Characteristics are Covariances? A Comment on Instrumented Principal Component Analysis

Lars Hornuf[a,*], Gerrit Liedtke[a], Thorsten Poddig[a], Christian Fieberg[a]

[a] *University of Bremen, Bremen, Germany*

Abstract

For more than thirty years now, there is an ongoing debate of whether the relationship between firm characteristics and stock returns is due to covariances with common risk factors or due to the characteristics themselves. In their paper "Characteristics are covariances: A unified model of risk and return", Kelly, Pruitt, and Su (2019) (KPS) propose a new factor modeling approach, called Instrumented Principal Component Analysis (IPCA). IPCA aims to find a reduced-form factor model by mapping a large number of firm characteristics to either risk factor exposures or an anomaly intercept, depending on whether return variation associated with firm characteristics is due to covariances with common risk factors or due to the characteristics themselves. Accordingly, the authors claim that IPCA can contribute to the ongoing "characteristics vs. covariances" debate. They find that a five-factor model without anomaly intercept explains the cross section of stock returns, and thus, the characteristics/expected returns relationship is due to covariances with common risk factors. We present analytical and simulation-based evidence that IPCA cannot reliably distinguish between the covariances and characteristics story. If IPCA estimates fewer factors than there are in the true model, the anomaly intercept subsumes the explanatory power of the omitted factors. Conversely, if IPCA estimates more factors than there are in the true model, covariances with risk factors subsume the explanatory power of a potential anomaly intercept. We refer to these observations as "beta-eating" and "alpha-eating", respectively. However, even though IPCA cannot identify the true model (i.e., whether covariances or characteristics explain returns), IPCA always identifies a model that has equivalent performance, making IPCA an upstanding tool for describing returns.

Keywords: Cross section of returns, Factor model, Anomaly, IPCA, Simulation

*Corresponding author, Max-von-Laue-Straße 1, 28359 Bremen, Germany.

Email addresses: hornuf@uni-bremen.de (Lars Hornuf), gliedtke@uni-bremen.de (Gerrit Liedtke), poddig@uni-bremen.de (Thorsten Poddig), cfieberg@uni-bremen.de (Christian Fieberg)

1. Introduction

Contrary to the presumption of the Capital Asset Pricing Model (CAPM), the literature identified more than 400 firm characteristics that appear to explain the cross section of (stock) returns (Hou et al., 2020), with a rising tendency (Harvey et al., 2016). Since these return driving phenomenons are not explained by the CAPM, they are referred to as "anomalies". At least since the literature has discovered such "anomalous returns", the literature questions what causes them. Some argue that these patterns may be found due to statistical biases, e.g., selection or survivorship bias (Kothari et al., 1995), or data mining (Lo and MacKinlay, 1990; Harvey et al., 2016). McLean and Pontiff (2016) and Jacobs and Müller (2020) find that the return variation associated with some characteristics vanishes in post-sample or post-publication periods, i.e., they are in-sample phenomenons. Furthermore, some anomalies are close cousins and do not provide incremental information (Green et al., 2017). Although some empirical studies might be biased, some anomalies do not disappear and resist the aforementioned explanations, resulting in characteristic-based return predictability (e.g., Lewellen, 2015; Gu et al., 2020).

There are two common theories about what causes these "anomalous returns", and even after thirty years, research cannot clearly answer the question. On the one hand, researchers argue that characteristics proxy for asset's covariances with common risk factors. For example, firms with high book-to-market equity ratios are more likely to be distressed and are therefore riskier, resulting in higher loadings on a common (unobserved) distress factor. Thus, the "anomalous returns" associated with different levels of book-to-market equity ratios are compensations for bearing systematic risk (e.g., Fama and French, 1993). Another fraction of researchers claims that "anomalous returns" are due to the characteristics themselves, i.e., anomalies represent mispricing, as a result of investors' over- or underreaction to (fundamental) information (e.g., De Bondt and Thaler, 1987; Lakonishok et al., 1994). To catch up with the example above, firms with high book-to-market equity ratios (so-called value stocks) tend to have poor past earnings growth. On the contrary, firms with low book-to-market equity ratios (so-called growth stocks) have strong past earnings growth. Investors incorrectly extrapolate past earnings; therefore, they are overly optimistic about well-performing firms (growth) and overly pessimistic about those that performed poorly (value). As a result, growth stocks have irrationally high prices, while value stocks have irrationally low prices. The correction of this mispricing results in low returns for growth and high returns for value stocks (Lakonishok et al., 1994).

A first test for distinguishing between the covariance- or characteristic-based explanation is proposed by Daniel and Titman (1997) and Daniel et al. (2001). They sort stocks into portfolios based on firm characteristics and loadings on common factors to test for return variation associated with characteristics that is orthogonal to risk factor loadings. They find that characteristics drive the return variation, not the covariance with common factors. However, Davis et al. (2000) present contrary findings when extending the sample period. A major drawback of the Daniel and Titman (1997) approach is that is suffers considerably from the curse of dimensionality. According to Green et al. (2017), the relationship between returns and firm characteristics is multidimensional. When sorting assets into portfolios based on multiple characteristics and/or factor loadings, the portfolios are not sufficiently diversified, making this approach infeasible.

Kelly, Pruitt, and Su (2019) (KPS) propose a new method, called Instrumented Principal Component Analysis (IPCA), which aims to map a potentially large number of characteristics to either risk factor exposures (beta) or an "anomaly intercept" (alpha). More precisely, IPCA identifies common latent factors from the covariance matrix of asset returns and allows factor loadings to depend on observable firm characteristics. If characteristics proxy for the compensation for bearing systematic risk,[1] IPCA identifies the corresponding (latent) risk factors and the conditional loadings (i.e., betas). If the characteristic story is true, IPCA will additionally find a significant (conditional) alpha, given some latent factors. They fit beta-only (i.e., only allowing for beta) and alpha-beta (i.e., allowing for alpha and beta) IPCA specifications to empirical stock return data and find that a beta-only model with five factors explains the same amount of variation as the alpha-beta model. They 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 at least five factors are included. Accordingly, they conclude that characteristics proxy for the covariance to common risk factors and that there are no "anomalous returns" in the sense that the characteristics themselves drive the returns. In a recent application, Buechner and Kelly (2021) apply IPCA to option prices and find that a three-factor model without alpha best explains option prices. Kelly et al. (2021) find that a beta-only five-factor IPCA is sufficient for explaining the cross section of bond returns. These results suggest that not only stock returns but also option and bond returns are driven by covariances.

However, despite KPS demonstrate that IPCA outperforms existing asset pricing models, they do not present evidence that IPCA can truly differentiate between covariances and characteristics. Therefore, this paper aims to test for the validity of their findings in two ways. First, we present analytical evidence that IPCA can always find a covariance-only- or characteristic-only-based explanation for returns if IPCA is misspecified. Second, we simulate asset return data according to different data generating processes (DGPs) that correspond to the covariance and/or characteristic story, respectively. Specifically, we create data according to a 1) multifactor model (i.e., covariance-only world), 2) characteristic model (i.e., characteristic-only world), and 3) a hybrid of a multifactor and characteristic model (i.e., covariance-characteristic world). For each of these simulated data sets, we fit three IPCA versions, one allowing for beta only, one allowing for alpha only, and a third one that allows for both alpha and beta. We use the asset pricing test proposed by KPS to test if IPCA can reliably identify the true model.

Our findings can be summarized as follows. First, if IPCA estimates the same number of factors as exist

[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).

in the true model, IPCA correctly identifies the true model and the relevant characteristics. Second, if IPCA estimates more factors than there are in the true model, we observe an "alpha-eating" effect, that is, factor exposures subsume the return variation that is actually attributable to an anomaly intercept. The "alpha-eating" effect biases the results such that the asset pricing 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 and characteristics story. Conversely, if IPCA allows for alpha and the number of estimated factors is below the number of true factors, a "beta-eating" effect occurs. That is, alpha subsumes some explanatory power of the omitted factors. All these findings combined 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. covariance" debate. Although IPCA may fail to identify the true model, it always finds a model that yields the same performance, making it a useful tool in, e.g., asset return prediction.

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" effect. Section 3 presents our simulation study. Section 4 concludes the paper.

2. A Model for the cross section of 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 $t+1$, $\boldsymbol{f}_{t+1}$ is a $K \times 1$ vector of contemporaneous (latent) factor realizations, with K denoting the number of factors, $\boldsymbol{\beta}_{i,t}$ is a corresponding $1 \times K$ vector of (conditional) factor loadings (i.e., betas), $\alpha_{i,t}$ is a (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's firm characteristics $\boldsymbol{z}_{i,t}$ (with L denoting the number of firm 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 firm characteristics. KPS point out that if all coefficients of $\boldsymbol{\Gamma}_{\boldsymbol{\alpha}}$ are zero, this does not imply that there are no alphas in general but it means that alphas are not linked to firm 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 divide 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 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., $\alpha_{i,t}$). 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., $\boldsymbol{\Gamma}_\alpha = \mathbf{0}$ and $\boldsymbol{\lambda} \neq \mathbf{0}$). Second, the true DGP may be a characteristic-only world in which the characteristics itself explain expected returns and covariances to common risk factors at most explain short-term return innovations (i.e., $\boldsymbol{\Gamma}_\alpha \neq \mathbf{0}$ 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 $\boldsymbol{\lambda} \neq \mathbf{0}$). Note that for each hypothetical world, we assume that there are systematic innovations in returns, otherwise assets are cross-sectionally uncorrelated. Thus, there are nonzero betas on common risk factors; however these factors do not necessarily have nonzero premia. In the following, we analytically show how IPCA misspecification can bias the results, leading to erroneous conclusions regarding the covariance-characteristics story.

