphas are not linked to asset characteristics and therefore no "anomalous returns" exist.

Without loss of generality, we can decompose asset (excess) returns as described in equation (1) into expected returns $E_t\left[r_{i,t+1}\right]$ and short-term return innovations $\widetilde{r}_{i,t+1}$. Specifically, we decompose the return that is due to the (latent) factors into constant factor premia $\boldsymbol{\lambda}$ and stochastic factor innovations $\widetilde{\boldsymbol{f}}_{t+1}$. Equation (4) presents the reformulated general IPCA model:

$$r_{i,t+1} = \underbrace{\alpha_{i,t} + \boldsymbol{\beta}_{i,t}\boldsymbol{\lambda}}_{E_t\left[r_{i,t+1}\right]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{4}$$

where $E_t\left[r_{i,t+1}\right]$ denotes the conditional expectation in t for the return of asset i at time $t+1$ and $\widetilde{r}_{i,t+1}$ denotes the corresponding (unexpected) return innovation. The conditional alpha and conditional betas are defined as in the equations (2) and (3), respectively. The return innovations $\widetilde{r}_{i,t+1}$ can be further decomposed into systematic return innovations $\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1}$ that explain the (systematic) covariance matrix across asset returns and idiosyncratic innovations $\epsilon_{i,t+1}$.

The term $E_t\left[r_{i,t+1}\right]$ is of central interest in asset pricing studies and subject to the question of whether expected returns are explained by covariances to common risk factors (i.e., $\boldsymbol{\beta}_{i,t}$) or characteristics (i.e., $\alpha_{i,t}$). Note that in what follows, we assume that there are systematic innovations in returns; otherwise, assets would be cross-sectionally uncorrelated. Thus, there are nonzero betas on common risk factors; however these factors do not necessarily have nonzero premia $\boldsymbol{\lambda}$. According to equation (4), three possible explanations for expected returns can be identified. First, in a covariance-only world, characteristics do not contribute to expected returns and are therefore fully described by covariances to common risk factors (i.e., $\boldsymbol{\Gamma}_\alpha = \boldsymbol{0}$ and $\boldsymbol{\lambda} \neq \boldsymbol{0}$). Second, the true DGP may be a characteristic-only world in which the characteristics themselves explain expected returns and covariances to common risk factors at most explain short-term return innovations (i.e., $\boldsymbol{\Gamma}_\alpha \neq \boldsymbol{0}$

and $\boldsymbol{\lambda} = \mathbf{0}$). Last, asset returns may be generated according to a hybrid covariance-characteristic world, in which both covariances and characteristics drive expected returns (i.e., $\boldsymbol{\Gamma}_\alpha \neq \mathbf{0}$ and $\boldsymbol{\lambda} \neq \mathbf{0}$). In the following, we analytically show how IPCA misspecification can bias the results, leading to erroneous conclusions regarding the "covariances vs. characteristics" story.

2.1 Beta-eating

We begin by assuming that asset returns are described by a covariance-only model, that is, assets earn different returns because of different compensations for bearing systematic risk. In a covariance-only world, the $\alpha_{i,t}$ from equation (4) vanishes as shown in equation (5):

$$r_{i,t+1} = \underbrace{\boldsymbol{\beta}_{i,t}\boldsymbol{\lambda}}_{E_t[r_{i,t+1}]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{5}$$

Inserting equation (3) into (5), the resulting (true) DGP can be written as:

$$r_{i,t+1} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_\beta\boldsymbol{\lambda} + \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_\beta\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1} \tag{6}$$

Assume that a researcher misspecifies IPCA and estimates an alpha-only IPCA, that is, an IPCA that estimates no factors but only a conditional alpha. We denote $\boldsymbol{\Gamma}_\alpha^*$ as the estimate of the characteristic coefficients for the conditional alpha, obtained from the alpha-only IPCA. Since alpha has no innovations, the alpha-only IPCA cannot model systematic return innovations $\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1}$. We therefore restrict attention to the expected return component, which is of central interest in asset pricing studies and the subject of the "covariance vs. characteristic" debate. The true DGP process for expected returns is described as:

$$E_t[r_{i,t+1}] = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_\beta\boldsymbol{\lambda} \tag{7}$$

Given a set of $l = 1, \ldots, L$ asset characteristics and $j = 1, \ldots, K$ risk factors, IPCA finds an equivalent alpha-only description of expected returns that satisfies

$$
\begin{aligned}
E_t\left[r_{i,t+1}\right] &= \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}\boldsymbol{\lambda} \\
&\mathrel{\widehat{=}} \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\alpha}^{*}
\end{aligned}
\tag{8}
$$

if there is a solution to the following system of linear equations:

$$
\begin{array}{ccccccccc}
\Gamma_{\beta,1,1}\lambda_1 & + & \Gamma_{\beta,1,2}\lambda_2 & + & \dots & + & \Gamma_{\beta,1,K}\lambda_K & = & \Gamma_{\alpha,1}^{*} \\
\Gamma_{\beta,2,1}\lambda_1 & + & \Gamma_{\beta,2,2}\lambda_2 & + & \dots & + & \Gamma_{\beta,2,K}\lambda_K & = & \Gamma_{\alpha,2}^{*} \\
 & \vdots & & & & & & & \\
\Gamma_{\beta,L,1}\lambda_1 & + & \Gamma_{\beta,L,2}\lambda_2 & + & \dots & + & \Gamma_{\beta,L,K}\lambda_K & = & \Gamma_{\alpha,L}^{*}
\end{array}
\tag{9}
$$

where the left-hand side represents the true DGP and the right-hand side shows the estimates from a misspecified alpha-only IPCA setting. The term $\Gamma_{\beta,l,j}$ denotes the characteristic coefficient of the l-th characteristic for the j-th beta, λ_j denotes the factor premium of the j-th factor, and $\Gamma_{\alpha,l}^{*}$ denotes the IPCA estimate of the alpha characteristic coefficient for the l-th characteristic. The linear system of equations is uniquely determined if the l-th element of $\boldsymbol{\Gamma}_{\alpha}^{*}$ is chosen such that it equals the product of the l-th row of $\boldsymbol{\Gamma}_{\beta}$ and the vector of factor premia $\boldsymbol{\lambda}$:

$$
\Gamma_{\alpha,l}^{*} = \sum_{j=1}^{K} \Gamma_{\beta,l,j}\lambda_j \tag{10}
$$

Denoting K^{*} as the estimated number of factors in an alpha-beta IPCA, the beta-eating effect in equation (11) can be generalized to any IPCA specification in which the estimated number of factors is below the true number of factors, i.e., $K^{*} < K$:

$$
\Gamma_{\alpha,l}^{*} = \left(\sum_{j=1}^{K} \Gamma_{\beta,l,j}\lambda_j\right) - \left(\sum_{j^{*}=1}^{K^{*}} \Gamma_{\beta,l,j^{*}}\lambda_{j^{*}}\right) \tag{11}
$$

Accordingly, a true covariance-only world can always be transformed into an equivalent indistinguishable characteristic-only world (or a mixed covariance-characteristic world) if no or too few factors are considered when estimating the IPCA. As a result, the effect that a characteristic has on the expected returns is captured, but it is not possible to distinguish whether this effect is due to the beta or to the characteristic itself. We refer to this effect as beta-eating.

2.2 Alpha-eating

Next, we assume that asset returns are generated according to a characteristic-only world, but we note that the derivations equivalently hold for a mixed covariance-characteristic world. We assume that there may be (latent) factors that drive systematic return innovations, although they have factor premia of zero, and thus do not explain expected returns. Otherwise, we could not model (systematic) covariances between assets, which would be an unrealistic assumption since (systematic) covariations can be observed in real-world data. The true DGP is presented in equation (12):

$$r_{i,t+1} = \underbrace{\alpha_{i,t}}_{E_t\left[r_{i,t+1}\right]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{12}$$

Inserting equations (2) and (3) into (12), the resulting (true) DGP can be written as

$$r_{i,t+1} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\alpha} + \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1} \tag{13}$$

Denote $\boldsymbol{\Gamma}_{\beta}^{*}$, $\boldsymbol{\lambda}^{*}$, and $\widetilde{\boldsymbol{f}}_{t+1}^{*}$ as beta-only IPCA estimates of the characteristic coefficients for the betas, factor premia, and factor innovations, respectively, resulting from a misspecified IPCA setting. A misspecified beta-only IPCA can find an equivalent description of equation (13) without alpha:

$$r_{i,t+1} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}^{*}\boldsymbol{\lambda}^{*} + \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}^{*}\widetilde{\boldsymbol{f}}_{t+1}^{*} + \epsilon_{i,t+1} \tag{14}$$

if

$$\boldsymbol{\Gamma}_{\alpha} \widehat{=} \boldsymbol{\Gamma}_{\beta}^{*}\boldsymbol{\lambda}^{*} \tag{15}$$

Denote $\Gamma_{\alpha,l}$ as the true characteristic coefficient for the alpha of the l-th characteristic, $\Gamma_{\beta,l,j}^{*}$ as the estimated characteristic coefficient for the j-th beta of the l-th characteristic, and λ_{j}^{*} as the estimated factor premium of factor j. The linear system of equations to obtain a decomposition of

$\mathbf{\Gamma}_\alpha$ into $\mathbf{\Gamma}_\beta^*$ and $\boldsymbol{\lambda}^*$ can be written as follows:

$$\begin{array}{lcccccccc}\Gamma_{\alpha,1} & = & \Gamma^*_{\beta,1,1}\lambda^*_1 & + & \Gamma^*_{\beta,1,2}\lambda^*_2 & + & \dots & + & \Gamma^*_{\beta,1,K}\lambda^*_K \\ \Gamma_{\alpha,2} & = & \Gamma^*_{\beta,2,1}\lambda^*_1 & + & \Gamma^*_{\beta,2,2}\lambda^*_2 & + & \dots & + & \Gamma^*_{\beta,2,K}\lambda^*_K \\ & \vdots & & & & & & & \\ \Gamma_{\alpha,L} & = & \Gamma^*_{\beta,L,1}\lambda^*_1 & + & \Gamma^*_{\beta,L,2}\lambda^*_2 & + & \dots & + & \Gamma^*_{\beta,L,K}\lambda^*_K \end{array} \tag{16}$$

The system of linear equations is characterized by L equations, with L being the number of characteristics. Since we assume that latent factors drive systematic return innovations, the coefficients $\mathbf{\Gamma}_\beta^*$ are fixed and are therefore no free parameters. Accordingly, the only free parameters are the K values of the factor premia $\boldsymbol{\lambda}^*$, with K denoting the number of factors estimated by IPCA.

Depending on the (unknown) rank r^* of the true but unknown system of linear equations and the specification of the IPCA, we can differentiate three cases that produce different outcomes. First, if the rank of the system exceeds the number of estimated IPCA factors ($K < r^*$), the system is underdetermined; hence, no solution exists. Accordingly, IPCA cannot find a full beta-only description of expected returns. Second, if the number of IPCA factors equals the rank of the system ($K = r^*$), there is an exact solution. In this case, we can fully represent the characteristic-only world by an indistinguishable beta-only model. Third, if the number of factors exceeds the rank ($K > r^*$), then the system of equations is overdetermined, and therefore an infinite number of solutions exists. Thus, we expect an alpha-eating effect if a beta-only IPCA estimates at least $K = r^*$ factors, resulting in a covariance-only explanation of expected returns. Note, that a sufficient but not necessary condition for this is to set the number of factors equal to the number of characteristics, i.e., $K = L$. At least under this condition, we will observe the complete “alpha-eating” effect. In real-world applications, however, one might face a much lower rank, so that the alpha will already be completely “eaten up” even when fewer factors are estimated.

3 Simulation study

After analytically demonstrating the conditions under which either an alpha-eating or beta-eating effect can occur when estimating IPCA, we conduct a simulation study to demonstrate these effects. We describe the creation of the simulation data in Section 3.1, Section 3.2 briefly reviews the estimation of IPCA, and Section 3.3 presents the simulation results. We conclude the simulation study with a discussion of the results and their implications for empirical asset pricing research.

3.1 Simulation data

This section describes our simulation setup in greater detail. To achieve a simulation setting that resembles real-world applications, we simulate asset return data using empirical data for calibration purposes. However, we note that the calibration with respect to empirical data serves to work with realistic values and does not qualitatively affect the results (see the analytical derivation above, which holds for any arbitrary parameter choice). Specifically, we use the empirical U.S. equity dataset studied in KPS to calibrate our simulation parameters. We generate asset returns for $N = 4100$ assets and $T = 1200$ observations.[2] Additionally, we set the number of true factors to $K = 2$ and the number of simulated characteristics to $L = 10$.[3] We obtain some calibration parameters directly from the empirical data while others are obtained from estimating IPCA with two factors and alpha on the empirical data; however, the calibration process does not qualitatively affect our results.

The simulation of asset returns is divided into five steps, i.e., 1) simulation of characteristics $\boldsymbol{z}_{i,t}$, 2) simulation of the conditional alpha $\alpha_{i,t}$ and betas $\boldsymbol{\beta}_{i,t}$, 3) simulation of factor returns $\boldsymbol{f}_{t+1}$, 4) simulation of residual returns $\epsilon_{i,t+1}$, and 5) determination of asset (excess) returns $r_{i,t+1}$ by combining all components according to equation (4). To establish the easiest and comfortable simulation for IPCA, we keep the DGPs as simple as possible and do not assume, for example, (vector) autocorrelated residuals or time-varying factor premia.

3.1.1 Simulation of characteristics

We simulate $L = 10$ characteristics and assume for ease of interpretation that two of them are perfect proxies for the true betas on two simulated risk factors. This simplifies verifying whether

[2]These numbers are based on the empirical sample size of the U.S. equity market. According to Jensen et al. (2021), the average number of stocks in the U.S. is roughly 4,100 with a maximum history of approximately 1,200 months. In unreported results, we also simulated asset return data that matches the time series and cross section of other equity markets, i.e., China, Great Britain, and Spain. We choose these markets because of their different sizes of the time series and/or cross section. However, the results remain qualitatively unchanged.

[3]Based on a similar setting, Kelly et al. (2020) provide simulation evidence that IPCA's estimation errors are well-approximated with a normal distribution, which allows assuming normality for confidence intervals and hypothesis tests. However, they do not test the ability of the IPCA to distinguish between covariances and characteristics.

IPCA can replicate the true betas and thus increases interpretability of the simulation results. The characteristics are assumed to follow a vector autoregressive (VAR) process of order one according to equation (17):

$$\boldsymbol{z}_{i,t} = \boldsymbol{z}_{0,i} + \boldsymbol{\Phi}_{1,i}\boldsymbol{z}_{i,t-1} + \boldsymbol{u}_{i,t} \tag{17}$$

The term $\boldsymbol{z}_{0,i}$ is an $L \times 1$ vector of constants for each characteristic, $\boldsymbol{\Phi}_{1,i}$ is a $L \times L$ matrix of VAR(1) coefficients, and $\boldsymbol{u}_{i,t}$ is a $L \times 1$ vector of characteristic innovations. To parameterize the simulation, we assume that the vector of constants is zero and that the characteristics follow a near unit process determined by $\boldsymbol{\Phi}_{1,i}$, which is a diagonal matrix with all diagonal elements equal to 0.999 for all assets $i = 1, \ldots, N$. Characteristic innovations are drawn from a multivariate normal distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}_z$:

$$\boldsymbol{u}_{i,t} \sim N\left(0, \boldsymbol{\Sigma}_z\right) \tag{18}$$

The $L \times L$ covariance matrix $\boldsymbol{\Sigma}_z$ is obtained by $\left(\boldsymbol{s}\boldsymbol{s}'\right) \odot \boldsymbol{\Omega}$ with $\boldsymbol{s}$ denoting an $L \times 1$ vector of the standard deviations of the innovations and $\odot$ denoting element-wise multiplication. We assume that the innovations have unit variance, i.e., $\boldsymbol{s} = \mathbf{1}$. The term $\boldsymbol{\Omega}$ denotes an $L \times L$ correlation matrix that we can manipulate to obtain specific properties of our simulated dataset. In our simulation we assume that characteristics are uncorrelated; therefore, $\boldsymbol{\Omega}$ is set to an identity matrix. However, in unreported results we find that the assumption of correlated characteristics does not qualitatively affect the results.

After simulating characteristics according to equation (17), we rescale them to match the properties of the empirical characteristics. Specifically, we calibrate the means and standard deviations of the first eight simulated characteristics by obtaining the time-series averages of the cross-sectional means and standard deviations of empirical characteristics.[4] Because we assume that characteristics nine and ten are perfect proxies for the betas on two common factors, we calibrate these

[4]Specifically, we use those eight empirical characteristics for calibration purposes that have the largest estimated slopes (in absolute terms) obtained from univariate Fama-MacBeth regressions. The characteristics are (slopes in parentheses) bid-ask spread (0.0680), short-term reversal (-0.0547), idiosyncratic volatility with respect to the Fama-French 3-factor model (-0.0338), capital intensity (0.0258), return on assets (0.0099), ratio of change in property, plants, and equipment to the change in total assets (0.0098), price relative to its 52-week high (0.0091), and intermediate momentum (0.0066). We again note that the actual values of the simulation parameters do not qualitatively affect the results. These values serve to work with realistic values; however, one may choose completely different, even unrealistic values without obtaining qualitatively different results (see again our analytical derivation).

characteristics with respect to the estimated betas obtained from a two-factor alpha-beta IPCA on empirical data.[5] We obtain the time-series averages of the cross-sectional means and standard deviations of these (individual asset) betas and use them for calibration. Table 1 presents the calibration parameters used for our simulation. The column "Mu" refers to the mean and "Sigma" is the standard deviation of the characteristics.

	Mu	Sigma
$\boldsymbol{Z}_1$	0.0574	0.0280
$\boldsymbol{Z}_2$	0.0113	0.1762
$\boldsymbol{Z}_3$	0.0390	0.0217
$\boldsymbol{Z}_4$	0.0451	0.0120
$\boldsymbol{Z}_5$	-0.0309	0.0772
$\boldsymbol{Z}_6$	0.1064	0.1295
$\boldsymbol{Z}_7$	0.6894	0.1693
$\boldsymbol{Z}_8$	0.0577	0.3874
$\boldsymbol{Z}_9$	0.4916	0.1334
$\boldsymbol{Z}_{10}$	-0.5362	0.1032

Table 1: Target values for simulated characteristics
The table reports the target values used for calibration of the simulated characteristics. The column "Mu" refers to the means and "Sigma" are the average standard deviations of the characteristics. The term $\boldsymbol{Z}_l$ refers to the l-th characteristic ($l = 1, \ldots, 10$). The first eight characteristics are calibrated with respect to empirical characteristics and the last two characteristics are calibrated to the estimated IPCA betas obtained from a two-factor IPCA analysis on empirical U.S. data. All target values are the same for each simulated data set. In our initial setup, we assume that characteristics are mutually uncorrelated.

3.1.2 Simulation of alpha and beta

The specification of $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ is central to our simulation because it allows us to distinguish between a covariance-only, characteristic-only, or a covariance-characteristic world. We note that, even if only characteristics determine expected returns (i.e., a characteristic-only world), we assume that factor innovations explain systematic return innovations. Otherwise, we could not model a covariance structure among the assets. Accordingly, in the characteristic-only world, risk factors do not explain (long-term) expected returns but the (short-term) covariance structure of the asset returns.

[5] Note again that we could have also used a beta-only IPCA to calibrate the beta simulation or any other arbitrary choice without qualitatively affecting our results.

Table 2 shows the true $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ coefficients of all simulated worlds. Assuming a covariance-only worlds, all entries of $\boldsymbol{\Gamma}_\alpha$ are set to zero such that "anomaly returns" do not exist. For the characteristic-only and covariance-characteristic world, we assume that the $\boldsymbol{\Gamma}_\alpha$ coefficients of the first and second characteristic are nonzero, while those of the other characteristics are always zero. We obtain calibration parameters for the nonzero entries by running univariate Fama-MacBeth (FM) regressions (with an intercept) of the returns from the U.S. stocks on the firm characteristics. We calculate the average slopes and calibrate the $\boldsymbol{\Gamma}_\alpha$ coefficients on the first and second characteristic with respect to the two highest (in absolute terms) FM-coefficients, i.e., bid-ask spread (0.0680) and short-term reversal (-0.0547), respectively. Because we assume that characteristics nine and ten are perfectly correlated with the first and second beta, respectively, we set the corresponding $\boldsymbol{\Gamma}_\beta$ coefficients of these characteristics to one.[6] The remaining entries are set to zero. We note that the $\boldsymbol{\Gamma}_\beta$ coefficients for all DGPs are the same because in the characteristic-only world, we still allow betas to describe the covariance of the returns. However, the three worlds differ in the calculation of the factor returns, which we discuss next.

	Covariance-Only			Characteristic-Only			Covariance-Characteristic		
	$\boldsymbol{\Gamma}_\alpha$	$\boldsymbol{\Gamma}_{\beta,1}$	$\boldsymbol{\Gamma}_{\beta,2}$	$\boldsymbol{\Gamma}_\alpha$	$\boldsymbol{\Gamma}_{\beta,1}$	$\boldsymbol{\Gamma}_{\beta,2}$	$\boldsymbol{\Gamma}_\alpha$	$\boldsymbol{\Gamma}_{\beta,1}$	$\boldsymbol{\Gamma}_{\beta,2}$
$\boldsymbol{Z}_1$	0.0000	0.0000	0.0000	0.0680	0.0000	0.0000	0.0680	0.0000	0.0000
$\boldsymbol{Z}_2$	0.0000	0.0000	0.0000	-0.0547	0.0000	0.0000	-0.0547	0.0000	0.0000
$\boldsymbol{Z}_3$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_4$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_5$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_6$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_7$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_8$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\boldsymbol{Z}_9$	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000
$\boldsymbol{Z}_{10}$	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000

Table 2: Simulation values for $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$
The table reports the simulation values of the $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ coefficients for each simulated data set, i.e., covariance-only, characteristic-only, and covariance-characteristic world. $\boldsymbol{\Gamma}_\alpha$ denotes the characteristic coefficients of the anomaly intercept, $\boldsymbol{\Gamma}_{\beta,1}$ are the characteristic coefficients of the beta on the first factor, and $\boldsymbol{\Gamma}_{\beta,2}$ are the characteristic coefficients of the beta on the second factor.

[6] This specification of the true $\boldsymbol{\Gamma}_\beta$ coefficients is in line with the IPCA algorithm because IPCA estimates orthonormal $\boldsymbol{\Gamma}_\beta$ coefficients, which implies that the columns are scaled to unit length.

3.1.3 Simulation of factor returns

According to equation (4), factor returns can be divided into a (long-term) factor premium $\boldsymbol{\lambda}$ and (short-term) factor innovations $\widetilde{\boldsymbol{f}}_{t+1}$. Despite the assumption that a constant factor premia may be restrictive and oversimplified for modeling financial markets,[7] we do not assume time-varying factor premia. Otherwise, IPCA might not be able to distinguish between time-varying factor premia and stochastic factor innovations. We assume that the factor innovations have zero mean and do not follow a time-dependent process. Specifically, we simulate factor returns by drawing them from a multivariate normal distribution according to equation (19):

$$\boldsymbol{f}_{t+1} \sim N\left(\boldsymbol{\lambda}, \boldsymbol{\Sigma}_f\right) \tag{19}$$

with $\boldsymbol{f}_{t+1}$ denoting the $K \times 1$ vector of factor returns, and $\boldsymbol{\Sigma}_f$ representing the $K \times K$ covariance matrix of the factor innovations. In a characteristic-only world that does not allow for priced risk factors, factor means are set to zero (i.e., $\boldsymbol{\lambda} = \mathbf{0}$). Otherwise, we calibrate the factor means with respect to the estimated premia of the factors obtained from a two-factor IPCA analysis on the U.S. equity dataset studied in KPS. Specifically, we obtain the time-series averages of the estimated factors (i.e., 0.0331 and 0.0145) and use these as calibration values for the simulated true factor premia. We assume that the covariance matrix of factor innovations $\boldsymbol{\Sigma}_f$ is a diagonal matrix with diagonal elements calibrated with respect to the time series standard deviations of the estimated IPCA factors (i.e., 0.1031 and 0.0653). Assuming that $\boldsymbol{\Sigma}_f$ is diagonal implies that the factors are orthogonal. This assumption is in line with the IPCA algorithm because IPCA always extracts orthogonal factors. The target values for each simulated dataset are summarized in Table 3. Note that we consider factor innovations in the characteristic-only world but set the factor premia to zero.

[7] For example, Ilmanen et al. (2021) find time-varying risk premia for empirical factors that may be conditional on macroeconomic state variables. Daniel and Titman (1997) highlight the possibility that the actual premium of a factor may depend on past factor innovations, thus resulting in time-varying factor premia.

	Covariance-Only		Characteristic-Only		Covariance-Characteristic	
Factor	Mu	Sigma	Mu	Sigma	Mu	Sigma
f_1	0.0331	0.1031	0.0000	0.1031	0.0331	0.1031
f_2	0.0145	0.0653	0.0000	0.0653	0.0145	0.0653

Table 3: Target values for factor returns
The table reports the target values for the factor returns obtained from a two-factor IPCA with alpha estimated on the empirical U.S. equity data set. The column "Mu" refers to the average of the factors (i.e., factor premium) and "Sigma" refers to the standard deviation of the factor returns (i.e., standard deviation of factor innovations).

3.1.4 Simulation of Idiosyncratic Return Innovations

Idiosyncratic return innovations (i.e., residual returns) $\epsilon_{i,t+1}$ are drawn from a multivariate normal distribution according to equation (20):

$$\boldsymbol{\epsilon}_{t+1} \sim N\left(\boldsymbol{\mu}_{\epsilon}, \boldsymbol{\Sigma}_{\epsilon}\right) \tag{20}$$

with $\boldsymbol{\epsilon}_{t+1}$ denoting the $N \times 1$ vector of residual returns. The distribution of $\boldsymbol{\epsilon}_{t+1}$ is characterized by the $N \times 1$ vector of expected residual returns $\boldsymbol{\mu}_{\epsilon}$ and the $N \times N$ covariance matrix $\boldsymbol{\Sigma}_{\epsilon}$. For simplicity, we assume that all assets have the same expected residual return of zero, i.e., $\boldsymbol{\mu}_{\epsilon} = \mathbf{0}$. $\boldsymbol{\Sigma}_{\epsilon}$ is a diagonal matrix with diagonal elements σ_{ϵ} for all assets, i.e., the residuals are uncorrelated. We calibrate σ_{ϵ} with respect to the results obtained from our IPCA analysis on empirical data. That is, we obtain fitted (individual) asset returns from IPCA and calculate the cross sectional average of the standard deviations of the residuals. Thus, we set all diagonal elements of $\boldsymbol{\Sigma}_{\epsilon}$ to 0.1629.