2.1. Beta-eating

We begin by showing how a misspecification of IPCA can lead to erroneous conclusions in a covariance-only world, that is, only covariances to priced factors drive expected returns. The DGP is described 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 DGP can be written as:

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

Denote $\boldsymbol{\Gamma}_{\alpha}^{*}$ as the estimate of the characteristic coefficients for the conditional alpha, obtained from an alpha-only IPCA, that is, an IPCA that estimates no factors but only a conditional alpha. Because the alpha-only IPCA assumes a constant environment (i.e., alpha has no innovations), 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, that is:

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

IPCA can find an alpha-only description of expected returns that satisfies:

$$E_t[r_{i,t+1}] = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}\boldsymbol{\lambda} \widehat{=} \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\alpha}^{*} \tag{8}$$

if

$$\boldsymbol{\Gamma}_{\beta}\boldsymbol{\lambda} \widehat{=} \boldsymbol{\Gamma}_{\alpha}^{*} \tag{9}$$

The linear system of equations is uniquely determined if the l-th element $1, \ldots, L$ 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 $\boldsymbol{\lambda}$. Accordingly, a true covariance-only world can always be transformed into an equivalent indistinguishable characteristic-only world. As a result, the effect 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 covariance-characteristic model. The DGP is presented in equation (10). Note that we assume that there may be (latent) factors that drive systematic return innovations, although they are unpriced, i.e., they have factor premia of zero, and thus do not explain expected returns:

$$r_{i,t+1} = \underbrace{\alpha_{i,t}}_{E_t[r_{i,t+1}]} + \underbrace{\boldsymbol{\beta}_{i,t}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1}}_{\widetilde{r}_{i,t+1}} \tag{10}$$

Inserting equations (2) and (3) into (10), the DGP can be written as

$$r_{i,t+1} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\alpha} + \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}\widetilde{\boldsymbol{f}}_{t+1} + \epsilon_{i,t+1} \tag{11}$$

Denote $\boldsymbol{\Gamma}_{\beta}^{*}$, $\boldsymbol{\lambda}^{*}$, and $\widetilde{\boldsymbol{f}}_{t}^{*}$ as beta-only IPCA estimates of characteristic coefficients for the betas, factor premia, and factor innovations, respectively, a beta-only IPCA can find an equivalent description of equation (11) without alpha:

$$r_{i,t+1} = \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}^{*}\boldsymbol{\lambda}^{*} + \boldsymbol{z}_{i,t}^{'}\boldsymbol{\Gamma}_{\beta}^{*}\widetilde{\boldsymbol{f}}_{t+1}^{*} + \epsilon_{i,t+1} \tag{12}$$

if

$$\boldsymbol{\Gamma}_{\alpha}\widehat{=}\boldsymbol{\Gamma}_{\beta}^{*}\boldsymbol{\lambda}^{*} \tag{13}$$

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

$$\begin{array}{lcccccccc} \Gamma_{\alpha,1} & = & \Gamma_{\beta,1,1}^{*}\lambda_{1}^{*} & + & \Gamma_{\beta,1,2}^{*}\lambda_{2}^{*} & + & \dots & + & \Gamma_{\beta,1,K}^{*}\lambda_{K}^{*} \\ \Gamma_{\alpha,2} & = & \Gamma_{\beta,2,1}^{*}\lambda_{1}^{*} & + & \Gamma_{\beta,2,2}^{*}\lambda_{2}^{*} & + & \dots & + & \Gamma_{\beta,2,K}^{*}\lambda_{K}^{*} \\ & \vdots & & & & & & & \\ \Gamma_{\alpha,L} & = & \Gamma_{\beta,L,1}^{*}\lambda_{1}^{*} & + & \Gamma_{\beta,L,2}^{*}\lambda_{2}^{*} & + & \dots & + & \Gamma_{\beta,L,K}^{*}\lambda_{K}^{*} \end{array} \tag{14}$$

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 $\boldsymbol{\Gamma}_{\beta}^{*}$ are fixed and are therefore no free parameters. Accordingly, there are only K free parameters, i.e., the factor premia $\boldsymbol{\lambda}^{*}$, with K denoting the number of factors.

Depending on the (unknown) rank r^{*} of the system of linear equations and the specification of the IPCA, we can differentiate three cases which 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. 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.

3. Simulation study

After analytically showing under which conditions 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 results. We conclude the simulation study with a discussion on the results and their implications for empirical asset pricing research.

3.1. Simulation data

This section describes our simulation setup in more detail. For a simulation setting that is close to 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. Specifically, we use the empirical U.S. equity data set studied in KPS to calibrate our simulation parameters. We generate asset returns for $N = 4100$ assets and $T = 1200$ periods.[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 by estimating an IPCA with two factors and alpha on the empirical data.

The simulation of asset returns divides into five steps, i.e., 1) simulation of characteristics $\boldsymbol{z}_{i,t}$, 2) simulation of conditional alphas $\alpha_{i,t}$ and betas $\boldsymbol{\beta}_{i,t}$, 3) simulation of factor returns $\boldsymbol{f}_t$, 4) simulation of residual returns $\epsilon_{i,t+1}$, and 5) determination of asset (excess) returns $r_{i,t+1}$ by putting all components together according to equation (4). To set up a simulation most easy and comfortable for the IPCA, we keep the DGPs as simple as possible (e.g., we do not assume vector autoregressive residuals, time-varying factor premia, etc.).

3.1.1. Simulation of characteristics

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

[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 4100 with a maximum history of approximately 1200 months. In Appendix A, we also simulate asset return data that matches the time series and cross section of other developed markets.

[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.

the true betas or not, 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 (15):

$$\boldsymbol{z}_{i,t} = \boldsymbol{z}_{0,i} + \boldsymbol{\Phi}_{1,i}\boldsymbol{z}_{i,t-1} + \boldsymbol{u}_{i,t} \tag{15}$$

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 set up 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 0.999 for all assets $i = 1, \ldots, N$. Characteristic innovations are drawn from a multivariate normal distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}_z$:

$$\boldsymbol{u}_{i,t} \sim N\left(0, \boldsymbol{\Sigma}_z\right) \tag{16}$$

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., $\boldsymbol{s} = \boldsymbol{1}$. The term $\boldsymbol{\Omega}$ denotes an $L \times L$ correlation matrix that we can manipulate to obtain specific properties of our simulated data set. We begin with a simulation for which we assume that characteristics are uncorrelated, therefore $\boldsymbol{\Omega}$ is set to an identity matrix.

After simulating autocorrelated characteristics according to equation (15), 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. We use those eight empirical characteristics that have the largest estimated slopes (in absolute terms), obtained from univariate Fama-MacBeth (FM) regressions with intercept, i.e., (slopes in brackets) 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).[4] Because we assume that characteristics nine and ten are perfect proxies for the betas on two common factors, we calibrate these characteristics with respect to the estimated betas obtained from a two-factor alpha-beta IPCA on empirical data. 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.[5]

[4]We point out again that the actual values of the simulation parameters do not qualitatively affect the results.

[5]Note that these values serve to work with realistic values and do not qualitatively affect the results. In unreported results, we find that the results are robust to the scaling and assumed autocorrelation 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
This table reports the target values used for calibration of the simulated characteristics. The column "Mu" refers to the averages 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 a correlation of zero between the characteristics.

3.1.2. Simulation of alpha and beta

The specification of $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ is central to our simulation because this 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., characteristic-only world), we assume that factor innovations explain systematic return innovations. Otherwise, we could not model a covariance structure between 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 assets.

Table 2 shows the true $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ coefficients of all simulated worlds. Assuming a covariance-only world, 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 on the first and second characteristics are nonzero, while those for the other characteristics are always zero. We obtain calibration parameters for the nonzero entries by running univariate Fama-MacBeth (FM) regressions (with intercept) of the returns of the U.S. stock returns 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 the 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. 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 for 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 $\boldsymbol{\Gamma}_\beta$
This table reports the target 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. $\boldsymbol{\Gamma}_\alpha$ denotes the characteristic coefficients for the anomaly intercept, $\boldsymbol{\Gamma}_{\beta,1}$ are the characteristic coefficients for the beta on the first factor, and $\boldsymbol{\Gamma}_{\beta,2}$ are the characteristic coefficients for the beta on the second factor.

3.1.3. Simulation of factor returns

According to equation (4), factor returns have two components, i.e., 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,[6] 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. Therefore, we can simulate factor returns by drawing them from a multivariate normal distribution according to equation (17):

$$\boldsymbol{f}_{t+1} \sim N\left(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f\right) \tag{17}$$

with $\boldsymbol{f}_{t+1}$ denoting the $K \times 1$ vector of factor returns, $\boldsymbol{\mu}_f$ is a $K \times 1$ vector of factor means that equals the factor premia $\boldsymbol{\lambda}$, and $\boldsymbol{\Sigma}_f$ is the $K \times K$ covariance matrix of the factor innovations.

If we do not allow for priced factors (i.e., characteristic-only world), we set the factor means $\boldsymbol{\mu}_f = \boldsymbol{\lambda} = \boldsymbol{0}$. 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 data set. Specifically, we calculate the time series averages of the factors (i.e., 0.0331 and 0.0145) and use these as calibration values for the simulated factor returns. 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

[6] 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) point out the idea that the actual premium of a factor may depend on past factor innovations, thus leading to time-varying factor premia.

simulated data set are reported in Table 3. Note that we also have factor innovations in the characteristic-only world but factor premia of zero.

	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
This table reports the target values for the factor returns obtained from a two-factor IPCA 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 Residual Returns

We assume that the residual returns $\epsilon_{i,t+1}$ are multivariate normally distributed according to equation (18):

$$\boldsymbol{\epsilon}_{t+1} \sim N\left(\boldsymbol{\mu}_{\epsilon}, \boldsymbol{\Sigma}_{\epsilon}\right) \tag{18}$$

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 obtained all simulated characteristics, conditional alphas and betas, factor returns, and residual returns, we put all components together and calculate asset returns as in 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 $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta} = \left[\hat{\boldsymbol{\Gamma}}_{\alpha}, \hat{\boldsymbol{\Gamma}}_{\beta}\right]$ coefficients as discussed in KPS. Note that we differentiate between simulation parameters and (estimated) IPCA parameters by $(\hat{\bullet})$. For each simulated data set (i.e., covariance-only world, characteristic-only world, and covariance-characteristic world), we fit 1) a beta-only IPCA that estimates only betas and no alpha (i.e., $\hat{\boldsymbol{\Gamma}}_{\alpha} = \mathbf{0}$, $\hat{\boldsymbol{\Gamma}}_{\beta} \neq \mathbf{0}$), 2) an alpha-only IPCA that estimates alpha but no betas (i.e., $\hat{\boldsymbol{\Gamma}}_{\alpha} \neq \mathbf{0}$, $\hat{\boldsymbol{\Gamma}}_{\beta} = \mathbf{0}$), and 3) an "unrestricted" IPCA that estimates both alpha and betas (i.e., $\hat{\boldsymbol{\Gamma}}_{\alpha} \neq \mathbf{0}$, $\hat{\boldsymbol{\Gamma}}_{\beta} \neq \mathbf{0}$). For specifications that estimate betas and thus factors, we estimate models with $K = 1, \ldots, 6$ factors.