After we have simulated characteristics, conditional alphas and betas, factor returns, and residual returns, we combine all components and determine the respective asset returns as described by equation (4).

3.2 Methodology

3.2.1 Estimation

This section briefly reviews the estimation procedure of the latent factors $\hat{\boldsymbol{f}}_{t+1}$ and the characteristic coefficients $\hat{\boldsymbol{\Gamma}}_\alpha$ and $\hat{\boldsymbol{\Gamma}}_\beta$, as discussed in KPS. Note that we differentiate between simulation parameters and (estimated) IPCA parameters by $\hat{(\cdot)}$. For each simulated dataset (i.e., the covariance-only, characteristic-only, and covariance-characteristic worlds), we fit 1) a beta-only IPCA that estimates only betas and no alpha (i.e., $\hat{\boldsymbol{\Gamma}}_\alpha = \boldsymbol{0}$, $\hat{\boldsymbol{\Gamma}}_\beta \neq \boldsymbol{0}$), 2) an alpha-only IPCA that estimates alpha but no betas (i.e., $\hat{\boldsymbol{\Gamma}}_\alpha \neq \boldsymbol{0}$, $\hat{\boldsymbol{\Gamma}}_\beta = \boldsymbol{0}$), and 3) an "unrestricted" IPCA that estimates both alpha and betas (i.e., $\hat{\boldsymbol{\Gamma}}_\alpha \neq \boldsymbol{0}$, $\hat{\boldsymbol{\Gamma}}_\beta \neq \boldsymbol{0}$). For specifications that estimate betas and thus factors, we follow KPS and estimate models with $K = 1, \ldots, 6$ factors. Since the IPCA estimation algorithm as proposed by KPS starts with a beta-only IPCA, we mainly discuss the procedure for estimating a beta-only IPCA here and highlight changes for an IPCA with alpha at relevant points.

IPCA finds estimates $\hat{\boldsymbol{f}}_{t+1}$ and $\hat{\boldsymbol{\Gamma}}_\beta$ according to a least squares criterion as described in equation (21):

$$\min_{\boldsymbol{\Gamma}_\beta, \boldsymbol{F}} \sum_{t=1}^{T-1} \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}_t \boldsymbol{\Gamma}_\beta \boldsymbol{f}_{t+1}\right)' \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}_t \boldsymbol{\Gamma}_\beta \boldsymbol{f}_{t+1}\right) \tag{21}$$

with $\boldsymbol{Z}_t$ denoting an $N \times L$ matrix of asset-level characteristics $\left(\boldsymbol{Z}_t = [\boldsymbol{z}_{1,t}, \ldots, \boldsymbol{z}_{N,t}]'\right)$, $\boldsymbol{f}_{t+1}$ is a $K \times 1$ vector of contemporaneous factor realizations, and $\boldsymbol{F}$ denotes the $K \times T$ matrix of stacked factor realizations. Since there is no closed-form solution to solve equation (21), KPS propose an alternating least squares algorithm to solve numerically for the model parameters. Given any arbitrary estimate for the latent factors, estimates for the $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients are obtained by a pooled ordinary least squares regression (POLS) of the asset returns on the asset characteristics interacted with the estimated factors:

$$\text{vec}\left(\hat{\boldsymbol{\Gamma}}'_\beta\right) = \left(\sum_{t=1}^{T-1} \boldsymbol{Z}'_t \boldsymbol{Z}_t \otimes \hat{\boldsymbol{f}}_{t+1} \hat{\boldsymbol{f}}'_{t+1}\right)^{-1} \left(\sum_{t=1}^{T-1} \left[\boldsymbol{Z}_t \otimes \hat{\boldsymbol{f}}_{t+1}\right]' \boldsymbol{r}_{t+1}\right) \tag{22}$$

with $\otimes$ denoting the Kronecker product. The estimates $\hat{\boldsymbol{\Gamma}}_\beta$ are arranged as an $(LK) \times 1$ vector; therefore, the vector has to be reshaped into an $L \times K$ matrix. The estimates for the characteristic coefficients in an IPCA that additionally allows for alpha are obtained equivalently to equation

(22) with the difference being that a constant factor is added to $\hat{\boldsymbol{f}}_{t+1}$. In this case, the estimates are arranged as an $(L(K+1)) \times 1$ vector and have to be reshaped into an $L \times 1$ vector and an $L \times K$ matrix of $\hat{\boldsymbol{\Gamma}}_\alpha$ and $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients, respectively.

Given the $\hat{\boldsymbol{\Gamma}}_\beta$ estimates, the factors are updated via periodical cross-sectional regressions of the asset returns on the conditional betas as described in equation (23):

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}_\beta' \boldsymbol{Z}_t' \boldsymbol{Z}_t \hat{\boldsymbol{\Gamma}}_\beta\right)^{-1} \left(\hat{\boldsymbol{\Gamma}}_\beta' \boldsymbol{Z}_t' \boldsymbol{r}_{t+1}\right) \forall t \tag{23}$$

In an IPCA that additionally allows for alpha, the factors are updated via cross-sectional regressions of the asset returns in excess of the anomaly returns $\boldsymbol{Z}_t \hat{\boldsymbol{\Gamma}}_\alpha$ on conditional betas:

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}_\beta' \boldsymbol{Z}_t' \boldsymbol{Z}_t \hat{\boldsymbol{\Gamma}}_\beta\right)^{-1} \left(\hat{\boldsymbol{\Gamma}}_\beta' \boldsymbol{Z}_t' \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}_t \hat{\boldsymbol{\Gamma}}_\alpha\right)\right) \forall t \tag{24}$$

KPS show that the entire IPCA algorithm can be written in terms of characteristic-managed portfolios instead of individual assets. This accelerates the computational performance and helps to circumvent the problem of missing values in the return and characteristic matrices of the individual assets. Characteristic-managed portfolios are the asset-level characteristics multiplied by the corresponding returns as described in equation (25):

$$\boldsymbol{x}_{t+1} = \frac{\boldsymbol{Z}_t' \boldsymbol{r}_{t+1}}{N_{t+1}} \tag{25}$$

with $\boldsymbol{x}_{t+1}$ denoting an $L \times 1$ vector of characteristic-managed portfolio returns and N_{t+1} denoting the number of nonmissing assets at $t+1$. The l-th element in $\boldsymbol{x}_{t+1}$ is the portfolio return at $t+1$, with asset weights determined by the l-th characteristic. Note that the characteristic-managed portfolios are created from those assets that have only nonmissing observations for any characteristic $\boldsymbol{Z}_t$. The product $\boldsymbol{Z}_t' \boldsymbol{Z}_t$ in equations (22) and (23) is invariant during numerical optimization, and therefore we can define:

$$\boldsymbol{W}_t = \frac{\boldsymbol{Z}_t' \boldsymbol{Z}_t}{N_{t+1}} \tag{26}$$

with $\boldsymbol{W}_t$ denoting an $L \times L$ matrix of cross-products of nonmissing characteristics. $\boldsymbol{W}_t$ replaces $\boldsymbol{Z}_t' \boldsymbol{Z}_t$ in equations (22) and (23), and the returns on characteristic-managed portfolios replace

$\boldsymbol{Z}_t'\boldsymbol{r}_{t+1}$ such that equations (22) and (23) can be rewritten as:

$$\text{vec}\left(\hat{\boldsymbol{\Gamma}}_\beta'\right) = \left(\sum_{t=1}^{T-1} N_{t+1}\left(\boldsymbol{W}_t \otimes \hat{\boldsymbol{f}}_{t+1}\hat{\boldsymbol{f}}_{t+1}'\right)\right)^{-1}\left(\sum_{t=1}^{T-1} N_{t+1}\left[\boldsymbol{x}_{t+1} \otimes \hat{\boldsymbol{f}}_{t+1}\right]\right) \tag{27}$$

and

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}_\beta'\boldsymbol{W}_t\hat{\boldsymbol{\Gamma}}_\beta\right)^{-1}\left(\hat{\boldsymbol{\Gamma}}_\beta\boldsymbol{x}_{t+1}\right)\forall t \tag{28}$$

Equivalently, we can rewrite equation (24):

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}_\beta'\boldsymbol{W}_t\hat{\boldsymbol{\Gamma}}_\beta\right)^{-1}\left(\hat{\boldsymbol{\Gamma}}_\beta\left(\boldsymbol{x}_{t+1} - \boldsymbol{W}_t\hat{\boldsymbol{\Gamma}}_\alpha\right)\right)\forall t \tag{29}$$

KPS propose a two-step estimation of the unrestricted alpha-beta IPCA. In the first step, a beta-only IPCA is estimated to obtain estimates for $\hat{\boldsymbol{\Gamma}}_\beta$. In the second step, the estimated $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients are used as initial values to estimate an alpha-beta IPCA. If there is a superior alpha-beta description of asset returns different from the beta-only world, IPCA will estimate a nonzero anomaly alpha. Thus, IPCA always begins by assuming a covariance-only world and will find this solution only if a suitable covariance-only representation exists, irrespective of whether it is true.

3.2.2 Validation

The performance evaluation is twofold. First, we compare the performance of the IPCA specifications on each simulated dataset by using the uncentered R^2s as defined in KPS. The R^2_{total} is a measure of the overall explanatory power for asset returns, i.e., it measures the fraction of return variation explained by conditional alphas, betas, and contemporaneous factor realizations:

$$R^2_{total} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - \left(\hat{\alpha}_{i,t} + \hat{\boldsymbol{f}}_{t+1}\hat{\boldsymbol{\beta}}_{i,t}\right)\right)^2}{\sum_{i,t} r_{i,t+1}^2} \tag{30}$$

with $\hat{\alpha}_{i,t}$ and $\hat{\boldsymbol{\beta}}_{i,t}$ denoting the estimates of the conditional alpha and the conditional betas as defined in equations (2) and (3), respectively. A second measure is the R^2_{pred} that quantifies the explanatory power for expected returns $E\left[r_{i,t+1}\right]$:

$$R^2_{pred} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - \left(\hat{\alpha}_{i,t} + \hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\beta}}_{i,t}\right)\right)^2}{\sum_{i,t} r^2_{i,t+1}} \tag{31}$$

with $\hat{\boldsymbol{\lambda}}$ denoting an estimate for the factor premia (i.e., the time-series average of the factors).

Second, to test whether there is a significant anomaly intercept given a set of K common factors, we use the statistical test proposed by KPS for testing the null hypothesis that all elements of $\hat{\boldsymbol{\Gamma}}_\alpha$ equal zero, i.e., alpha does not depend on characteristics. Thus, the null hypothesis assumes that asset returns are generated according to a covariance-only world and can be formalized as:

$$H_0 : \boldsymbol{\Gamma}_\alpha = \mathbf{0}$$

The alternative hypothesis is:

$$H_1 : \boldsymbol{\Gamma}_\alpha \neq \mathbf{0}$$

which implies that the characteristics themselves drive asset returns, i.e., asset returns are generated according to either a characteristic-only or a covariance-characteristic world.

KPS first estimate the alpha-beta (i.e., unrestricted) IPCA to obtain estimates for the $\hat{\boldsymbol{\Gamma}}_\alpha$ vector and calculate the test statistic as the sum of squared elements in $\hat{\boldsymbol{\Gamma}}_\alpha$:

$$W_\alpha = \hat{\boldsymbol{\Gamma}}'_\alpha \hat{\boldsymbol{\Gamma}}_\alpha \tag{32}$$

They obtain a distribution of W_α under the null hypothesis using a "residual bootstrap" procedure. Their residual bootstrap is based on resampling characteristic-managed portfolio returns as described in equation (33):

$$\boldsymbol{x}^b_{t+1} = \boldsymbol{W}_t \hat{\boldsymbol{\Gamma}}_\beta \hat{\boldsymbol{f}}_{t+1} + \widetilde{\boldsymbol{d}}_{t+1} \tag{33}$$

with $\boldsymbol{x}^b_{t+1}$ denoting an $L \times 1$ vector of resampled characteristic-managed portfolio returns of bootstrap $b = 1, \ldots, 1000$. The first part of equation (33) are the fitted returns obtained from a beta-only IPCA. The second term $\widetilde{\boldsymbol{d}}_{t+1}$ is an $L \times 1$ vector of residuals obtained from the alpha-beta IPCA, multiplied with a Student t random variable with unit variance and five degrees of freedom.[8] Using the bootstrapped samples, they re-estimate the alpha-beta IPCA and calculate a bootstrapped test statistic W_α^b as described in equation (32). The p-values are calculated as the fraction of bootstrapped W_α^b statistics that exceed the value of W_α from the original dataset. They reject the null hypothesis if the p-value is below the 1 percent significance level; otherwise, they accept the null hypothesis.

3.3 Simulation results

This section reports the results for the mixed covariance-characteristic world because this represents the alternative hypothesis of the statistical test proposed by KPS and is capable of demonstrating both the alpha- and beta-eating effects. Although, considering the extreme cases, i.e., the covariance-only and characteristic-only worlds, allows for a more isolated analysis of the beta-eating and alpha-eating effect, respectively, we save space and do not discuss these results in detail here because no further insights are gained by considering these two cases. For the interested reader, we refer to Internet Appendices A and B, respectively. The true DGP in the covariance-characteristic world can be written as follows:

$$r_{i,t+1} = \underbrace{\alpha_{i,t} + \boldsymbol{\beta}_{i,t}\boldsymbol{\lambda}}_{E_t\left[r_{i,t+1}\right]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{34}$$

where the true factor premia are $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$ and the corresponding $\boldsymbol{\Gamma}_\beta$ coefficients of $\boldsymbol{Z}_9$ and $\boldsymbol{Z}_{10}$ are equal to one. Additionally, we calibrate the true $\boldsymbol{\Gamma}_\alpha$ coefficients of $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ to be 0.0680 and -0.0547. According to this DGP, the asset pricing test should identify the two-factor alpha-beta IPCA as the true model and IPCA should return appropriate estimates of $\boldsymbol{\lambda}$, $\boldsymbol{\Gamma}_\beta$ and $\boldsymbol{\Gamma}_\alpha$. Note that in reality, the true DGP is unknown, and a researcher will apply different specifications of the IPCA model without knowing which is the true one.

Panel A and Panel B of Table 4 report the performance evaluation metrics R^2_{total} and R^2_{pred} as defined in equations (30) and (31) for individual asset returns r_t and characteristic-managed portfolio returns x_t, respectively. When estimating an alpha-beta IPCA with the true but in reality

[8] Multiplying the residuals by a random t variable improves the efficiency of bootstrapping in heteroskedastic data (Gonçalves and Kilian, 2004; Kelly et al., 2019).

unknown number of factors (i.e., $K = 2$), the IPCA achieves an R^2_{total} of 14.3 (100.0) percent and an R^2_{pred} of 0.83 (3.51) percent for individual asset (characteristic-managed portfolio) returns. The results in Panels A and B indicate a "beta-eating" effect when estimating too few factors. Dropping one factor from the alpha-beta IPCA, the R^2_{total} reduces to 14.0 (99.9) percent, indicating that the one-factor alpha-beta IPCA cannot fully describe the covariance structure among the asset returns. However, the R^2_{pred}s are not distinguishable from those of the true model. Similarly, removing the remaining factor from the alpha-beta IPCA and estimating an alpha-only model, the R^2_{pred}s remain at the same level as for the true model. Using the Clark and West (2007) test (unreported), we find that the two-factor alpha-beta model neither statistically significantly outperforms the alpha-only (p-value = 50.0 percent) nor the one-factor alpha-beta (p-value = 48.9 percent) model. As analytically derived in Section 2.1, the conditional alpha compensates for the omitted factor(s), and hence, both specifications can describe expected returns $E_t[r_{t+1}]$ exactly as well as the true model, although they cannot capture the entire contemporaneous covariation among the assets. This holds equally if no alpha is included in the true model and asset returns are generated according to a covariance-only world, as reported in Internet Appendix A. Taking into consideration that such a world does not assume an anomaly intercept, it highlights the "beta-eating" effect in an even more isolated way.

The results on the covariance-characteristic world also reveal an "alpha-eating" effect. Not surprisingly, the one- and two-factor beta-only IPCA models cannot fully describe expected asset returns because they are misspecified. However, when adding a third factor to the model, both the R^2_{total} and R^2_{pred} are equal to those of the true model, indicating that IPCA compensates for the omitted intercept with an additional factor. As the R^2_{pred} is indistinguishable from that of the true model (Clark and West (2007) test p-value = 39.2 percent), IPCA provides at least four models that have the same performance in describing expected returns.

Panel C reports the sum of squared $\hat{\mathbf{\Gamma}}_\alpha$ coefficients, denoted by W_α, and the corresponding bootstrapped p-values of the asset pricing test for testing the null hypothesis of an anomaly intercept equal to zero. Following the asset pricing test, the null hypothesis of a zero conditional alpha beyond betas is correctly rejected when estimating one or two factors. However, the "alpha-eating" effect is present not only in the beta-only IPCA, but also in the alpha-beta IPCA, as indicated by the declining sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$ with an increasing number of factors. When estimating at least three factors, the null hypothesis cannot be rejected at any conventional significance level (p-value = 29.1 percent). Accordingly, the test concludes that asset returns are driven by the characteristics themselves when estimating up to two factors, but characteristics are covariances when estimating at least three risk factors. In real-world applications, a researcher does not know which test result to trust because the performance of the two-factor alpha-beta and three-factor beta-only models is indistinguishable. In Internet Appendix B, we demonstrate the alpha-eating effect in a more isolated way assuming a characteristic-only world. The results hold equally for a world without priced risk factors, as already shown in Section 2.2.

IPCA	Beta						Alpha-Beta						Alpha
K	1	2	3	4	5	6	1	2	3	4	5	6	0
Panel A: Individual assets (r_t)													
R^2_{total}	13.6	14.1	14.3	14.3	14.3	14.4	14.0	14.3	14.3	14.3	14.3	14.4	0.83
R^2_{pred}	0.51	0.75	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83
Panel B: Managed portfolios (x_t)													
R^2_{total}	99.8	99.9	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	3.51
R^2_{pred}	3.43	3.49	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51
Panel C: Asset pricing test													
W_α							0.81	0.71	0.61	0.32	0.32	0.17	
p-value							0.00	0.00	29.1	52.5	15.0	49.9	

Table 4: IPCA model performance (covariance-characteristic world)
Panel A and B report the performance evaluation metrics R^2_{total} and R^2_{pred} for individual assets (r_t) and characteristic-managed portfolios (x_t), respectively, using a simulated covariance-characteristic data set. Panel C reports the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$ (W_α) (multiplied by 100 for presentation purposes) and bootstrapped p-values in percent for the test of $\hat{\mathbf{\Gamma}}_\alpha = \mathbf{0}$. The column "Beta" refers to an IPCA without alpha, "Alpha-Beta" refers to an IPCA with both alpha and beta, and "Alpha" is an IPCA that only estimates alpha and no betas.

To gain additional insights regarding the estimated parameters, Table 6 reports the estimated factor premia for the beta-only and alpha-beta IPCAs, respectively. The estimated premia for the two-factor alpha-beta IPCA are 0.0348 and 0.0157 and therefore close to those assumed in the true DGP (i.e., $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$). However, the two-factor beta-only IPCA considerably overestimates the factor premia ($\hat{\lambda}_1 = 0.0768$ and $\hat{\lambda}_2 = 0.0294$), indicating that the "alpha-eating" effect already occurs when two factors are estimated. This is also reflected in the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients reported in Table 7. According to the true DGP, we expect coefficients on characteristics one through eight to be zero for any beta and coefficients on characteristics nine and ten equal to one. However, in the beta-only IPCA specification, the coefficients on the first and second characteristics are nonzero and those on the ninth and tenth characteristics are not equal to one, suggesting that the beta-only IPCA cannot model the true betas because the latter subsume some explanatory power of the alpha. The picture becomes similar for the alpha-beta IPCA when estimating at least three factors. For example, the coefficient of the second characteristic to the third beta $\hat{\Gamma}_{\beta,2,3}$ is -0.45. This effect is related to the fact that in the IPCA estimation the beta-only model forms the initial solution for the estimation of the alpha-beta model.

Table 5 reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only IPCA specifications for the covariance-characteristic world. From Table 4 we find that the alpha-only IPCA achieves the same explanatory power for expected returns as the true two-factor alpha-beta IPCA. The alpha-only IPCA estimates for the true $\mathbf{\Gamma}_\alpha$ coefficients on the first and second characteristics (i.e., 0.0643 and -0.0549) are very close to the true values reported in Table 2. The coefficients for characteristics three to eight are not distinguishable from zero, which is also in line with the true DGP. However, the coefficients on characteristics nine and ten are 0.0334 and 0.0138, respectively, and therefore very close to the true factor premia of $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$. These results reflect the beta-eating effect as discussed in Section 2.1. There, we show that an alpha-only IPCA can find a full characteristic-only description of a true covariance-only world if the l-th $\hat{\mathbf{\Gamma}}_\alpha$ element is chosen to be equal to the product of the l-th row of $\hat{\mathbf{\Gamma}}_\beta$ and the true factor premia. Since the true $\mathbf{\Gamma}_\beta$ coefficients of the ninth and tenth characteristics are set to one, IPCA estimates $\hat{\mathbf{\Gamma}}_\alpha$ coefficients that simply equal the true factor premia. The $\hat{\Gamma}_{\alpha,1}$ and $\hat{\Gamma}_{\alpha,2}$ coefficients are also close to the true values when estimating a one- or two-factor alpha-beta IPCA. However, when estimating a third factor, the coefficients begin to decline, while we document increasing $\hat{\mathbf{\Gamma}}_\beta$ coefficients on those characteristics in Table 7. This, again, reflects the "alpha-eating" effect as discussed above.

	K					
	1	2	3	4	5	6
Panel A: Beta-Only IPCA						
$\hat{\lambda}_1$	**0.0192**	**0.0768**	**0.0702**	**0.0811**	**0.0826**	**0.0844**
$\hat{\lambda}_2$		**0.0294**	**0.0557**	0.0008	0.0000	0.0018
$\hat{\lambda}_3$			**0.0221**	**0.0267**	**0.0280**	**0.0262**
$\hat{\lambda}_4$				**0.0348**	0.0006	0.0017
$\hat{\lambda}_5$					**0.0300**	**0.0061**
$\hat{\lambda}_6$						**0.0266**
Panel B: Alpha-Beta IPCA						
$\hat{\lambda}_1$	**0.0226**	**0.0348**	**0.0382**	**0.0661**	**0.0681**	**0.0747**
$\hat{\lambda}_2$		**0.0157**	**0.0197**	0.0031	0.0027	0.0015
$\hat{\lambda}_3$			**0.0250**	**0.0163**	**0.0230**	**0.0265**
$\hat{\lambda}_4$				**0.0271**	0.0006	0.0002
$\hat{\lambda}_5$					**0.0154**	**0.0243**
$\hat{\lambda}_6$						**0.0043**

Table 6: Estimated factor premia (covariance-characteristic world)
The table reports the estimated factor premia $\hat{\lambda}_j$, calculated as the time series average of each factor. If a factor premium significantly differs from zero at the 1 percent level, we print the premium in bold letters.