IPCA finds estimates $\hat{\boldsymbol{f}}_{t+1}$ and $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$ according to a least squares criterion as described in equation (19):

$$\min_{\boldsymbol{\Gamma}_{\alpha,\beta},\boldsymbol{F}} = \sum_{t=1}^{T-1} \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}_t \boldsymbol{\Gamma}_{\alpha,\beta} \boldsymbol{f}_{t+1}\right)' \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}_t \boldsymbol{\Gamma}_{\alpha,\beta} \boldsymbol{f}_{t+1}\right) \tag{19}$$

with $\boldsymbol{Z}_t$ denoting an $N \times L$ matrix of asset-level characteristics, $\boldsymbol{f}_t$ is a $K \times 1$ vector of contemporaneous factor realizations, and $\boldsymbol{F}$ denotes the $K \times T$ matrix of stacked factor realizations. Note that for simplicity, the factors include a constant factor if IPCA allows for an anomaly alpha. Since there is no closed-form solution for equation (19), KPS propose an alternating least squares algorithm to solve numerically for $\hat{\boldsymbol{f}}_{t+1}$ and $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$. Given an initial estimate for the latent factors, the $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$ coefficients are obtained by a pooled ordinary least squares regression (POLS) of the asset returns on the firm characteristics interacted with the estimated factors (and a constant factor if IPCA allows for alpha):

$$\text{vec}\left(\hat{\boldsymbol{\Gamma}}'_{\alpha,\beta}\right) = \left(\sum_{t=1}^{T-1} \boldsymbol{Z}'_t \boldsymbol{Z}_t \otimes \hat{\boldsymbol{f}}_{t+1} \hat{\boldsymbol{f}}'_{t+1}\right)^{-1} \left(\sum_{t=1}^{T-1} \left[\boldsymbol{Z}_t \otimes \hat{\boldsymbol{f}}_{t+1}\right]' \boldsymbol{r}_{t+1}\right) \tag{20}$$

with $\otimes$ denoting the Kronecker product. For a beta-only IPCA, the estimates $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$ are arranged as an $(LK) \times 1$ vector; therefore, the vector has to be reshaped into an $L \times K$ matrix. However, if IPCA allows for alpha, the vector is of dimension $(L(K+1)) \times 1$ and has 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 $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$ estimates, the (non-constant) factors are updated via cross-sectional regressions of the asset returns in excess of the "anomaly return" on the conditional betas as described in equation (21):

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}'_\beta \boldsymbol{Z}'_t \boldsymbol{Z}_t \hat{\boldsymbol{\Gamma}}_\beta\right)^{-1} \left(\hat{\boldsymbol{\Gamma}}'_\beta \boldsymbol{Z}'_t \left(\boldsymbol{r}_{t+1} - \boldsymbol{Z}'_t \hat{\boldsymbol{\Gamma}}_\alpha\right)\right) \forall t \tag{21}$$

with $\boldsymbol{Z}'_t \hat{\boldsymbol{\Gamma}}_\alpha$ denoting the "anomaly returns". However, if IPCA does not allow for an anomaly alpha, the update rule in equation (21) reduces to cross-sectional regressions of the asset returns on the conditional betas.

KPS suggest to create "characteristic-managed portfolios" for accelerating computational performance and for dealing with the missing-value problem in the return and characteristic matrices of the individual assets. Characteristic-managed portfolios are simply the asset-level characteristics interacted with the corresponding returns as defined in equation (22):

$$\boldsymbol{x}_{t+1} = \frac{\boldsymbol{Z}'_t \boldsymbol{r}_{t+1}}{N_{t+1}} \tag{22}$$

with $\boldsymbol{x}_{t+1}$ denoting an $L \times 1$ vector of characteristic-managed portfolio returns and N_{t+1} denoting the number of non-missing observations at month $t+1$. 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. Similarly, missing values in the matrix of firm characteristics are eliminated by defining a matrix of cross-products of the characteristics:

$$\boldsymbol{W}_t = \frac{\boldsymbol{Z}_t^{'}\boldsymbol{Z}_t}{N_{t+1}} \tag{23}$$

with $\boldsymbol{W}_t$ denoting an $L \times L$ matrix of cross-products of non-missing characteristics. $\boldsymbol{W}_t$ is non-changing during the numerical optimization and therefore replaces $\boldsymbol{Z}_t^{'}\boldsymbol{Z}_t$ in equations (20) and (21). The returns on characteristic-managed portfolios replace $\boldsymbol{Z}_t^{'}\boldsymbol{r}_{t+1}$ such that IPCA estimates for $\hat{\boldsymbol{f}}_{t+1}$ and $\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}$ are obtained by:

$$\text{vec}\left(\hat{\boldsymbol{\Gamma}}_{\alpha,\beta}^{'}\right) = \left(\sum_{t=1}^{T-1} N_{t+1}\left(\boldsymbol{W}_t \otimes \hat{\boldsymbol{f}}_{t+1}\hat{\boldsymbol{f}}_{t+1}^{'}\right)\right)^{-1}\left(\sum_{t=1}^{T-1} N_{t+1}\left[\boldsymbol{x}_{t+1} \otimes \hat{\boldsymbol{f}}_{t+1}\right]\right) \tag{24}$$

and

$$\hat{\boldsymbol{f}}_{t+1} = \left(\hat{\boldsymbol{\Gamma}}_{\beta}^{'}\boldsymbol{W}_t\hat{\boldsymbol{\Gamma}}_{\beta}\right)^{-1}\left(\hat{\boldsymbol{\Gamma}}_{\beta}\left(\boldsymbol{x}_{t+1} - \boldsymbol{W}_t\hat{\boldsymbol{\Gamma}}_{\alpha}\right)\right)\forall t \tag{25}$$

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.

3.2.2. Validation

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

$$R^2_{total} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - \left(\hat{\alpha}_{i,t} + \hat{\boldsymbol{f}}_{t+1}\hat{\boldsymbol{\beta}}_{i,t}\right)\right)^2}{\sum_{i,t} r_{i,t+1}^2} \tag{26}$$

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 $E\left[r_{i,t+1}\right]$:

$$R^2_{pred} = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - \left(\hat{\alpha}_{i,t} + \hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\beta}}_{i,t}\right)\right)^2}{\sum_{i,t} r_{i,t+1}^2} \tag{27}$$

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, we test the null hypothesis:

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

against the alternative:

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

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

$$W_\alpha = \hat{\mathbf{\Gamma}}_\alpha' \hat{\mathbf{\Gamma}}_\alpha \tag{28}$$

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 (29):

$$\boldsymbol{x}_t^b = \boldsymbol{W}_t \hat{\mathbf{\Gamma}}_\beta \boldsymbol{f}_{t+1} + \widetilde{\boldsymbol{d}}_{t+1} \tag{29}$$

with $\boldsymbol{x}_t^b$ denoting an $L \times 1$ vector of resampled characteristic-managed portfolio returns of bootstrap $b = 1, \ldots, 1000$. The first part of equation (29) are the fitted returns obtained from the 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.[7] Using the bootstrapped samples, they re-estimate the alpha-beta IPCA and calculate the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha^b$ as described in equation (28). The p-values are calculated as the fraction of bootstrapped W_α^b statistics that exceed the value of W_α from the original data set. 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

3.3.1. Covariance-only world

In our first simulation setup, we assume that asset returns are fully described by covariances to (two) priced risk factors and no anomaly alpha exists. As shown in Table 3, the true factor premia are $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$, respectively, and the corresponding $\mathbf{\Gamma}_\beta$ coefficients to $\boldsymbol{Z}_9$ and $\boldsymbol{Z}_{10}$ are one. Accordingly, the asset pricing test should identify the beta-only IPCA as the true model and the IPCA should return appropriate estimates of the factor premia, $\mathbf{\Gamma}_\alpha$, and $\mathbf{\Gamma}_\beta$ coefficients.

In Table 4 we report the performance evaluation metrics as defined in equations (26) and (27) for individual asset returns r_t (Panel A) and characteristic-managed portfolio returns x_t (Panel B). We also

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

report the sum of squared $\hat{\mathbf{\Gamma}}_\alpha$ coefficients, denoted as W_α, and the corresponding bootstrapped p-values of the asset pricing test for testing the null hypothesis of an anomaly intercepts 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 1 percent level if IPCA estimates at least two factors (p-value = 23.0), suggesting that the test correctly identifies the true model, if IPCA estimates at least as many factors as there are in the true model.

The result of the asset pricing test that there is a significant anomaly alpha if IPCA estimates only one factor indicates an "beta-eating" effect, as shown in Section 2.1. The effect is also reflected in the performance measures presented in Panels A and B. The correctly specified two-factor beta-only IPCA achieves a R^2_{total} of 13.9 percent and a R^2_{pred} of 0.31 percent at the asset-level. Not surprisingly, the one-factor beta-only IPCA has lower model performance, as a result of the omitted factor. However, estimating only one factor and additionally allowing for alpha increases the R^2_{pred} to 0.31 percent, which equals that of the true model. Thus, IPCA compensates for the omitted factor by estimating a non-zero anomaly intercept. Even if IPCA estimates no factors (i.e., alpha-only IPCA), IPCA achieves the same R^2_{pred} suggesting that alpha is sufficient for describing expected returns. From the analytical analysis in Section 2.1, these findings were to be expected.

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 4: 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 5 reports the estimated factor premia (i.e., the time series average of the estimated factors) and Table 6 reports the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients. If a factor premium significantly differs from zero at the 1 percent level, we print the respective premium in bold letters. 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 findings that the IPCA identifies the true model, if IPCA is correctly specified.[8]

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

Table 5: 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 print the premium in bold letters.

[8] According to the findings in Table 4, including additional factors does not affect IPCA's performance. However, estimating more factors makes the interpretation of the model more difficult. First, the influence of a truly relevant characteristic seems to be divided among several betas. Considering the results for the three-factor IPCA, the characteristic nine has large $\hat{\mathbf{\Gamma}}_\beta$ coefficients for the first and second beta of 0.75 and 0.66, respectively. Thus, the estimation results are not robust to a larger number of factors. Second, IPCA estimates "pseudo-betas" (and thus "pseudo-factors") when including more factors than exist in the true model. For example, in the three-factor beta-only IPCA, the fourth characteristic has a large $\hat{\mathbf{\Gamma}}_\beta$ coefficient of 0.66 on the second beta. Since the corresponding factor premium significantly differs from zero, it appears that the fourth characteristic is relevant for explaining asset returns. Although these issues do not affect the model performance, they may make model interpretation more difficult and possibly hinder the identification of the true model.

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 6: Estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients (covariance-only world)

This 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.

The results in Table 4 show 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. Table 7 presents the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only ICPA. In the alpha-only IPCA, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients to these characteristics are very close to the calibrated factor premia presented in Table 3. Specifically, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient on characteristic nine is 0.0336 and the coefficient on characteristic ten is 0.0138. This is because we assigned $\mathbf{\Gamma}_\beta$ coefficients of one to these characteristics. 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 not the covariance among the assets. Even in the misspecified one-factor unrestricted IPCA, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients to characteristics nine and ten are nonzero (i.e., 0.0114 and 0.0235). Thus, IPCA assigns components of both characteristics (i.e., 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 7: Estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients (covariance-only world)
This 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.

3.3.2. Characteristic-only world

Next, 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 performance measures and the results of the asset pricing test are reported in Table 8.

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 $\hat{\mathbf{\Gamma}}_\alpha$ 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 1 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. There we showed that there is a beta-only explanation for expected returns once the number of estimated factors equals or exceeds the number of characteristics. If the number of factors is below the number of characteristics, then no solution exists and IPCA only finds a least-squares approximation. However, this least-squares approximation is already sufficient for the asset pricing test to fail to reject the null hypothesis of a zero anomaly alpha. The decreasing test statistic W_α indicates that the approximation improves with increasing number of factors. In the extreme case $K = L$ (unreported) the alpha disappears entirely.

The true IPCA (i.e., two factors and unrestricted) achieves a R^2_{total} of 13.3 percent and a 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 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 8: 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 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.

Again, we analyze the IPCA results in more detail by reporting the estimated factor premia, 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 the characteristics nine and ten, and $\hat{\mathbf{\Gamma}}_\alpha$ coefficients of 0.0680 and -0.0543 for the 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 for both the beta-only and alpha-beta IPCA. However, the factor premium of the second factor in the alpha-beta IPCA is statistically significant at the 1 percent level but small in magnitude; 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 here indicates the existence of the "alpha-eating" effect. When estimating three (or more) factors in the alpha-beta IPCA, the most factor premia become distinguishable from zero, supporting the finding that IPCA constructs a covariance-based explanation for asset returns.

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

Table 9: 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 print the premium in bold letters.

Table 10 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 to 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, the 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 8. As described in Section 2.2, this is because the system of equations is underdetermined with $L = 10$ equations and only two optimizable parameters. In the three-factor beta-only IPCA, the system is still underdetermined but the least squares approximation is sufficient for filtering the relevant characteristics out of the noisy characteristics, resulting in a good description of expected returns.

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 10: $\hat{\mathbf{\Gamma}}_\beta$ coefficients (characteristic-only world)
This 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.

Table 11 reports the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only IPCA. We note that the true coefficients for the first and second characteristic 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. However, there is large decrease in the magnitude of the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient of the first characteristic when including a third factor.[9] This is in line with the findings from Table 10 that reports increasing $\hat{\mathbf{\Gamma}}_\beta$ 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 11: $\hat{\mathbf{\Gamma}}_\alpha$ coefficients (characteristic-only world)
This table reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ 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.

3.3.3. Covariance-characteristic world

In the previous analyses, we showed that there is an "alpha-eating" effect in a characteristic-only world if IPCA omits the alpha. The "alpha-eating' effect is also present when estimating an alpha-beta IPCA with more factors than actually exist in the true model. However, if the IPCA includes an intercept but estimates fewer factors, the alpha eats beta. Table 12 repeats our analyses for a hybrid world in which both covariances and characteristics drive expected asset returns. Accordingly, the true model is the unrestricted IPCA with two priced factors.

The results of the asset pricing test reported in Table 12 suggest that the null hypothesis for the one- and two-factor models can be rejected at the 1 percent level (p-value = 0.00). However, as in the characteristic-only world, including additional factors changes the results. The null hypothesis cannot be rejected at any conventional significance level if IPCA estimates at least three factors (p-value = 29.1), suggesting that the characteristic-return relationship is described by covariances, not by the characteristics themselves. This finding supports the "alpha-eating" effect we derive in Section 2.2 and also find in Section 3.3.2.

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

The three-factor beta-only IPCA achieves the same performance as the true two-factor alpha-beta IPCA. However, only including alpha and no factors yields the same R^2_{pred} of 0.83 percent as for the true model, suggesting once more that alpha is sufficient for modeling expected returns.

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 12: 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.

We turn our focus to the estimated factor premia and the $\hat{\mathbf{\Gamma}}_\alpha$ and $\hat{\mathbf{\Gamma}}_\beta$ coefficients. Table 13 reports the estimated factor premia for the beta-only and alpha-beta IPCA, respectively. The estimated premia for the two-factor alpha-beta IPCA are close to those assumed in the true DGP (see Table 3). However, the beta-only IPCA considerably overestimates the factor premia, indicating that the "alpha-eating" effect already begins when estimating two factors. This is also reflected in the corresponding $\hat{\mathbf{\Gamma}}_\beta$ coefficients reported in Table 14. Here, we observe the same as in Table 10, that is, the coefficients of both the first and second characteristic to any beta are nonzero and those of characteristics nine and ten are not close to one. Thus, IPCA cannot model the entire covariance of the returns because the betas eat the alpha.

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

Table 13: Estimated factor premia (covariance-characteristic world)
The table reports the estimated factor premia, 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 14: $\hat{\mathbf{\Gamma}}_\beta$ coefficients (covariance-characteristic world)
This 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.

Table 15 reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for the alpha-beta and alpha-only IPCA specifications for the covariance-characteristic world. From Table 12 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 ones reported in Table 2. The coefficients for the characteristics three to eight are not distinguishable from zero. However, the coefficients on characteristics nine and ten (i.e., 0.0334 and 0.0138) are very close to the true factor premia of $\lambda_1 = 0.0331$ and $\lambda_2 = 0.0145$, supporting the finding that alpha eats beta if IPCA estimates fewer factors than exist in the true model.

Including additional factors, we observe the same declining relationship between the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient of the first characteristic with every additional factor as reported in Table 11. However, the magnitude of the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient of the second characteristic also monotonically declines from two to six factors, although the decrease is smaller than that of the coefficients of the first characteristic.

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 15: $\hat{\mathbf{\Gamma}}_\alpha$ coefficients (covariance-characteristic world)
This 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. 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 alphas. Therefore the conclusion drawn by KPS, Buechner and Kelly (2021), and Kelly et al. (2021) that only covariances explain stock, option, or bond returns may be inappropriate.

From our simulation, we find that the IPCA alpha is always sufficient for modeling expected returns, irrespective of whether alpha determines returns or not. What is striking about the results in KPS is that the R^2_{pred} is always higher for the unrestricted IPCA. In Appendix B, 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 percent for individual stock returns and 2.67 percent for portfolio returns. Thus, the effect that alpha is enough for modeling expected returns is not only a phenomenon from our simulation but can also be observed on empirical data.

There is some indication that the empirical results presented in KPS may be subject to 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 characteristic-only world (Table 8), the R^2_{pred}'s are 0.04 percent for the one-factor, 0.20 percent for the two-factor, and 0.35 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 portfolios, 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 have 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 the IPCA to distinguish between covariances and characteristics, therefore, the question of whether covariances, characteristics, or both drive asset returns remains unanswered. Although IPCA does not give an answer to the debate, it is a powerful tool for estimating returns. Even if misspecified, IPCA always finds a model with explanatory power for returns that is equivalent to that of the true model.

Our simulation results are based on a single setup and one might argue that these findings only hold for that particular experimental setup. In the appendix, we therefore present robustness analyses in which we change the sample size, assume correlated characteristics and betas, and decrease and increase the anomaly returns. We find that the results are qualitatively unchanged, that is, the "alpha-eating" and "beta-eating" effects still occur.

4. Conclusion

Kelly et al. (2019), Buechner and Kelly (2021), and Kelly et al. (2021) perform IPCA on stock, option, and bond return data, respectively, and conclude that returns are determined by covariances with common risk factors and no "anomalous returns" due to characteristics exist. These findings are a response to the

[10] In Buechner and Kelly (2021), 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 ($R^2_{pred} = 6.19$).

ongoing "covariance vs. characteristic" debate that has been extensively discussed in the literature over the last thirty years. 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 confirm this finding by simulating data according to a covariance-only world, a characteristic-only world, and a covariance-characteristic world.

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. Thus, the asset pricing test proposed by KPS can correctly differentiate between covariances and characteristics. However, if the IPCA is misspecified, then the IPCA will find a model that has the same performance, but the interpretation becomes more difficult or even the wrong inferences are drawn. For example, omitting factors and allowing for an anomaly alpha results in a "beta-eating" effect, meaning that alpha subsumes 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 look like factor premia, resulting in the rejection of the characteristic story. KPS and Buechner and Kelly (2021) 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.

However, our simulation study cannot help to answer the question of whether covariances or characteristics explain asset returns. Due to the "beta-eating" effect, alpha is always enough for modeling expected returns, independent of the number of factors included. However, these superior fits that we also observe for real-world data do not indicate that asset returns are driven by anomaly returns as shown in our simulation. The "beta-eating" effect will always subsume explanatory power of omitted betas. By contrast, the "alpha-eating" effect will always create a covariance-based explanation of returns if IPCA estimates a sufficient number of factors. Thus, both effects distort IPCA's results.

References

Buechner, M., Kelly, B., 2021. A factor model for option returns doi:10.3386/w29369.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. The Journal of Finance 52, 1–33. doi:10.1111/j.1540-6261.1997.tb03806.x.