IPCA	Beta						Alpha-Beta					
K	1	2	3	4	5	6	1	2	3	4	5	6
$\hat{\Gamma}_{\beta,1,1}$	0.09	0.39	0.35	0.69	0.72	0.77	-0.04	0.00	0.01	0.57	0.61	0.68
$\hat{\Gamma}_{\beta,2,1}$	-0.12	-0.33	-0.30	-0.29	-0.29	-0.28	-0.01	-0.01	-0.06	-0.08	-0.08	-0.15
$\hat{\Gamma}_{\beta,3,1}$	-0.01	0.00	0.00	-0.11	-0.06	-0.13	-0.02	-0.01	0.08	-0.03	0.03	0.01
$\hat{\Gamma}_{\beta,4,1}$	-0.04	-0.01	-0.01	-0.35	-0.32	-0.24	-0.03	0.02	0.03	-0.51	-0.48	-0.38
$\hat{\Gamma}_{\beta,5,1}$	0.01	0.00	0.00	0.01	-0.01	0.00	0.01	0.01	-0.01	-0.01	-0.04	0.05
$\hat{\Gamma}_{\beta,6,1}$	0.00	-0.00	-0.00	0.03	0.03	0.01	0.00	-0.00	0.01	0.08	0.07	0.13
$\hat{\Gamma}_{\beta,7,1}$	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.01	-0.01	-0.02	0.05
$\hat{\Gamma}_{\beta,8,1}$	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01	-0.00	0.00	-0.00	-0.03	-0.03	-0.03
$\hat{\Gamma}_{\beta,9,1}$	0.92	0.85	0.88	0.54	0.53	0.48	0.91	1.00	0.99	0.61	0.60	0.58
$\hat{\Gamma}_{\beta,10,1}$	-0.35	0.11	0.13	0.12	0.12	0.13	-0.41	0.09	0.09	0.13	0.14	0.13
$\hat{\Gamma}_{\beta,1,2}$		0.40	0.60	0.41	0.40	-0.40		0.09	0.10	0.43	0.43	-0.36
$\hat{\Gamma}_{\beta,2,2}$		-0.27	-0.47	0.00	0.01	-0.02		0.00	-0.07	0.02	0.02	-0.03
$\hat{\Gamma}_{\beta,3,2}$		0.03	0.03	-0.13	-0.16	0.19		0.02	0.15	-0.14	-0.16	0.24
$\hat{\Gamma}_{\beta,4,2}$		0.05	-0.01	-0.39	-0.40	0.30		0.09	0.10	-0.43	-0.43	0.49
$\hat{\Gamma}_{\beta,5,2}$		-0.01	-0.02	0.00	0.01	-0.02		0.00	-0.03	0.00	0.01	-0.02
$\hat{\Gamma}_{\beta,6,2}$		-0.01	-0.00	0.04	0.05	-0.03		-0.01	0.01	0.05	0.05	-0.05
$\hat{\Gamma}_{\beta,7,2}$		-0.00	-0.00	0.00	0.00	-0.01		-0.00	-0.02	0.00	0.00	0.00
$\hat{\Gamma}_{\beta,8,2}$		-0.00	-0.00	-0.02	-0.02	0.01		0.00	-0.01	-0.02	-0.02	0.02
$\hat{\Gamma}_{\beta,9,2}$		-0.39	-0.47	-0.81	-0.81	0.84		-0.09	-0.11	-0.78	-0.78	0.75
$\hat{\Gamma}_{\beta,10,2}$		0.78	0.44	0.01	0.00	-0.01		0.99	0.97	0.01	0.01	0.03
$\hat{\Gamma}_{\beta,1,3}$			0.30	0.08	0.28	0.28			0.08	0.03	0.29	0.42
$\hat{\Gamma}_{\beta,2,3}$			-0.32	-0.30	-0.23	-0.21			-0.45	-0.08	-0.09	-0.11
$\hat{\Gamma}_{\beta,3,3}$			-0.00	0.15	0.64	0.50			0.83	0.19	0.74	0.63
$\hat{\Gamma}_{\beta,4,3}$			-0.11	0.39	0.47	0.64			0.07	0.17	0.33	0.44
$\hat{\Gamma}_{\beta,5,3}$			-0.01	-0.02	-0.21	-0.16			-0.18	-0.03	-0.24	-0.12
$\hat{\Gamma}_{\beta,6,3}$			0.01	-0.04	-0.10	-0.13			0.12	0.00	-0.06	0.01
$\hat{\Gamma}_{\beta,7,3}$			0.00	-0.01	-0.05	-0.04			-0.08	-0.02	-0.07	0.04
$\hat{\Gamma}_{\beta,8,3}$			-0.00	0.01	0.01	0.02			-0.05	-0.01	-0.01	0.01
$\hat{\Gamma}_{\beta,9,3}$			-0.11	-0.17	-0.23	-0.22			-0.08	-0.10	-0.17	-0.31
$\hat{\Gamma}_{\beta,10,3}$			-0.88	0.84	0.35	0.36			-0.20	0.96	0.39	0.32
$\hat{\Gamma}_{\beta,1,4}$				0.18	0.18	0.06				0.11	0.17	0.17
$\hat{\Gamma}_{\beta,2,4}$				-0.67	-0.01	-0.03				-0.46	-0.02	-0.02
$\hat{\Gamma}_{\beta,3,4}$				0.21	0.36	0.74				0.81	0.35	0.34
$\hat{\Gamma}_{\beta,4,4}$				0.42	0.00	-0.41				0.01	0.00	-0.00
$\hat{\Gamma}_{\beta,5,4}$				-0.04	-0.15	-0.26				-0.18	-0.14	-0.13
$\hat{\Gamma}_{\beta,6,4}$				-0.04	-0.03	0.06				0.14	-0.02	-0.02
$\hat{\Gamma}_{\beta,7,4}$				-0.00	-0.03	-0.06				-0.07	-0.03	-0.03
$\hat{\Gamma}_{\beta,8,4}$				0.02	-0.00	-0.04				-0.05	-0.01	-0.00
$\hat{\Gamma}_{\beta,9,4}$				-0.16	0.01	-0.00				-0.09	0.01	0.01
$\hat{\Gamma}_{\beta,10,4}$				-0.52	-0.90	-0.45				-0.23	-0.91	-0.92
$\hat{\Gamma}_{\beta,1,5}$					-0.08	0.23					-0.38	-0.04
$\hat{\Gamma}_{\beta,2,5}$					-0.81	-0.09					-0.73	-0.62
$\hat{\Gamma}_{\beta,3,5}$					-0.37	-0.22					0.25	-0.15
$\hat{\Gamma}_{\beta,4,5}$					0.33	0.47					-0.35	0.23
$\hat{\Gamma}_{\beta,5,5}$					0.19	0.05					0.10	0.47
$\hat{\Gamma}_{\beta,6,5}$					0.02	-0.11					0.34	0.45
$\hat{\Gamma}_{\beta,7,5}$					0.06	0.02					-0.02	0.31
$\hat{\Gamma}_{\beta,8,5}$					0.03	0.04					-0.08	-0.00
$\hat{\Gamma}_{\beta,9,5}$					-0.13	-0.03					-0.05	-0.09
$\hat{\Gamma}_{\beta,10,5}$					-0.19	-0.81					0.02	-0.14
$\hat{\Gamma}_{\beta,1,6}$						-0.30						-0.40
$\hat{\Gamma}_{\beta,2,6}$						-0.92						-0.49
$\hat{\Gamma}_{\beta,3,6}$						-0.05						0.44
$\hat{\Gamma}_{\beta,4,6}$						-0.10						-0.52
$\hat{\Gamma}_{\beta,5,6}$						0.12						-0.21
$\hat{\Gamma}_{\beta,6,6}$						0.12						0.11
$\hat{\Gamma}_{\beta,7,6}$						0.04						-0.23
$\hat{\Gamma}_{\beta,8,6}$						-0.01						-0.09
$\hat{\Gamma}_{\beta,9,6}$						-0.11						-0.02
$\hat{\Gamma}_{\beta,10,6}$						-0.03						0.13

Table 7: $\hat{\mathbf{\Gamma}}_\beta$ coefficients (covariance-characteristic world)
The table reports the estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients for the beta-only and alpha-beta (i.e., unrestricted) IPCA, respectively, using a simulated covariance-characteristic data set. The term $\hat{\Gamma}_{\beta,l,j}$ denotes the estimated coefficient of the l-th characteristic for the beta on the j-th factor.

IPCA	Alpha-Beta						Alpha
K	1	2	3	4	5	6	0
$\hat{\Gamma}_{\alpha,1}$	0.0665	0.0641	0.0609	0.0225	0.0217	0.0067	0.0643
$\hat{\Gamma}_{\alpha,2}$	-0.0546	-0.0546	-0.0400	-0.0355	-0.0359	-0.0236	-0.0549
$\hat{\Gamma}_{\alpha,3}$	0.0018	0.0013	-0.0258	-0.0217	-0.0215	-0.0151	0.0010
$\hat{\Gamma}_{\alpha,4}$	-0.0031	-0.0061	-0.0088	0.0282	0.0281	0.0089	-0.0033
$\hat{\Gamma}_{\alpha,5}$	-0.0012	-0.0014	0.0046	0.0055	0.0058	-0.0116	-0.0008
$\hat{\Gamma}_{\alpha,6}$	-0.0002	0.0003	-0.0037	-0.0094	-0.0089	-0.0216	0.0001
$\hat{\Gamma}_{\alpha,7}$	-0.0002	-0.0001	0.0024	0.0029	0.0030	-0.0119	-0.0003
$\hat{\Gamma}_{\alpha,8}$	-0.0001	-0.0001	0.0014	0.0036	0.0035	0.0020	-0.0001
$\hat{\Gamma}_{\alpha,9}$	0.0127	-0.0000	-0.0003	-0.0007	-0.0008	-0.0003	0.0334
$\hat{\Gamma}_{\alpha,10}$	0.0231	-0.0049	-0.0039	-0.0043	-0.0042	-0.0014	0.0138

Table 5: $\hat{\mathbf{\Gamma}}_\alpha$ coefficients (covariance-characteristic world)
The table reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta (i.e., unrestricted) and alpha-only IPCA, using a simulated covariance-characteristic data set. The term $\hat{\Gamma}_{\alpha,l}$ denotes the estimated coefficient of the l-th characteristic for the alpha.

3.4 Discussion

Both analytical evidence and simulation results suggest that IPCA cannot reliably distinguish between covariances and characteristics. A researcher who applies IPCA faces a joint hypothesis problem, since one must determine the correct number of factors and whether there is an alpha. When adding more factors to the model, IPCA will always create a covariance-based explanation of asset returns for which the asset pricing test cannot reject the null hypothesis of zero alpha. Therefore, the conclusion drawn by KPS, Büchner and Kelly (2022), and Kelly et al. (2021) that only covariances explain stock, option, or bond returns may be subject to the alpha-eating effect and therefore inappropriate.[9]

From our simulation, we find that the conditional alpha is always sufficient for modeling expected returns, irrespective of whether the alpha truly drives asset returns. What is striking about the results in KPS is that the R^2_{pred} is always higher for the unrestricted IPCA. In Internet Appendix C, we extend Table 1 in KPS by allowing for an alpha-only IPCA. We find that the alpha-only IPCA achieves the highest R^2_{pred} among all other IPCA specifications presented in KPS, i.e., 0.77

[9]The joint hypothesis problem cannot even be solved by considering information criteria, which are commonly used in the literature for model selection. Information criteria account for both the statistical model performance and its complexity. In unreported results, we find that information criteria, e.g., the Akaike Information Criterion (AIC), prefer models with alpha, indicating that the model selection based on information criteria is biased by the beta-eating effect.

percent for individual stock returns and 2.67 percent for portfolio returns. Thus, the effect that alpha is sufficient for modeling expected returns is not only a phenomenon from our simulation but can also be observed in empirical data.

There is some indication that the empirical results presented in KPS may be biased toward either the beta-eating or alpha-eating effect. In the simulation, we also find that as soon as the alpha-eating effect occurs, there are large jumps in the R^2_{pred}s for the beta-only IPCA that close the gap between the R^2_{pred}s of the beta-only and alpha-beta IPCA. For example, in the covariance-characteristic world (Table 4), the R^2_{pred}s are 0.51 percent for the one-factor, 0.75 percent for the two-factor, and 0.83 percent for the three-factor models, respectively. Such jumps can also be observed in the results presented in KPS. When increasing the number of factors from four to five, the R^2_{pred} increases from 0.41 to 0.69 percent for individual assets. For characteristic-managed portfolio returns, there is also a large difference between the four- and five-factor specifications, i.e., 2.13 percent for the four-factor and 2.41 percent for the five-factor model.[10] From our simulation, we can provide two possible explanations for these jumps. First, the alpha-beta (or alpha-only) IPCA has superior explanatory power for returns in low-dimensional factor models because the alpha "eats" the beta. Thus, estimating the fifth factor eliminates the beta-eating effect, and IPCA identifies the true covariance-only model. Second, a low-factor characteristic-only or covariance-characteristic model is the true model, and five factors in KPS are sufficient to capture the entire explanatory power of the omitted alpha (i.e., alpha-eating explanation). However, both the alpha-eating and the beta-eating effect make it impossible for IPCA to distinguish between covariances and characteristics; therefore, the question of whether covariances, characteristics, or both drive asset returns remains unanswered. Although IPCA cannot distinguish between characteristics and covariances, IPCA always finds a model with explanatory power for returns that is equivalent to that of the true model, making it a powerful tool for return estimation and asset allocation.

4 Conclusion

KPS propose IPCA, a new factor modeling approach, that finds latent risk factors and estimates factor loadings as a function of observable asset characteristics. By achieving a dimensionality

[10] In Büchner and Kelly (2022), such a jump can also be observed for individual option returns, when increasing the number of factors from three ($R^2_{pred} = 5.85$) to four ($R^2_{pred} = 6.19$).

reduction from a potentially large number of asset characteristics to a smaller number of conditional factor loadings, IPCA allows modeling returns as a function of many characteristics while being parsimoniously parameterized. There is a growing strand of literature suggesting that IPCA does not only outperform previously studied factor models in modeling stock returns (Kelly et al., 2019; Windmüller, 2021), but also contributes to other asset classes. Büchner and Kelly (2022) and Goyal and Saretto (2022) document IPCA's superiority in explaining delta-hedged option returns and Kelly et al. (2021) and Bianchi and Babiak (2022) extend this evidence to corporate bond and cryptocurrency returns, respectively. However, KPS claim that IPCA not only has superior explanatory power for returns, but can also distinguish between covariances and characteristics. By applying IPCA to stock return data, they conclude that returns are determined by covariances with common risk factors and no "anomalous returns" driven by the characteristics themselves exist. These findings are a response to the ongoing "covariance vs. characteristic" debate that has been extensively discussed in the literature over the past thirty years.

By presenting analytical evidence, we show that IPCA cannot reliably distinguish between covariances and characteristics because the results may be subject to a beta-eating or alpha-eating effect. We find supporting evidence for this using simulation data. Specifically, we find that a researcher who applies IPCA faces a joint hypothesis problem, as the question of whether returns are driven by covariances or characteristics cannot be answered without addressing the question of how many factors have to be estimated. We find that a correctly specified IPCA, that is, an IPCA that estimates the correct number of factors and estimates either alpha or betas or both, can identify the true model parameters. As a result, the asset pricing test proposed by KPS can correctly differentiate between covariances and characteristics as long as the true model is known in advance. However, the true number is not known in reality. If the IPCA is misspecified, then it will find a model that has the same performance as the true model, but the interpretation becomes more difficult or potentially incorrect inferences are drawn. Omitting factors and allowing for an anomaly alpha results in a beta-eating effect, meaning that alpha subsumes the explanatory power of returns that is actually attributable to betas. For example, an alpha-only IPCA identifies the true factor premia and the true anomaly returns associated with the respective characteristics but cannot distinguish between covariances or characteristics because it treats both as alpha. If IPCA estimates more factors than exist in the true model and there are anomaly returns in the true model, there is an alpha-eating effect, i.e., beta subsumes the returns associated with conditional alphas. This in turn makes anomaly returns resemble factor premia, resulting in the rejection of the characteristic story. KPS do not address the question of how many factors are necessary to model asset prices. Instead, they add factors to the model until the covariance-story is true, although the performance evaluation shows that the alpha-beta IPCA has comparable or even superior model performance. This simulation study shows that one can always generate a model that fits the covariance story that also has similar performance as the possibly true characteristic model.

References

Bianchi, D. and Babiak, M. (2022). A factor model for cryptocurrency returns. *SSRN Electronic Journal.*

Büchner, M. and Kelly, B. (2022). A factor model for option returns. *Journal of Financial Economics*, 143(3):1140–1161.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108.

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

Daniel, K., Titman, S., and Wei, K. J. (2001). Explaining the cross-section of stock returns in japan: Factors or characteristics? *Journal of Finance*, 56(2):743–766.

Davis, J. L., Fama, E. F., and French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *Journal of Finance*, 55(1):389–406.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33(5):2326–2377.

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

Goyal, A. and Saretto, A. (2022). Are equity option returns abnormal? ipca says no. *Working Paper*.

Green, J., Hand, J. R. M., and Zhang, X. F. (2017). The characteristics that provide independent information about average u.s. monthly stock returns. *Review of Financial Studies*, 30(12):4389–4436.