Daniel, K., Titman, S., Wei, K.J., 2001. Explaining the cross-section of stock returns in japan: Factors or characteristics? Journal of Finance 56, 743–766. doi:10.1111/0022-1082.00344.

Davis, J.L., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55, 389–406. doi:10.1111/0022-1082.00209.

De Bondt, W.F.M., Thaler, R.H., 1987. Further evidence on investor overreaction and stock market seasonality. Journal of Finance 42, 557–581. doi:10.1111/j.1540-6261.1987.tb04569.x.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56. doi:10.1016/0304-405X(93)90023-5.

Gonçalves, S., Kilian, L., 2004. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. Journal of Econometrics 123, 89–120. doi:10.1016/j.jeconom.2003.10.030.

Green, J., Hand, J.R.M., Zhang, X.F., 2017. The characteristics that provide independent information about average u.s. monthly stock returns. Review of Financial Studies 30, 4389–4436. doi:10.1093/rfs/hhx019.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Review of Financial Studies 33, 2223–2273. doi:10.1093/rfs/hhaa009.

Harvey, C.R., Liu, Y., Zhu, H., 2016. . . . and the cross-section of expected returns. Review of Financial Studies 29, 5–68. doi:10.1093/rfs/hhv059.

Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. Review of Financial Studies 33, 2019–2133. doi:10.1093/rfs/hhy131.

Ilmanen, A.S., Israel, R., Moskowitz, T.J., Thapar, A.K., Wang, F., 2021. Factor premia and factor timing: A century of evidence. SSRN Electronic Journal doi:10.2139/ssrn.3400998.

Jacobs, H., Müller, S., 2020. Anomalies across the globe: Once public, no longer existent? Journal of Financial Economics 135, 213–230. doi:10.1016/j.jfineco.2019.06.004.

Jensen, T.I., Kelly, B.T., Pedersen, L.H., 2021. Is there a replication crisis in finance? SSRN Electronic Journal doi:10.2139/ssrn.3774514.

Kelly, B., Pruitt, S., Su, Y., 2020. Instrumented principal component analysis .

Kelly, B.T., Palhares, D., Pruitt, S., 2021. Modeling corporate bond returns. SSRN Electronic Journal doi:10.2139/ssrn.3720789.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524. doi:10.1016/j.jfineco.2019.05.001.

Kothari, S.P., Shanken, J., Sloan, R.G., 1995. Another look at the cross-section of expected stock returns. The Journal of Finance 50, 185–224. doi:10.2307/2329243.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. The Journal of Finance 49, 1541–1578. doi:10.1111/j.1540-6261.1994.tb04772.x.

Lewellen, J., 2015. The cross-section of expected stock returns. Critical Finance Review 4, 1–44. doi:10.1561/104.00000024.

Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–467. doi:10.1093/rfs/3.3.431.

McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? The Journal of Finance 71, 5–32. doi:10.1111/jofi.12365.

Appendix A. Robustness checks

The main results show that IPCA favors a covariance-only explanation for asset returns if IPCA estimates a sufficient number of factors. However, the findings are based on a simple DGP with a time series and cross section that matches that of the U.S. In this section, we change some settings of the initial simulation setup to investigate the behavior of the IPCA when 1) varying the sample size, 2) allowing for correlated characteristics, and 3) decreasing/increasing the anomaly returns.

Appendix A.1. Varying sample size

Most asset pricing studies are performed on the U.S. equity market, which is characterized by a large time series and cross section. In our initial simulation setup, we study IPCA's performance and ability to distinguish between covariances and characteristics in an environment that matches the U.S. equity market. In this analysis, we examine the extent to which the size of the time series or cross section changes the results. To examine how decreasing the size of the time series dimension affects the estimation results, we simulate data according to the Chinese equity market, which is characterized by a large cross section similar to that of the U.S. equity market but has a much smaller time series ($N = 3600$ and $T = 400$). We examine the extent to which the results change in dependence of decreasing the cross section by simulating data according to the British (i.e., $N = 1700$, $T = 400$) and Spanish equity markets (i.e., $N = 200$, $T = 400$).[11] All other simulation parameters are unchanged.

To save space, we only report the bootstrapped p-values of the asset pricing test for testing the null hypothesis of a zero anomaly intercept against the alternative of a nonzero intercept. Panel A of Table A.16 reports the results for the "China-like" market. Assuming a covariance-only world, the null hypothesis of a zero anomaly intercept is rejected at the 1 percent level if IPCA estimates only one factor (p-value = 0.70). However, estimating at least two factors, the test shows that the null hypothesis cannot be rejected, suggesting that anomalous returns do not exist. Hence, the test correctly identifies the true model. This is in line with the results from our initial setup presented in Table 4. Assuming either a characteristic-only or a covariance-characteristic world, the test correctly rejects the null hypothesis if IPCA estimates at most two factors. However, as in our initial simulation setup presented in Table 8 and Table 12, respectively, the "alpha-eating" effect occurs so that the betas eat the alphas. As a consequence, ICPA cannot distinguish between covariances and characteristics if at least three factors are included, resulting to an erroneous acceptance of the zero anomaly return hypothesis. In summary, the results do not change by reducing the time series dimension, provided the cross section remains large. Therefore, we next analyze how the reduction of the size of the cross section affects the results.

[11] All numbers are based on those presented in the Jensen et al. (2021).

Panel B presents the results for the "Britain-like" market, which is characterized by a time series as in the Chinese market but a cross section of half the size of the cross section in China. Assuming a covariance-only world, the asset pricing test cannot reject the null hypothesis at the 1 percent level for any number of the IPCA factors, suggesting that anomalous returns do not exist. This is consistent with the underlying DGP. For both the covariance-only and the covariance-characteristic worlds, the asset pricing test correctly rejects the null hypothesis for the one- and two-factor model but not for models with additional factors. These results suggest that both the "beta-eating" and "alpha-eating" effect are also existent here, with the "alpha-eating" effect being more apparent.

Next, we further reduce the size of the cross section and simulate returns according to a "Spain-like" market. The results are reported in Panel C. In the covariance-only world, the "beta-eating" effect we observed for the one-factor IPCA does not seem to occur here. The asset pricing test cannot reject the null hypothesis (p-value = 43.9), suggesting that the true covariance-only world is correctly identified. In the characteristic-only world, the null hypothesis is rejected at the 1 percent level (p-value = 0.40) for the one-factor model but not for the two- (or more) factor model (p-value = 7.40). This shows that IPCA has even more problems to identify the true model when modeling a very small cross section. The results for the covariance-characteristic world show more clearly that IPCA has problems with modeling a small cross-section because the test can never identify the true model.

K	1	2	3	4	5	6
Panel A: China ($N = 3600$, $T = 400$)						
Covariance-Only	0.70	94.7	85.7	78.4	96.4	85.2
Characteristic-Only	0.00	0.00	28.9	21.8	42.3	65.4
Covariance-Characteristic	0.00	0.00	11.9	34.0	14.0	2.50
Panel B: Great Britain ($N = 1700$, $T = 400$)						
Covariance-Only	2.80	17.5	52.9	75.4	56.7	33.3
Characteristic-Only	0.00	0.10	17.8	87.2	96.9	40.5
Covariance-Characteristic	0.00	0.10	13.7	5.40	58.1	68.0
Panel B: Spain ($N = 200$, $T = 400$)						
Covariance-Only	43.9	59.6	24.5	20.5	8.60	67.3
Characteristic-Only	0.40	7.40	99.9	98.6	69.2	11.5
Covariance-Characteristic	11.9	47.6	98.0	93.9	82.5	10.5

Table A.16: Varying sample size
This table reports the bootstrapped p-values of the asset pricing test suggested by KPS for asset return data simulated according to three different DGPs, i.e., covariance-only world, characteristic-only world, and covariance characteristic world using different sample sizes. The sample sizes are chosen according to real-world equity markets, i.e., China, Great Britain, and Spain. All p-values are reported in percent.

To examine these results in more detail, Table A.17 reports the model performance of IPCA in a covariance-characteristic "Spain-like" world. Once again, we find that only allowing for an IPCA alpha is sufficient for modeling expected returns, as indicated by the high R^2_{pred} of 0.81 for the alpha-only IPCA. However, the true model, i.e., alpha-beta IPCA with two factors, achieves a R^2_{total} of 15.2 percent which is lower than that of, for example, the six-factor alpha-beta IPCA of 17 percent. This finding suggests, that IPCA requires more factors than there are in the true model to capture the entire covariation of the asset returns.

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}	14.1	15.0	15.6	16.1	16.6	17.0	14.4	15.2	15.7	16.2	16.6	17.0	0.81
R^2_{pred}	0.50	0.72	0.80	0.80	0.80	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
Panel B: Managed portfolios (x_t)													
R^2_{total}	99.1	99.3	99.4	99.5	99.6	99.7	99.2	99.3	99.4	99.5	99.6	99.7	3.47
R^2_{pred}	3.39	3.44	3.48	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47
Panel C: Asset pricing test													
W_α							0.51	0.54	0.28	0.27	0.22	0.22	
p-value							11.9	47.6	98.0	93.9	82.5	10.5	

Table A.17: IPCA model performance for a "Spain-like" world (covariance-characteristic world)
Panel A and B report the performance evaluation metrics R^2_{total} and R^2_{pred} for the "Spain-like" simulated data set for both individual assets (r_t) and characteristic-managed portfolios (x_t), respectively. 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.

The corresponding $\hat{\mathbf{\Gamma}}_\beta$ and $\hat{\mathbf{\Gamma}}_\alpha$ coefficients are presented in Table A.18 and Table A.19, respectively. According to the true DGP, the $\mathbf{\Gamma}_\beta$ coefficients to characteristic nine and ten should be one. However, the $\hat{\mathbf{\Gamma}}_\beta$ estimates for the true two-factor alpha-beta IPCA are far from one, or the influence of a true beta is not clearly attributable to one IPCA beta. The ninth characteristic has coefficients of 0.46 and 0.80 for the first and second IPCA beta, respectively. Characteristic ten has $\hat{\mathbf{\Gamma}}_\beta$ coefficients of 0.56 and -0.23. Therefore, IPCA seems to fail to model the true betas, which may explain the relatively low R^2_{total} of the correctly specified IPCA in Table A.17.