Ilmanen, A. S., Israel, R., Moskowitz, T. J., Thapar, A. K., and Wang, F. (2021). Factor premia and factor timing: A century of evidence. *SSRN Electronic Journal*.

Jensen, T. I., Kelly, B. T., and Pedersen, L. H. (2021). Is there a replication crisis in finance? *SSRN Electronic Journal*.

Kelly, B., Pruitt, S., and Su, Y. (2020). Instrumented principal component analysis. *Working Paper*.

Kelly, B. T., Palhares, D., and Pruitt, S. (2021). Modeling corporate bond returns. *SSRN Electronic Journal*.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Windmüller, S. (2021). Firm Characteristics and Global Stock Returns: A Conditional Asset Pricing Model. *The Review of Asset Pricing Studies*, 12(2):447–499.

Internet Appendices: Are Characteristics Covariances? A Comment on Instrumented Principal Component Analysis

A Covariance-only world

In our simulation presented in Section 3.3, we generated data according to a mixed covariance-characteristic world that represents the alternative hypothesis of the asset pricing test proposed by KPS. In this section, we analyze the beta-eating effect in greater detail by generating data according to a covariance-only world. Specifically, we assume that asset returns are fully described by covariances with (two) priced risk factors and no anomaly alpha exists. The data-generating process can be written as follows:

$$r_{i,t+1} = \underbrace{\boldsymbol{\beta}_{i,t}\boldsymbol{\lambda}}_{E_t\left[r_{i,t+1}\right]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{35}$$

As shown in Table 3, the true factor premia are $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$, and the corresponding $\boldsymbol{\Gamma}_\beta$ coefficients of $\boldsymbol{Z}_9$ and $\boldsymbol{Z}_{10}$ equal 1. All $\boldsymbol{\Gamma}_\alpha$ coefficients are zero. Accordingly, the asset pricing test should identify the beta-only IPCA as the true model and IPCA should return appropriate estimates of the factor premia, $\boldsymbol{\Gamma}_\alpha$, and $\boldsymbol{\Gamma}_\beta$ coefficients.

In Table 8 we report the performance evaluation metrics as defined in equations (30) and (31) for individual asset returns r_t (Panel A) and characteristic-managed portfolio returns x_t (Panel B). We also report the sum of squared $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients, denoted as W_α, and the corresponding bootstrapped p-values of the asset pricing test for testing the null hypothesis of an anomaly intercept equal to zero (Panel C). The results of the asset pricing test in Panel C suggest that the null hypothesis cannot be rejected at the one percent level if IPCA estimates at least two factors (p-value = 23.0 percent), suggesting that the test correctly identifies the true model if IPCA estimates at least as many factors as there are in the true model.

However, the asset pricing test suggests that there is a significant anomaly alpha if IPCA estimates only one factor (p-value = 0.00 percent), indicating a beta-eating effect as analytically shown in Section 2.1. This effect is also reflected in the performance measures presented in Panels A and B. The correctly specified two-factor beta-only IPCA achieves an R^2_{total} of 13.9 percent and an R^2_{pred} of 0.31 percent at the asset level. Not surprisingly, the one-factor beta-only IPCA has weaker performance, as a result of the omitted factor. However, estimating only one factor and additionally allowing for an alpha increases the R^2_{pred} to 0.31 percent, which equals that of the true model. By compensating for the omitted factor with a conditional alpha, the one-factor alpha-beta IPCA can describe expected returns $E_t[r_{i,t+1}]$ exactly as well as the true model, although it cannot capture the entire contemporaneous covariation among the assets as indicated by the lower R^2_{total} of 13.6 percent. Even when estimating an alpha-only IPCA that does not allow for any factors, IPCA achieves the same R^2_{pred} suggesting that alpha is sufficient for describing expected returns. These findings are in line with the analytical analysis in Section 2.1 and those obtained for a covariance-characteristic world in Table 4.

IPCA	Beta						Alpha-Beta						Alpha
K	1	2	3	4	5	6	1	2	3	4	5	6	0
Panel A: Individual assets (r_t)													
R^2_{total}	13.6	13.9	13.9	14.0	14.0	14.0	13.6	13.9	13.9	14.0	14.0	14.0	0.31
R^2_{pred}	0.27	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Panel B: Managed portfolios (x_t)													
R^2_{total}	99.9	100.0	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	1.84
R^2_{pred}	1.83	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84
Panel C: Asset pricing test													
W_α							0.07	0.01	0.00	0.00	0.00	0.00	
p-value							0.00	23.0	50.4	71.5	33.1	95.8	

Table 8: IPCA model performance (covariance-only world)
Panel A and B report the performance evaluation metrics R^2_{total} and R^2_{pred} for individual assets (r_t) and characteristic-managed portfolios (x_t), respectively, using a simulated covariance-only data set. Panel C reports the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$ (W_α) (multiplied by 100 for presentation purposes) and bootstrapped p-values in percent for testing the null hypothesis of $\hat{\mathbf{\Gamma}}_\alpha = \mathbf{0}$. The column "Beta" refers to an IPCA without alpha, "Alpha-Beta" refers to an IPCA with both alpha and beta, and "Alpha" is an IPCA that only estimates alpha and no betas.

To gain additional insights from the estimated IPCA specifications, Table 9 reports the estimated factor premia (i.e., the time-series averages of the estimated factors) and Table 10 reports the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients. If a factor premium significantly differs from zero at the one percent level, we report the respective premium in bold. In the two-factor beta-only IPCA, the estimated factor premia are $\hat{\lambda}_1 = 0.0333$ and $\hat{\lambda}_2 = 0.0124$, which are close to the true factor premia of $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$, respectively (see Table 3). As assumed in the underlying DGP, the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients (i.e., $\hat{\Gamma}_{\beta,9,1}$ and $\hat{\Gamma}_{\beta,10,1}$) are close to one, which confirms the finding that the IPCA identifies the true model, if it is correctly specified.

	K					
	1	2	3	4	5	6
Panel A: Beta-Only IPCA						
$\hat{\lambda}_1$	**0.0138**	**0.0333**	**0.0262**	0.0064	0.0051	0.0045
$\hat{\lambda}_2$		**0.0124**	**0.0200**	**0.0323**	**0.0188**	**0.0125**
$\hat{\lambda}_3$			**0.0126**	**0.0125**	**0.0267**	**0.0302**
$\hat{\lambda}_4$				0.0005	**0.0123**	**0.0121**
$\hat{\lambda}_5$					0.0001	0.0017
$\hat{\lambda}_6$						0.0006
Panel B: Alpha-Beta IPCA						
$\hat{\lambda}_1$	**0.0236**	**0.0319**	**0.0237**	0.0028	0.0022	0.0022
$\hat{\lambda}_2$		**0.0122**	**0.0225**	**0.0329**	**0.0168**	**0.0118**
$\hat{\lambda}_3$			**0.0128**	**0.0125**	**0.0285**	**0.0309**
$\hat{\lambda}_4$				0.0009	**0.0121**	**0.0122**
$\hat{\lambda}_5$					0.0011	0.0019
$\hat{\lambda}_6$						0.0003

Table 9: Estimated factor premia (covariance-only world)
The table reports the estimated factor premia of the beta-only and alpha-beta IPCA specifications, assuming a covariance-only world. If the factor premium significantly differs from zero at the 1 percent level, we report the respective premium in bold letters.

The results in Table 8 suggest that a one-factor alpha-beta IPCA and an alpha-only IPCA have the same explanatory power for expected returns as the true two-factor beta-only IPCA. In Section 2.1, we show that this is the case if the l-th element in $\hat{\mathbf{\Gamma}}_\alpha$ is chosen such that it equals the product of the l-th row of the true $\mathbf{\Gamma}_\beta$ and the vector of true factor premia $\boldsymbol{\lambda}$ that is not explained by the estimated betas and factor premia. Since we assume that the true $\mathbf{\Gamma}_\beta$ coefficients of both characteristics are one, we expect for the alpha-only IPCA that the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients of characteristics nine and ten equal the true factor premia. Table 11 presents the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only ICPA. As expected for the alpha-only IPCA, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients on these characteristics are very close to the calibrated factor premia presented in Table 3. Specifically, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient of characteristic nine is 0.0336 and the coefficient of characteristic ten is 0.0138. Therefore, the alpha-only IPCA identifies the true factor premia but assigns them to alpha. By doing this, an alpha-only IPCA can fully describe expected returns but

IPCA	Beta						Alpha-Beta					
K	1	2	3	4	5	6	1	2	3	4	5	6
$\hat{\Gamma}_{\beta,1,1}$	-0.02	-0.00	-0.12	-0.20	-0.18	-0.16	-0.02	-0.00	-0.11	-0.20	-0.18	-0.16
$\hat{\Gamma}_{\beta,2,1}$	0.01	0.01	0.02	-0.04	-0.04	-0.03	0.01	0.01	0.02	-0.04	-0.04	-0.03
$\hat{\Gamma}_{\beta,3,1}$	-0.00	0.01	0.16	0.08	0.39	0.35	-0.00	0.01	0.15	0.08	0.39	0.36
$\hat{\Gamma}_{\beta,4,1}$	-0.11	-0.13	-0.61	-0.95	-0.89	-0.91	-0.12	-0.16	-0.59	-0.95	-0.89	-0.90
$\hat{\Gamma}_{\beta,5,1}$	-0.02	-0.01	-0.15	-0.09	-0.06	-0.06	-0.02	-0.01	-0.14	-0.09	-0.07	-0.06
$\hat{\Gamma}_{\beta,6,1}$	-0.01	-0.01	-0.01	-0.01	0.00	0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.01
$\hat{\Gamma}_{\beta,7,1}$	-0.00	0.00	0.04	0.05	0.04	0.04	0.00	0.00	0.03	0.05	0.04	0.04
$\hat{\Gamma}_{\beta,8,1}$	-0.00	-0.00	-0.01	-0.01	-0.01	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01	-0.01
$\hat{\Gamma}_{\beta,9,1}$	0.92	0.99	0.75	0.19	0.15	0.15	0.91	0.99	0.77	0.20	0.15	0.15
$\hat{\Gamma}_{\beta,10,1}$	-0.37	0.04	0.04	-0.01	0.01	0.01	-0.41	0.03	0.04	-0.01	0.01	0.01
$\hat{\Gamma}_{\beta,1,2}$		0.04	0.17	0.06	-0.05	-0.00		0.04	0.18	0.07	-0.05	-0.00
$\hat{\Gamma}_{\beta,2,2}$		-0.00	-0.02	0.02	-0.02	-0.02		-0.00	-0.02	0.02	-0.02	-0.02
$\hat{\Gamma}_{\beta,3,2}$		0.04	-0.22	-0.02	-0.77	-0.86		0.04	-0.24	-0.02	-0.77	-0.86
$\hat{\Gamma}_{\beta,4,2}$		-0.02	0.66	0.18	-0.22	-0.26		-0.05	0.68	0.18	-0.23	-0.26
$\hat{\Gamma}_{\beta,5,2}$		0.02	0.20	0.03	-0.04	-0.03		0.03	0.21	0.03	-0.04	-0.03
$\hat{\Gamma}_{\beta,6,2}$		-0.00	0.00	-0.01	-0.04	-0.03		-0.00	0.00	-0.01	-0.03	-0.03
$\hat{\Gamma}_{\beta,7,2}$		0.00	-0.05	-0.02	0.02	0.03		0.00	-0.05	-0.02	0.02	0.03
$\hat{\Gamma}_{\beta,8,2}$		0.00	0.01	0.00	-0.00	-0.01		0.00	0.01	0.00	-0.00	-0.01
$\hat{\Gamma}_{\beta,9,2}$		-0.04	0.66	0.98	0.59	0.43		-0.04	0.63	0.98	0.59	0.43
$\hat{\Gamma}_{\beta,10,2}$		1.00	-0.01	0.04	-0.05	-0.05		1.00	-0.01	0.04	-0.05	-0.05
$\hat{\Gamma}_{\beta,1,3}$			0.05	0.03	0.10	0.05			0.05	0.04	0.11	0.04
$\hat{\Gamma}_{\beta,2,3}$			-0.00	0.02	0.03	0.02			-0.00	0.02	0.04	0.02
$\hat{\Gamma}_{\beta,3,3}$			0.02	0.09	0.50	0.35			0.01	0.09	0.49	0.36
$\hat{\Gamma}_{\beta,4,3}$			0.02	-0.02	0.32	0.27			0.02	-0.02	0.32	0.27
$\hat{\Gamma}_{\beta,5,3}$			0.04	-0.02	0.05	0.03			0.05	-0.02	0.06	0.03
$\hat{\Gamma}_{\beta,6,3}$			-0.00	-0.00	0.02	0.00			-0.00	-0.00	0.02	0.00
$\hat{\Gamma}_{\beta,7,3}$			-0.00	0.00	-0.03	-0.03			-0.00	0.00	-0.03	-0.03
$\hat{\Gamma}_{\beta,8,3}$			0.00	0.00	0.00	0.01			0.00	0.00	0.00	0.01
$\hat{\Gamma}_{\beta,9,3}$			-0.03	-0.04	0.79	0.89			-0.03	-0.04	0.79	0.89
$\hat{\Gamma}_{\beta,10,3}$			1.00	0.99	0.09	0.07			1.00	0.99	0.08	0.07
$\hat{\Gamma}_{\beta,1,4}$				0.09	0.03	0.09				-0.09	0.03	0.10
$\hat{\Gamma}_{\beta,2,4}$				-0.26	0.00	0.01				0.26	0.01	0.01
$\hat{\Gamma}_{\beta,3,4}$				-0.73	-0.08	-0.06				0.73	-0.08	-0.06
$\hat{\Gamma}_{\beta,4,4}$				-0.13	-0.04	-0.04				0.13	-0.04	-0.04
$\hat{\Gamma}_{\beta,5,4}$				0.60	-0.00	0.02				-0.60	-0.00	0.03
$\hat{\Gamma}_{\beta,6,4}$				-0.00	-0.01	0.00				0.00	-0.01	0.00
$\hat{\Gamma}_{\beta,7,4}$				-0.05	0.01	0.01				0.05	0.01	0.01
$\hat{\Gamma}_{\beta,8,4}$				-0.01	0.00	-0.01				0.01	0.00	-0.01
$\hat{\Gamma}_{\beta,9,4}$				-0.02	-0.04	-0.04				0.02	-0.04	-0.04
$\hat{\Gamma}_{\beta,10,4}$				0.07	0.99	0.99				-0.07	0.99	0.99
$\hat{\Gamma}_{\beta,1,5}$					0.18	-0.80					-0.17	-0.81
$\hat{\Gamma}_{\beta,2,5}$					-0.35	-0.00					0.35	-0.00
$\hat{\Gamma}_{\beta,3,5}$					-0.03	-0.03					0.03	-0.03
$\hat{\Gamma}_{\beta,4,5}$					-0.11	0.17					0.11	0.17
$\hat{\Gamma}_{\beta,5,5}$					0.91	-0.55					-0.91	-0.53
$\hat{\Gamma}_{\beta,6,5}$					0.04	-0.15					-0.03	-0.15
$\hat{\Gamma}_{\beta,7,5}$					-0.10	0.00					0.10	-0.00
$\hat{\Gamma}_{\beta,8,5}$					-0.01	0.10					0.01	0.10
$\hat{\Gamma}_{\beta,9,5}$					-0.01	0.01					0.01	0.01
$\hat{\Gamma}_{\beta,10,5}$					-0.01	0.09					0.01	0.10
$\hat{\Gamma}_{\beta,1,6}$						-0.46						-0.44
$\hat{\Gamma}_{\beta,2,6}$						-0.48						-0.47
$\hat{\Gamma}_{\beta,3,6}$						-0.03						-0.03
$\hat{\Gamma}_{\beta,4,6}$						0.03						0.03
$\hat{\Gamma}_{\beta,5,6}$						0.73						0.74
$\hat{\Gamma}_{\beta,6,6}$						-0.09						-0.08
$\hat{\Gamma}_{\beta,7,6}$						-0.13						-0.13
$\hat{\Gamma}_{\beta,8,6}$						0.08						0.08
$\hat{\Gamma}_{\beta,9,6}$						0.01						0.00
$\hat{\Gamma}_{\beta,10,6}$						0.03						0.03

Table 10: Estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients (covariance-only world)
The table reports the estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients for the beta-only and alpha-beta (i.e., unrestricted) IPCA, respectively, using a simulated covariance-only data set. The term $\hat{\Gamma}_{\beta,l,j}$ denotes the estimated coefficient of the l-th characteristic for the beta on the j-th factor.

not the covariance among the assets. Even in the misspecified one-factor unrestricted IPCA, the

$\hat{\mathbf{\Gamma}}_\alpha$ coefficients of characteristics nine and ten are nonzero (i.e., 0.0114 and 0.0235). Thus, IPCA assigns components of both true betas to the anomaly intercept that are important for modeling expected returns but that cannot be explained by a single beta on a common factor.

IPCA			Alpha-Beta				Alpha
K	1	2	3	4	5	6	0
$\hat{\mathbf{\Gamma}}_{\alpha,1}$	0.0010	-0.0000	-0.0018	-0.0018	-0.0018	0.0002	-0.0020
$\hat{\mathbf{\Gamma}}_{\alpha,2}$	-0.0002	-0.0002	-0.0000	-0.0011	-0.0012	-0.0005	0.0002
$\hat{\mathbf{\Gamma}}_{\alpha,3}$	0.0007	0.0002	0.0025	-0.0002	0.0001	0.0000	-0.0030
$\hat{\mathbf{\Gamma}}_{\alpha,4}$	0.0064	0.0092	0.0019	0.0005	0.0004	0.0000	0.0024
$\hat{\mathbf{\Gamma}}_{\alpha,5}$	0.0001	-0.0004	-0.0025	-0.0003	-0.0000	-0.0002	-0.0007
$\hat{\mathbf{\Gamma}}_{\alpha,6}$	-0.0006	-0.0005	-0.0006	-0.0006	-0.0006	-0.0002	-0.0007
$\hat{\mathbf{\Gamma}}_{\alpha,7}$	-0.0004	-0.0004	0.0001	-0.0000	-0.0001	0.0001	-0.0006
$\hat{\mathbf{\Gamma}}_{\alpha,8}$	0.0003	0.0003	0.0002	0.0002	0.0002	-0.0001	0.0001
$\hat{\mathbf{\Gamma}}_{\alpha,9}$	0.0114	0.0014	0.0002	0.0001	0.0001	0.0000	0.0336
$\hat{\mathbf{\Gamma}}_{\alpha,10}$	0.0235	0.0005	0.0002	0.0001	0.0001	-0.0000	0.0138

Table 11: Estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients (covariance-only world)
The table reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta (i.e., unrestricted) and alpha-only IPCA, using a simulated covariance-only data set. The term $\hat{\Gamma}_{\alpha,l}$ denotes the estimated coefficient of the l-th characteristic for the alpha.

B Characteristic-only world

Here, we simulate asset returns assuming that the expected returns are fully determined by the characteristics themselves. The betas only describe systematic covariances among the assets. Accordingly, the true IPCA model is the alpha-beta IPCA with two factors, where the factor premia are zero. The true DGP is described by:

$$r_{i,t+1} = \underbrace{\alpha_{i,t}}_{E_t[r_{i,t+1}]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{36}$$

where all parameters are calibrated as described in Tables 1, 2, and 3. The performance measures and the results of the asset pricing test are reported in Table 12.[11]

The asset pricing test in Panel C indicates that there is a significant anomaly alpha after controlling for one or two factors (p-value = 0.00). However, the magnitude of the $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients, measured by the sum of squared elements W_α, declines with every additional factor and the null hypothesis of an anomaly intercept equal to zero cannot be rejected at the one percent level if IPCA estimates at least three factors (p-value = 9.40). Thus, the asset pricing test of KPS concludes that asset returns are determined by covariances, not characteristics, if IPCA estimates at least three factors. This finding shows that there is an "alpha-eating" effect, which we derived in Section 2.2.

The true IPCA (i.e., two-factor alpha-beta) achieves an R^2_{total} of 13.3 percent and an R^2_{pred} of 0.35 percent for individual assets. Omitting one or even both factors results in a decline in the R^2_{total} to 13.0 or 0.35 percent, respectively, but does not affect the R^2_{pred}. This is to be expected because, according to the underlying DGP, the betas should not explain expected returns but only covariances among the assets. However, the one- or two-factor beta-only IPCA cannot describe expected returns, as indicated by the relatively low R^2_{pred} of 0.04 and 0.20 percent, respectively, compared to 0.35 percent achieved by the true alpha-beta IPCA. When including a third factor, the R^2_{pred} for the beta-only IPCA increases to 0.35 percent, which is equivalent to that of the true model. Thus, IPCA attributes the explanatory power of the missing alpha to the betas, indicating an "alpha-eating" effect. This "alpha-eating" effect is evident not only in the beta-only IPCA but

[11] In unreported results, we also simulate data according to a characteristic-only world, in which no factor innovations exist. The results remain qualitatively unchanged. However, this scenario is unrealistic because we would not observe covariations among asset returns; therefore, we do not consider this scenario in this paper.

also in the alpha-beta IPCA, as indicated by the declining sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$ with an increasing number of factors.

IPCA	Beta						Alpha-Beta						Alpha
K	1	2	3	4	5	6	1	2	3	4	5	6	0
Panel A: Individual assets (r_t)													
R^2_{total}	12.7	13.0	13.3	13.4	13.4	13.4	13.0	13.3	13.3	13.4	13.4	13.4	0.35
R^2_{pred}	0.04	0.20	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35
Panel B: Managed portfolios (x_t)													
R^2_{total}	99.8	99.9	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	0.35
R^2_{pred}	0.28	0.32	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35
Panel C: Asset pricing test													
W_α							0.76	0.76	0.64	0.50	0.40	0.27	
p-value							0.00	0.00	9.40	6.60	6.30	45.0	

Table 12: IPCA model performance (characteristic-only world)
Panel A and B report the performance evaluation metrics R^2_{total} and R^2_{pred} for individual assets (r_t) and characteristic-managed portfolios (x_t), respectively, using a simulated characteristic-only data set. Panel C reports the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$ (W_α) (multiplied by 100 for presentation purposes) and bootstrapped p-values in percent for testing the null hypothesis of $\hat{\mathbf{\Gamma}}_\alpha = \mathbf{0}$. The column "Beta" refers to an IPCA without alpha, "Alpha-Beta" refers to an IPCA with both alpha and beta, and "Alpha" is an IPCA that only estimates alpha and no betas.

Again, we analyze the IPCA results in greater detail by reporting the estimated factor premia and the $\hat{\mathbf{\Gamma}}_\beta$ and the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients. According to the true DGP, we expect factor premia not distinguishable from zero, $\hat{\mathbf{\Gamma}}_\beta$ coefficients of one for characteristics nine and ten, and $\hat{\mathbf{\Gamma}}_\alpha$ coefficients of 0.0680 and -0.0543 for characteristics one and two. All other $\hat{\mathbf{\Gamma}}_{\alpha,\beta}$ coefficients should be zero. Estimating a one-factor IPCA, the estimated factor premium is not significant at the 1 percent level, neither for the beta-only nor for the alpha-beta specification. However, in the two-factor alpha-beta IPCA, the factor premium of the second factor in the alpha-beta IPCA is statistically significant at the 1 percent level but small in magnitude and therefore not economically significant. However, the picture is different for the beta-only IPCA. The premia of $\hat{\lambda}_1 = 0.0390$ and $\hat{\lambda}_2 = 0.0426$ are statistically significant and large in magnitude, which already indicates the existence of the "alpha-eating" effect. When estimating three (or more) factors in the alpha-beta IPCA, most factor premia become distinguishable from zero, indicating that these factors become (more) relevant. This supports the finding that IPCA constructs a covariance-based explanation for asset returns if

IPCA estimates more factors than there are in the true model.

	K					
	1	2	3	4	5	6
Panel A: Beta-only IPCA						
$\hat{\lambda}_1$	0.0049	**0.0390**	0.0008	0.0063	0.0034	**0.0263**
$\hat{\lambda}_2$		**0.0426**	**0.0877**	**0.0835**	**0.0859**	0.0015
$\hat{\lambda}_3$			**0.0091**	**0.0179**	**0.0113**	**0.0824**
$\hat{\lambda}_4$				**0.0203**	**0.0073**	**0.0089**
$\hat{\lambda}_5$					**0.0133**	0.0015
$\hat{\lambda}_6$						**0.0129**
Panel B: Alpha-Beta IPCA						
$\hat{\lambda}_1$	0.0031	0.0003	0.0016	0.0007	0.0074	**0.0229**
$\hat{\lambda}_2$		**0.0070**	**0.0334**	**0.0446**	0.0008	**0.0122**
$\hat{\lambda}_3$			**0.0107**	**0.0235**	**0.0567**	**0.0604**
$\hat{\lambda}_4$				**0.0129**	**0.0202**	**0.0198**
$\hat{\lambda}_5$					0.0045	**0.0145**
$\hat{\lambda}_6$						**0.0063**

Table 13: Estimated factor premia (characteristic-only world)
The table reports the estimated factor premia of the beta-only and alpha-beta IPCA specifications, assuming a characteristic-only world. If the factor premium significantly differs from zero at the 1 percent level, we report the respective premium in bold letters.

Table 15 presents the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients. In the true two-factor alpha-beta IPCA, the $\hat{\mathbf{\Gamma}}_\beta$ coefficients on characteristics nine and ten are close to one and those on all other characteristics are close to zero. Even if the factor premium of the second factor ($\lambda_2 = 0.0070$) is statistically significant, it is very small and may not be economically significant. Therefore, IPCA correctly identifies the true model. However, when estimating three factors and alpha, the $\hat{\mathbf{\Gamma}}_\beta$ coefficient of the first characteristic on the second beta increases ($\hat{\Gamma}_{\beta,1,2}$) and the corresponding factor premium is large in magnitude ($\lambda_2 = 0.0334$), suggesting that the relationship between the second characteristic and returns is described by conditional betas, not by the characteristic itself. This erroneous conclusion is the result of the "alpha-eating" effect that becomes stronger with additional factors.

The "alpha-eating" effect already appears in the two-factor beta-only IPCA. Here, the coefficients of both the first and second characteristics to the first (i.e., $\hat{\Gamma}_{\beta,1,1} = 0.27$ and $\hat{\Gamma}_{\beta,2,1} =$ -0.21) and second beta (i.e., $\hat{\Gamma}_{\beta,1,2} = 0.50$ and $\hat{\Gamma}_{\beta,2,2} =$ -0.45) are large, while the coefficient of characteristic ten for the second beta (i.e., $\hat{\Gamma}_{\beta,10,2} =$ -0.70) is smaller than assumed in the true DGP, suggesting that IPCA fails to identify the true betas because the betas eat the alpha. However, two factors are not enough to eat up all the alpha, as indicated by the relatively low R^2_{pred} reported in Table 12. As described in Section 2.2, this is because the system of equations is underdetermined with $L = 10$ equations and only two optimizable parameters. However, when estimating three factors, IPCA can provide a full covariance-only description of asset returns.