IPCA	Beta						Alpha-Beta					
K	1	2	3	4	5	6	1	2	3	4	5	6
$\hat{\Gamma}_{\beta,1,1}$	-0.15	0.29	0.41	0.08	0.14	0.18	-0.19	-0.05	0.58	0.18	0.17	-0.06
$\hat{\Gamma}_{\beta,2,1}$	-0.04	-0.40	-0.20	-0.03	-0.02	-0.01	0.04	-0.07	-0.11	-0.03	-0.04	-0.01
$\hat{\Gamma}_{\beta,3,1}$	-0.04	0.28	-0.50	-0.42	-0.61	-0.41	-0.08	0.50	-0.53	-0.48	-0.50	-0.31
$\hat{\Gamma}_{\beta,4,1}$	-0.60	-0.26	-0.64	-0.89	-0.77	-0.88	-0.60	0.46	-0.55	-0.85	-0.84	-0.94
$\hat{\Gamma}_{\beta,5,1}$	-0.12	0.10	0.15	0.04	0.02	0.09	-0.12	0.00	0.20	0.07	0.07	0.07
$\hat{\Gamma}_{\beta,6,1}$	-0.01	0.02	0.06	-0.09	-0.05	-0.02	-0.01	-0.01	0.11	-0.06	-0.05	-0.02
$\hat{\Gamma}_{\beta,7,1}$	0.03	0.04	0.00	0.02	0.01	0.01	0.02	0.02	-0.02	0.01	0.01	0.04
$\hat{\Gamma}_{\beta,8,1}$	0.00	-0.00	0.03	-0.01	0.01	0.00	0.00	-0.03	0.03	0.00	0.00	0.01
$\hat{\Gamma}_{\beta,9,1}$	0.73	0.67	-0.17	0.02	0.04	0.03	0.71	0.46	-0.03	0.03	0.03	0.03
$\hat{\Gamma}_{\beta,10,1}$	-0.25	0.38	-0.27	-0.09	-0.07	-0.06	-0.28	0.56	-0.10	-0.08	-0.08	-0.05
$\hat{\Gamma}_{\beta,1,2}$		0.35	0.23	-0.43	-0.43	0.00		-0.20	0.22	0.75	0.02	-0.51
$\hat{\Gamma}_{\beta,2,2}$		-0.18	-0.40	0.07	-0.04	0.00		0.03	-0.15	-0.16	-0.03	0.05
$\hat{\Gamma}_{\beta,3,2}$		0.22	0.29	0.36	0.64	0.88		-0.04	0.25	-0.32	0.84	0.82
$\hat{\Gamma}_{\beta,4,2}$		0.50	-0.27	-0.22	-0.56	-0.39		-0.50	-0.07	0.33	-0.47	-0.24
$\hat{\Gamma}_{\beta,5,2}$		0.19	0.08	-0.14	0.05	0.23		-0.13	0.08	0.25	0.25	0.05
$\hat{\Gamma}_{\beta,6,2}$		0.02	0.02	-0.14	-0.24	0.02		-0.01	0.05	0.28	-0.03	0.05
$\hat{\Gamma}_{\beta,7,2}$		-0.01	0.04	0.04	0.07	0.04		0.02	0.01	-0.07	0.03	0.06
$\hat{\Gamma}_{\beta,8,2}$		-0.01	-0.01	-0.04	-0.08	-0.05		0.00	-0.01	0.06	-0.06	-0.02
$\hat{\Gamma}_{\beta,9,2}$		-0.49	0.70	0.71	0.14	0.01		0.80	0.78	-0.24	0.02	0.01
$\hat{\Gamma}_{\beta,10,2}$		0.52	0.37	0.30	-0.02	0.05		-0.23	0.49	0.01	0.03	0.05
$\hat{\Gamma}_{\beta,1,3}$			0.49	0.62	0.33	0.79			-0.46	0.25	0.81	0.79
$\hat{\Gamma}_{\beta,2,3}$			-0.29	-0.57	-0.35	-0.12			0.13	-0.16	-0.19	-0.15
$\hat{\Gamma}_{\beta,3,3}$			0.11	0.18	0.08	-0.01			-0.09	0.22	-0.01	0.43
$\hat{\Gamma}_{\beta,4,3}$			0.32	-0.04	-0.02	0.17			-0.44	-0.08	0.17	-0.17
$\hat{\Gamma}_{\beta,5,3}$			0.25	0.24	0.04	0.31			-0.22	0.11	0.35	0.37
$\hat{\Gamma}_{\beta,6,3}$			0.04	0.13	0.12	0.38			-0.06	0.08	0.31	0.02
$\hat{\Gamma}_{\beta,7,3}$			-0.01	0.02	0.02	-0.04			0.02	-0.00	-0.05	-0.03
$\hat{\Gamma}_{\beta,8,3}$			0.00	0.02	0.02	0.06			-0.00	-0.00	0.05	-0.04
$\hat{\Gamma}_{\beta,9,3}$			-0.52	0.25	0.71	-0.27			0.50	0.76	-0.23	-0.04
$\hat{\Gamma}_{\beta,10,3}$			0.48	0.35	0.49	0.09			-0.52	0.52	0.07	0.03
$\hat{\Gamma}_{\beta,1,4}$				0.12	0.50	0.16				-0.00	0.14	0.05
$\hat{\Gamma}_{\beta,2,4}$				-0.11	-0.49	-0.22				-0.05	-0.13	-0.09
$\hat{\Gamma}_{\beta,3,4}$				-0.13	0.28	-0.00				0.17	-0.02	0.01
$\hat{\Gamma}_{\beta,4,4}$				0.14	-0.14	0.03				-0.16	0.03	-0.01
$\hat{\Gamma}_{\beta,5,4}$				-0.09	0.37	-0.07				0.11	-0.04	-0.10
$\hat{\Gamma}_{\beta,6,4}$				0.25	0.05	0.11				-0.23	0.10	-0.00
$\hat{\Gamma}_{\beta,7,4}$				-0.02	0.02	0.01				0.03	-0.00	0.00
$\hat{\Gamma}_{\beta,8,4}$				0.07	-0.00	0.02				-0.06	0.02	-0.01
$\hat{\Gamma}_{\beta,9,4}$				0.55	-0.53	0.82				-0.57	0.85	0.88
$\hat{\Gamma}_{\beta,10,4}$				-0.75	-0.03	0.48				0.74	0.49	0.45
$\hat{\Gamma}_{\beta,1,5}$					0.10	-0.06					-0.09	-0.11
$\hat{\Gamma}_{\beta,2,5}$					-0.20	-0.84					-0.04	-0.18
$\hat{\Gamma}_{\beta,3,5}$					0.09	-0.02					-0.07	-0.14
$\hat{\Gamma}_{\beta,4,5}$					0.04	-0.01					-0.05	0.02
$\hat{\Gamma}_{\beta,5,5}$					-0.00	0.02					0.01	0.24
$\hat{\Gamma}_{\beta,6,5}$					0.25	-0.39					-0.28	0.70
$\hat{\Gamma}_{\beta,7,5}$					-0.01	0.09					0.03	0.03
$\hat{\Gamma}_{\beta,8,5}$					0.06	-0.05					-0.06	0.17
$\hat{\Gamma}_{\beta,9,5}$					0.39	-0.29					-0.44	-0.25
$\hat{\Gamma}_{\beta,10,5}$					-0.85	0.22					0.84	0.54
$\hat{\Gamma}_{\beta,1,6}$						0.07						0.01
$\hat{\Gamma}_{\beta,2,6}$						-0.40						-0.02
$\hat{\Gamma}_{\beta,3,6}$						0.07						0.02
$\hat{\Gamma}_{\beta,4,6}$						0.05						0.04
$\hat{\Gamma}_{\beta,5,6}$						-0.01						0.05
$\hat{\Gamma}_{\beta,6,6}$						0.14						0.61
$\hat{\Gamma}_{\beta,7,6}$						0.03						-0.03
$\hat{\Gamma}_{\beta,8,6}$						0.05						0.14
$\hat{\Gamma}_{\beta,9,6}$						0.34						0.36
$\hat{\Gamma}_{\beta,10,6}$						-0.83						-0.69

Table A.18: $\hat{\mathbf{\Gamma}}_\beta$ coefficients for a "Spain-like" world (covariance-characteristic world)
This table reports the estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients for for the beta-only and alpha-beta (i.e., unrestricted) IPCA, respectively, using a "Spain-like" simulated 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.

Even the examination of the $\hat{\mathbf{\Gamma}}_\alpha$ vectors in Table A.19 shows that IPCA cannot identify the correct model. As a reminder, according to Table 2, the true $\mathbf{\Gamma}_\alpha$ coefficients of the first and second characteristics are 0.0680 and -0.0547, respectively, while the coefficients of all other characteristics are zero. However, IPCA systematically underestimates the coefficients of the first characteristic, so that, for example, in the correctly specified IPCA the coefficient is estimated as only half as large (i.e., $\hat{\Gamma}_{\alpha,1} = 0.0333$) as in the true model. At the same time, the $\hat{\mathbf{\Gamma}}_\alpha$ coefficient on the fourth characteristic, which is expected to be zero, is -0.0296, although characteristic one and four are uncorrelated (the average correlation is only 0.0053). We qualitatively observe the same for the alpha-only IPCA.