Table 14 reports the $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only IPCA. We note that the true coefficients for the first and second characteristics are 0.0680 and -0.0547, respectively, while those for the other characteristics are zero (see Table 2). When estimating an alpha-only or unrestricted IPCA with up to two factors, the estimated coefficients are very close to the true coefficients, suggesting that IPCA identifies the true model. However, there is large decrease in the magnitude of the $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficient of the first characteristic when including a third factor.[12] This is in line with the findings from Table 15 that reports increasing $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients on the first characteristic when increasing the number of IPCA factors.

IPCA	Alpha-Beta						Alpha
K	1	2	3	4	5	6	0
$\hat{\Gamma}_{\alpha,1}$	0.0680	0.0675	0.0493	0.0324	0.0155	-0.0006	0.0673
$\hat{\Gamma}_{\alpha,2}$	-0.0543	-0.0543	-0.0540	-0.0528	-0.0530	-0.0508	-0.0542
$\hat{\Gamma}_{\alpha,3}$	-0.0049	-0.0047	-0.0302	-0.0060	-0.0015	-0.0003	-0.0058
$\hat{\Gamma}_{\alpha,4}$	0.0065	0.0063	-0.0033	-0.0234	0.0018	-0.0021	0.0068
$\hat{\Gamma}_{\alpha,5}$	-0.0001	-0.0000	-0.0105	-0.0224	-0.0283	0.0008	-0.0000
$\hat{\Gamma}_{\alpha,6}$	0.0006	0.0006	-0.0023	0.0044	0.0092	0.0024	0.0005
$\hat{\Gamma}_{\alpha,7}$	-0.0000	0.0000	-0.0003	0.0005	0.0035	0.0104	0.0001
$\hat{\Gamma}_{\alpha,8}$	-0.0000	-0.0000	0.0016	0.0026	0.0018	-0.0030	-0.0000
$\hat{\Gamma}_{\alpha,9}$	0.0017	-0.0002	-0.0006	-0.0005	-0.0003	-0.0004	-0.0009
$\hat{\Gamma}_{\alpha,10}$	-0.0020	-0.0069	-0.0051	-0.0024	-0.0015	0.0003	-0.0006

Table 14: $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients (characteristic-only world)
The table reports the estimated $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients for the alpha-beta (i.e., unrestricted) and alpha-only IPCA, using a simulated characteristic-only data set. The term $\hat{\Gamma}_{\alpha,l}$ denotes the estimated coefficient of the l-th characteristic for the alpha.

[12]Furthermore, the coefficient on the third characteristic increases in magnitude to -0.0302. However, the average correlation of the first and third characteristics is only 0.0052. Accordingly, we can rule out the possibility that the third characteristic subsumes the explanatory power of the first characteristic.

IPCA	Beta						Alpha-Beta					
K	1	2	3	4	5	6	1	2	3	4	5	6
$\hat{\Gamma}_{\beta,1,1}$	0.01	0.27	0.01	0.07	0.04	0.44	0.04	0.01	0.02	0.02	0.25	0.18
$\hat{\Gamma}_{\beta,2,1}$	-0.04	-0.21	-0.01	-0.03	-0.01	-0.10	0.01	-0.01	0.00	0.01	0.02	-0.04
$\hat{\Gamma}_{\beta,3,1}$	-0.01	-0.01	-0.01	0.05	0.08	0.58	0.00	-0.01	0.05	0.08	0.57	-0.40
$\hat{\Gamma}_{\beta,4,1}$	-0.02	0.05	-0.01	0.02	0.01	-0.67	0.02	-0.01	0.02	0.01	-0.78	0.89
$\hat{\Gamma}_{\beta,5,1}$	0.01	-0.00	0.00	0.02	0.01	0.05	-0.00	0.00	0.01	0.01	0.03	-0.03
$\hat{\Gamma}_{\beta,6,1}$	0.01	-0.00	0.00	0.00	0.01	0.05	-0.01	0.00	-0.00	0.01	0.04	-0.01
$\hat{\Gamma}_{\beta,7,1}$	0.00	-0.00	0.00	-0.00	-0.00	-0.05	-0.00	0.00	-0.00	-0.00	-0.04	-0.01
$\hat{\Gamma}_{\beta,8,1}$	-0.00	0.00	-0.00	0.00	0.00	0.02	0.00	-0.00	0.00	0.00	0.02	-0.01
$\hat{\Gamma}_{\beta,9,1}$	0.93	-0.94	1.00	-0.99	-0.99	0.03	-0.93	1.00	-1.00	-0.99	0.04	-0.04
$\hat{\Gamma}_{\beta,10,1}$	-0.38	0.03	0.07	-0.05	-0.05	-0.02	0.37	0.07	-0.06	-0.06	-0.04	0.06
$\hat{\Gamma}_{\beta,1,2}$		0.50	0.78	0.81	0.82	0.02		0.10	0.51	0.63	0.01	0.19
$\hat{\Gamma}_{\beta,2,2}$		-0.45	-0.61	-0.51	-0.52	-0.00		0.00	-0.01	-0.02	0.01	-0.00
$\hat{\Gamma}_{\beta,3,2}$		-0.04	-0.06	0.17	0.03	0.06		-0.01	0.64	0.40	0.07	0.15
$\hat{\Gamma}_{\beta,4,2}$		0.05	0.10	0.13	0.16	0.03		0.03	0.26	0.43	0.01	-0.01
$\hat{\Gamma}_{\beta,5,2}$		0.01	0.00	0.09	0.10	0.01		-0.00	0.26	0.36	0.01	0.02
$\hat{\Gamma}_{\beta,6,2}$		0.02	0.00	0.03	-0.00	0.00		-0.02	0.07	0.00	0.00	0.01
$\hat{\Gamma}_{\beta,7,2}$		0.00	-0.00	0.00	-0.00	0.00		-0.01	0.01	-0.00	-0.00	-0.02
$\hat{\Gamma}_{\beta,8,2}$		0.00	-0.00	-0.01	-0.01	0.00		-0.00	-0.04	-0.05	0.00	0.00
$\hat{\Gamma}_{\beta,9,2}$		0.22	-0.02	0.07	0.04	-1.00		-0.07	0.03	0.03	-1.00	-0.97
$\hat{\Gamma}_{\beta,10,2}$		-0.70	0.10	0.14	0.11	-0.05		0.99	0.44	0.37	-0.06	-0.04
$\hat{\Gamma}_{\beta,1,3}$			0.04	0.07	-0.01	0.73			0.17	0.23	0.79	0.89
$\hat{\Gamma}_{\beta,2,3}$			-0.11	-0.23	-0.11	-0.52			-0.01	-0.02	-0.02	-0.03
$\hat{\Gamma}_{\beta,3,3}$			-0.00	-0.20	-0.93	-0.24			0.36	-0.85	0.07	0.35
$\hat{\Gamma}_{\beta,4,3}$			-0.02	-0.04	0.09	0.35			0.11	0.38	0.31	-0.02
$\hat{\Gamma}_{\beta,5,3}$			0.01	-0.07	-0.02	0.09			0.15	0.18	0.44	0.08
$\hat{\Gamma}_{\beta,6,3}$			0.02	-0.00	-0.17	-0.04			0.05	-0.19	-0.09	0.03
$\hat{\Gamma}_{\beta,7,3}$			0.01	0.00	-0.02	0.01			0.01	-0.03	-0.03	-0.11
$\hat{\Gamma}_{\beta,8,3}$			0.00	0.01	0.02	-0.02			-0.02	-0.01	-0.04	0.02
$\hat{\Gamma}_{\beta,9,3}$			0.07	0.05	-0.06	0.00			0.08	-0.06	0.01	0.23
$\hat{\Gamma}_{\beta,10,3}$			-0.99	-0.95	-0.29	0.10			-0.89	-0.09	0.27	0.10
$\hat{\Gamma}_{\beta,1,4}$				-0.04	0.06	0.05				0.19	0.28	0.26
$\hat{\Gamma}_{\beta,2,4}$				-0.35	-0.09	-0.10				-0.01	-0.01	-0.02
$\hat{\Gamma}_{\beta,3,4}$				-0.83	0.29	-0.40				0.24	-0.45	-0.48
$\hat{\Gamma}_{\beta,4,4}$				-0.09	-0.05	-0.28				0.15	-0.20	-0.22
$\hat{\Gamma}_{\beta,5,4}$				-0.30	0.01	-0.03				0.17	0.17	0.02
$\hat{\Gamma}_{\beta,6,4}$				-0.09	0.07	-0.08				0.03	-0.14	-0.12
$\hat{\Gamma}_{\beta,7,4}$				-0.02	0.01	-0.03				0.01	-0.05	-0.08
$\hat{\Gamma}_{\beta,8,4}$				0.05	-0.00	0.02				-0.02	0.01	0.03
$\hat{\Gamma}_{\beta,9,4}$				-0.06	0.08	0.02				0.08	0.01	0.01
$\hat{\Gamma}_{\beta,10,4}$				0.28	-0.94	-0.86				-0.92	-0.79	-0.79
$\hat{\Gamma}_{\beta,1,5}$					-0.20	-0.02					0.07	0.20
$\hat{\Gamma}_{\beta,2,5}$					-0.56	-0.02					0.00	-0.00
$\hat{\Gamma}_{\beta,3,5}$					-0.00	0.63					-0.65	-0.65
$\hat{\Gamma}_{\beta,4,5}$					-0.43	0.56					-0.49	-0.37
$\hat{\Gamma}_{\beta,5,5}$					-0.65	0.09					-0.05	0.00
$\hat{\Gamma}_{\beta,6,5}$					0.16	0.16					-0.17	-0.18
$\hat{\Gamma}_{\beta,7,5}$					0.01	0.08					-0.07	-0.07
$\hat{\Gamma}_{\beta,8,5}$					0.08	-0.05					0.04	0.03
$\hat{\Gamma}_{\beta,9,5}$					-0.01	0.08					-0.08	-0.09
$\hat{\Gamma}_{\beta,10,5}$					0.06	-0.49					0.54	0.60
$\hat{\Gamma}_{\beta,1,6}$						-0.29						0.06
$\hat{\Gamma}_{\beta,2,6}$						-0.62						-0.04
$\hat{\Gamma}_{\beta,3,6}$						0.09						0.00
$\hat{\Gamma}_{\beta,4,6}$						-0.06						-0.05
$\hat{\Gamma}_{\beta,5,6}$						-0.68						-0.95
$\hat{\Gamma}_{\beta,6,6}$						0.22						0.22
$\hat{\Gamma}_{\beta,7,6}$						0.05						-0.16
$\hat{\Gamma}_{\beta,8,6}$						0.06						0.14
$\hat{\Gamma}_{\beta,9,6}$						-0.00						-0.00
$\hat{\Gamma}_{\beta,10,6}$						0.03						-0.00

Table 15: $\hat{\mathbf{\Gamma}}_\beta$ coefficients (characteristic-only world)

The table reports the estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients for the beta-only and alpha-beta (i.e., unrestricted) IPCA, respectively, using a simulated characteristic-only data set. The term $\hat{\Gamma}_{\beta,l,j}$ denotes the estimated coefficient of the l-th characteristic for the beta on the j-th factor.

C Empirical results

This appendix replicates Table 1 in KPS and extends it by including results for the alpha-only IPCA (i.e., $K = 0$). The results suggest that the estimation of a conditional alpha leads to a better description of the conditional expected returns for both individual stock and characteristic-managed portfolio returns. The alpha-only IPCA that estimates no factors but only a conditional alpha achieves the highest R^2_{pred} among all models, with is in line with our simulation results.

		K						
		0	1	2	3	4	5	6
Panel A: Individual stocks (r_t)								
R^2_{total}	$\mathbf{\Gamma}_\alpha = \mathbf{0}$		14.8	16.4	17.4	18.0	18.6	18.9
	$\mathbf{\Gamma}_\alpha \neq \mathbf{0}$	0.77	15.2	16.8	17.7	18.4	18.7	19.0
R^2_{pred}	$\mathbf{\Gamma}_\alpha = \mathbf{0}$		0.35	0.34	0.41	0.42	0.69	0.68
	$\mathbf{\Gamma}_\alpha \neq \mathbf{0}$	0.77	0.76	0.75	0.75	0.74	0.74	0.72
Panel B: Managed portfolios (x_t)								
R^2_{total}	$\mathbf{\Gamma}_\alpha = \mathbf{0}$		90.3	95.3	97.1	98.0	98.4	98.8
	$\mathbf{\Gamma}_\alpha \neq \mathbf{0}$	2.67	90.8	95.7	97.3	98.2	98.6	98.9
R^2_{pred}	$\mathbf{\Gamma}_\alpha = \mathbf{0}$		2.01	2.00	2.10	2.13	2.41	2.39
	$\mathbf{\Gamma}_\alpha \neq \mathbf{0}$	2.67	2.61	2.56	2.54	2.51	2.50	2.46

Table 16: IPCA model performance (in-sample)
Panel A and B report R^2_{total} and R^2_{pred} in percent for the beta-only $\mathbf{\Gamma}_\alpha = \mathbf{0}$ and alpha-beta $\mathbf{\Gamma}_\alpha \neq \mathbf{0}$ IPCA model using the same U.S. equity data set as in KPS. We include a special case, that is, an alpha-only IPCA that includes no common factors (i.e., $K = 0$).