IPCA			Alpha-Beta				Alpha
K	1	2	3	4	5	6	0
$\hat{\Gamma}_{\alpha,1}$	0.0293	0.0333	-0.0031	0.0038	-0.0020	-0.0029	0.0257
$\hat{\Gamma}_{\alpha,2}$	-0.0555	-0.0546	-0.0440	-0.0441	-0.0437	-0.0441	-0.0541
$\hat{\Gamma}_{\alpha,3}$	0.0158	0.0078	0.0189	0.0194	0.0022	0.0019	0.0063
$\hat{\Gamma}_{\alpha,4}$	-0.0128	-0.0296	-0.0147	-0.0077	-0.0001	-0.0003	-0.0320
$\hat{\Gamma}_{\alpha,5}$	0.0063	0.0074	-0.0066	-0.0061	-0.0135	-0.0130	0.0042
$\hat{\Gamma}_{\alpha,6}$	0.0020	0.0024	-0.0040	-0.0018	-0.0060	-0.0043	0.0026
$\hat{\Gamma}_{\alpha,7}$	0.0065	0.0063	0.0076	0.0071	0.0070	0.0069	0.0069
$\hat{\Gamma}_{\alpha,8}$	0.0004	0.0009	-0.0002	0.0004	0.0006	0.0013	0.0006
$\hat{\Gamma}_{\alpha,9}$	0.0117	-0.0024	-0.0105	-0.0064	-0.0029	-0.0031	0.0302
$\hat{\Gamma}_{\alpha,10}$	0.0222	0.0152	-0.0055	-0.0141	-0.0057	-0.0052	0.0143

Table A.19: $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for a "Spain-like" world (covariance-characteristic world)
This table reports the estimated $\hat{\mathbf{\Gamma}}_\alpha$ coefficients for each estimated IPCA, i.e., beta-only ("Beta"), unrestricted ("Alpha-Beta"), and alpha-only ("Alpha") using a "Spain-like" simulated data set. The term $\hat{\Gamma}_{\alpha,l}$ denotes the estimated coefficient of the l-th characteristic for the alpha.

To summarize, decreasing the size of the time series does not much affect the results. However, the problems that ICPA and hence the asset pricing test have in identifying the true model become more pronounced as the cross-section size decreases.

Appendix A.2. Correlated Characteristics

In our initial simulation setup, we assume that the characteristics are mutually uncorrelated. However, this assumption may be oversimplified, therefore we deviate from the assumption and allow for significant correlations. Specifically, we are interested in the results of the IPCA in a covariance-characteristic world in which a beta is highly correlated with a characteristic that proxies for alpha.[12]

Table A.20 presents the bootstrapped p-values of the asset pricing test for different correlation coefficients of characteristics 1 and 10 ($\rho_{1,10}$). To study the behavior of the IPCA at different correlation levels, we successively increase the correlation coefficient between the characteristics from 0 to 0.9. We find that,

[12] In unreported results, we also tested for high correlations between a beta and an irrelevant characteristic and find that the results are qualitatively unchanged.

regardless of the level of correlation, the null hypothesis is correctly rejected at the 1 percent level if IPCA estimates at most two factors. In our initial setup (i.e., $\rho_{1,10} = 0$), we find that the null hypothesis cannot be rejected if IPCA estimates three factors ($p = 29.1$). However, allowing for correlation between beta and one characteristic, we observe a monotonically declining p-value when increasing the correlation coefficient. With four estimated factors, we observe the same monotonic decreasing p-values with increasing correlations, however the null hypothesis can never be rejected here. The situation is different for the five- and six-factor models. If the correlations are extremely high (i.e., $\rho_{1,10} \geq 0.7$), then the IPCA recognizes that there is significant anomaly alpha. To examine this in more detail, we report the associated test statistics, i.e., the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$, in Table A.21. We find that the test statistic declines with every additional factor, suggesting that the $\hat{\mathbf{\Gamma}}_\alpha$ estimates are too small (note that the sum of squared elements in the true $\mathbf{\Gamma}_\alpha$ vector is 0.76). However, the magnitude of the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients does not differ between the four- and five-factor model but the five-factor IPCA alpha becomes statistically significant. This indicates that although the magnitude of the coefficients is small, the dispersion under the null hypothesis is decreased for the five-factor model, thus making these small alphas statistically significant. Decreasing p-values despite an increasing number of factors can also be observed in Buechner and Kelly (2021). We additionally perform the asset pricing test assuming seven and eight factors, respectively, and find that the null hypothesis cannot be rejected for any correlation level if at least seven factors are included. These results show that even if unrealistically high correlations between betas and characteristics are assumed, the beta always eats up the alpha, if a sufficient large number of factors is estimated.

	K							
$\rho_{1,9}$	1	2	3	4	5	6	7	8
0.0	0.00	0.00	29.1	52.5	15.0	49.9	100.0	99.8
0.1	0.00	0.00	20.5	50.2	11.2	37.3	100.0	99.7
0.2	0.00	0.00	15.0	40.3	7.70	22.6	100.0	100.0
0.3	0.00	0.00	11.6	35.0	6.10	16.2	100.0	99.9
0.4	0.00	0.00	8.80	31.5	3.70	8.20	99.9	99.7
0.5	0.00	0.00	6.60	26.9	3.10	3.80	100.0	100.0
0.6	0.00	0.00	4.60	23.2	1.90	1.50	100.0	100.0
0.7	0.00	0.00	3.20	20.9	1.00	0.50	98.5	99.9
0.8	0.00	0.00	3.10	13.0	1.30	0.20	90.4	99.8
0.9	0.00	0.00	3.20	10.6	0.20	0.00	68.1	99.9

Table A.20: Bootstrapped p-values assuming correlated characteristics (covariance-characteristic world)
This table reports bootstrapped p-values from the asset pricing test for testing the null hypothesis of zero anomaly intercepts. We assume a covariance-characteristic world and vary the correlation coefficients of the first and tenth characteristic from 0 to 0.9. All results are presented in percent.

				K				
$\rho_{1,9}$	1	2	3	4	5	6	7	8
0.0	0.81	0.71	0.61	0.32	0.32	0.17	0.02	0.02
0.1	0.81	0.72	0.64	0.33	0.33	0.19	0.02	0.02
0.2	0.82	0.72	0.68	0.37	0.37	0.24	0.02	0.02
0.3	0.82	0.72	0.68	0.37	0.37	0.24	0.02	0.02
0.4	0.82	0.73	0.69	0.38	0.38	0.25	0.02	0.02
0.5	0.82	0.73	0.71	0.39	0.39	0.27	0.02	0.02
0.6	0.82	0.73	0.72	0.41	0.40	0.30	0.03	0.02
0.7	0.82	0.74	0.73	0.42	0.42	0.34	0.05	0.02
0.8	0.83	0.74	0.74	0.44	0.44	0.39	0.08	0.02
0.9	0.82	0.74	0.74	0.47	0.47	0.44	0.12	0.03

Table A.21: Test statistics assuming correlated characteristics (covariance-characteristic world)
This table reports the test statistics from the asset pricing test for testing the null hypothesis of zero anomaly intercepts. We assume a covariance-characteristic world and vary the correlation coefficients of the first and tenth characteristic from 0 to 0.9. The test statistic W_α is the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$. For representation purposes, we multiply the test statistic by 100.

Appendix A.3. Varying anomaly returns

We calibrated the anomaly returns with respect to univariate Fama-MacBeth regressions of the empirical U.S. returns on the corresponding firm characteristics. However, these anomaly returns may be too high for simulating real-world asset returns. Therefore, we shrink the $\mathbf{\Gamma}_\alpha$ coefficients to 10 percent of the values from our initial setup and simulate asset return data according to a covariance-characteristic world. Thus, the $\mathbf{\Gamma}_\alpha$ coefficients for the first and second characteristics are 0.0068 and -0.0054, respectively. We keep all other simulation parameters as in our initial simulation setup.

Table A.22 presents the results. Again, we observe in Panel C that W_α, i.e., the sum of squared elements in $\hat{\mathbf{\Gamma}}_\alpha$, declines with every additional factor. If IPCA estimates at least two factors, the null hypothesis of zero anomaly intercepts is not rejected. As opposed to the results of our initial setting, the asset pricing test cannot reliably identify the true two-factor model with alpha because the betas already subsumed the alpha when allowing for two factors (p-value = 25.1). Thus, we conclude that the "alpha-eating" effect is more pronounced if the anomaly returns are smaller.

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}	13.5	13.8	13.9	13.9	13.9	13.9	13.5	13.8	13.9	13.9	13.9	13.9	0.33
R^2_{pred}	0.29	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
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	1.99
R^2_{pred}	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99
Panel C: Asset pricing test													
W_α							0.08	0.01	0.01	0.00	0.00	0.00	
p-value							0.00	25.1	66.5	98.9	97.0	91.8	

Table A.22: Reduced $\mathbf{\Gamma}_\alpha$ coefficients (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 assuming smaller true $\mathbf{\Gamma}_\alpha$ coefficients. 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.

Next, we make it easier for IPCA and simulate data according to a setting in which we scale the anomaly returns upwards. For this purpose, we multiply the $\mathbf{\Gamma}_\alpha$ coefficients of our initial setup by 10 and obtain coefficients of 0.6804 and -0.5473. We use unrealistic large values for the $\mathbf{\Gamma}_\alpha$ coefficients to make it as easy as possible for IPCA to identify the anomaly returns. The results for the covariance-characteristic world are reported in Table A.23.

The results of the asset pricing test (Panel C) are qualitatively similar to those presented in Table 12, i.e., the null hypothesis is (correctly) rejected at the 1 percent level if IPCA estimates at most two factors but the null cannot be rejected when allowing for additional factors. The true two-factor alpha-beta IPCA achieves an R^2_{pred} of 27 percent which is also achieved when excluding the intercept (i.e., beta-only). Thus, alpha is not necessary for modeling expect returns, which already suggests that the "alpha-eating" effect is also present and probably more pronounced.

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}	29.1	36.7	36.9	37.0	37.0	37.0	36.7	36.9	36.9	37.0	37.0	37.0	27.0
R^2_{pred}	26.3	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0
Panel B: Managed portfolios (x_t)													
R^2_{total}	62.4	99.9	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	33.2
R^2_{pred}	29.6	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2	33.2
Panel C: Asset pricing test													
W_α							76.1	75.4	65.5	37.3	37.0	18.2	
p-value							0.00	0.60	25.6	39.4	8.00	41.5	

Table A.23: Increased $\mathbf{\Gamma}_\alpha$ coefficients (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 assuming larger true $\mathbf{\Gamma}_\alpha$ coefficients. 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.

Next, we examine the estimated $\hat{\mathbf{\Gamma}}_\beta$ coefficients and the estimated factor premia in Table A.24 and Table A.25, respectively. For the one-factor beta-only IPCA, the $\hat{\mathbf{\Gamma}}_\beta$ coefficients of the first two characteristics are 0.78 and -0.63, respectively, while those to characteristics nine and ten are indistinguishable from zero. The corresponding factor premium is large and statistically significant ($\hat{\lambda}_1 = 0.7867$), suggesting that the "alpha-eating" effect already begins when estimating one factor. The $\hat{\mathbf{\Gamma}}_\beta$ coefficient of characteristic nine increases to 0.91 when allowing for a second factor but the factor premium of the second factor is zero. Irrespective of the number of factors, only the factor premium of the first factor is statistically significant while those of the other factors are not. This finding suggests that the beta-only IPCA fails to capture the expected returns that are attributable to the betas.

IPCA	Beta						Alpha-Beta					
K	1	2	3	4	5	6	1	2	3	4	5	6
B1_Z1	0.78	0.78	0.78	0.78	0.78	0.78	-0.04	0.04	0.11	0.67	0.70	0.80
B1_Z2	-0.63	-0.63	-0.63	-0.63	-0.62	-0.62	-0.01	-0.01	-0.44	-0.30	-0.26	-0.39
B1_Z3	0.00	0.00	0.00	-0.00	0.00	0.00	-0.02	0.00	0.82	0.35	0.36	0.21
B1_Z4	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01	-0.03	0.06	0.11	-0.53	-0.50	-0.13
B1_Z5	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.01	0.01	-0.18	-0.11	-0.12	0.15
B1_Z6	-0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	-0.01	0.12	0.15	0.12	0.29
B1_Z7	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.08	-0.05	-0.05	0.18
B1_Z8	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.05	-0.06	-0.05	-0.02
B1_Z9	0.07	0.04	0.04	0.04	0.04	0.04	0.91	0.85	0.12	0.07	0.07	0.05
B1_Z10	0.00	0.02	0.02	0.02	0.02	0.02	-0.41	0.52	0.19	0.10	0.10	0.04
B2_Z1		-0.04	0.04	-0.06	-0.06	-0.06		0.08	0.02	-0.06	-0.06	0.03
B2_Z2		-0.01	-0.02	-0.02	-0.02	-0.02		0.01	-0.04	-0.01	-0.01	-0.02
B2_Z3		-0.02	0.01	0.01	0.02	0.03		0.02	0.10	0.01	0.02	0.19
B2_Z4		-0.03	-0.01	0.07	0.08	0.05		0.07	0.00	0.06	0.06	0.19
B2_Z5		0.01	-0.01	0.01	0.01	0.00		-0.00	-0.03	0.01	0.01	-0.04
B2_Z6		0.00	0.00	-0.01	-0.01	-0.01		-0.01	0.02	-0.01	-0.01	-0.00
B2_Z7		0.00	0.00	-0.00	-0.00	-0.00		-0.00	-0.01	-0.00	-0.00	0.01
B2_Z8		-0.00	-0.00	0.00	0.00	0.00		0.00	-0.01	0.00	0.00	0.01
B2_Z9		0.91	-1.00	1.00	0.99	1.00		-0.53	-0.99	1.00	1.00	-0.96
B2_Z10		-0.41	0.00	0.00	0.00	-0.00		0.84	0.03	0.00	0.00	0.06
B3_Z1			-0.02	-0.31	-0.20	-0.19			-0.06	-0.23	-0.09	-0.09
B3_Z2			-0.05	-0.41	-0.27	-0.25			-0.11	-0.30	-0.12	-0.10
B3_Z3			-0.02	0.29	0.53	0.52			0.17	0.64	0.77	0.59
B3_Z4			-0.09	0.79	0.75	0.77			-0.07	0.43	0.55	0.72
B3_Z5			-0.00	-0.03	-0.13	-0.12			-0.04	-0.12	-0.22	-0.08
B3_Z6			0.01	-0.08	-0.10	-0.11			0.04	0.04	-0.07	-0.03
B3_Z7			0.00	-0.00	-0.03	-0.03			-0.01	-0.05	-0.06	0.04
B3_Z8			-0.00	0.03	0.03	0.03			-0.01	-0.02	0.00	0.02
B3_Z9			-0.00	-0.09	-0.09	-0.07			-0.01	-0.05	-0.05	0.27
B3_Z10			-0.99	0.08	0.08	0.09			-0.97	0.49	0.17	0.16
B4_Z1				-0.07	-0.13	-0.07				-0.16	-0.02	-0.09
B4_Z2				-0.11	-0.18	-0.10				-0.22	-0.05	-0.14
B4_Z3				0.02	-0.07	0.57				0.42	0.20	0.17
B4_Z4				0.01	0.06	-0.38				0.08	-0.02	0.06
B4_Z5				-0.01	0.04	-0.19				-0.09	-0.07	0.01
B4_Z6				0.00	0.01	0.07				0.06	0.01	0.07
B4_Z7				0.00	0.01	-0.05				-0.03	-0.02	0.03
B4_Z8				0.00	0.01	-0.03				-0.02	-0.01	-0.00
B4_Z9				-0.01	-0.01	-0.00				-0.02	-0.00	-0.02
B4_Z10				-0.99	-0.97	-0.69				-0.85	-0.98	-0.97
B5_Z1					-0.37	-0.17					-0.55	-0.38
B5_Z2					-0.47	-0.23					-0.68	-0.48
B5_Z3					-0.66	-0.55					0.09	-0.37
B5_Z4					0.23	0.36					-0.29	0.23
B5_Z5					0.30	0.21					0.16	0.47
B5_Z6					0.05	-0.03					0.33	0.36
B5_Z7					0.08	0.06					0.00	0.26
B5_Z8					0.03	0.04					-0.07	0.00
B5_Z9					-0.04	-0.02					-0.02	-0.04
B5_Z10					0.21	-0.66					0.06	0.09
B6_Z1						-0.55						-0.40
B6_Z2						-0.68						-0.49
B6_Z3						-0.05						0.44
B6_Z4						-0.32						-0.52
B6_Z5						0.13						-0.21
B6_Z6						0.18						0.12
B6_Z7						0.04						-0.23
B6_Z8						-0.02						-0.09
B6_Z9						-0.03						-0.01
B6_Z10						0.28						0.15

Table A.24: $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients for the increased $\boldsymbol{\Gamma}_\alpha$ setting (covariance-characteristic world)
This table reports the estimated $\hat{\boldsymbol{\Gamma}}_\beta$ coefficients for each estimated IPCA, i.e., beta-only ("Beta"), unrestricted ("Alpha-Beta"), and alpha-only ("Alpha") for our simulation setup with upward scaled $\boldsymbol{\Gamma}_\alpha$ coefficients. The term $\hat{\Gamma}_{\beta,l,j}$ denotes the estimated coefficient of the l-th characteristic for the beta on the j-th factor.

K	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$
Panel A: Beta-Only IPCA						
1	**0.7867**					
2	**0.8719**	0.0000				
3	**0.8718**	0.0005	0.0004			
4	**0.8717**	0.0002	0.0054	0.0009		
5	**0.8717**	0.0002	0.0040	0.0015	0.0030	
6	**0.8718**	0.0002	0.0038	0.0008	0.0018	0.0031
Panel B: Alpha-Beta IPCA						
1	0.0020					
2	**0.0662**	**0.0418**				
3	**0.3244**	0.0034	0.0026			
4	**0.6219**	0.0001	**0.0066**	0.0038		
5	**0.6249**	0.0001	0.0042	0.0009	**0.0037**	
6	**0.7600**	0.0005	0.0024	0.0018	**0.0044**	0.0004

Table A.25: Estimated factor premia for the increased $\mathbf{\Gamma}_\alpha$ setting (covariance-characteristic world)
The table reports the estimated factor premia, 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.

Finally, we shed light on the estimation of the $\hat{\mathbf{\Gamma}}_\alpha$ coefficients in Table A.26. In the true two-factor alpha-beta IPCA, the $\hat{\mathbf{\Gamma}}_\alpha$ are close to the true ones. However, allowing for additional factors, we observe the "alpha-eating" effect for both the first and second characteristic. At the same time, the coefficients on the third and fourth characteristic increase in magnitude, although the correlation with the truly relevant characteristics are low (i.e., all correlations are below 0.01).

This robustness analysis reveals that even if the true $\mathbf{\Gamma}_\alpha$ coefficients take unrealistic high values, the "alpha-eating" effects results in the acceptance of the null hypothesis if a sufficient number of factors is included, supporting the findings that IPCA will always create a covariance-based explanation for returns, even if the "anomalous returns" are unrealistic high.

IPCA			Alpha-Beta				Alpha
K	1	2	3	4	5	6	0
$\hat{\Gamma}_{\alpha,1}$	0.6780	0.6720	0.6417	0.2642	0.2396	0.0715	0.6767
$\hat{\Gamma}_{\alpha,2}$	-0.5474	-0.5472	-0.4039	-0.3595	-0.3830	-0.2460	-0.5475
$\hat{\Gamma}_{\alpha,3}$	0.0014	0.0003	-0.2663	-0.2255	-0.2298	-0.1597	0.0010
$\hat{\Gamma}_{\alpha,4}$	-0.0038	-0.0109	-0.0390	0.3249	0.3104	0.0954	-0.0033
$\hat{\Gamma}_{\alpha,5}$	-0.0010	-0.0015	0.0574	0.0656	0.0762	-0.1191	-0.0008
$\hat{\Gamma}_{\alpha,6}$	-0.0001	0.0010	-0.0377	-0.0937	-0.0786	-0.2212	0.0001
$\hat{\Gamma}_{\alpha,7}$	-0.0002	0.0001	0.0245	0.0296	0.0313	-0.1359	-0.0003
$\hat{\Gamma}_{\alpha,8}$	-0.0001	-0.0002	0.0147	0.0361	0.0333	0.0163	-0.0001
$\hat{\Gamma}_{\alpha,9}$	0.0314	-0.0010	-0.0035	-0.0079	-0.0084	-0.0036	0.0334
$\hat{\Gamma}_{\alpha,10}$	0.0147	-0.0560	-0.0466	-0.0500	-0.0463	-0.0149	0.0138

Table A.26: $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients for the increased $\boldsymbol{\Gamma}_\alpha$ setting (covariance-characteristic world)
This table reports the estimated $\hat{\boldsymbol{\Gamma}}_\alpha$ coefficients for each estimated IPCA, i.e., beta-only ("Beta"), unrestricted ("Alpha-Beta"), and alpha-only ("Alpha") for our simulation setup with upward scaled $\boldsymbol{\Gamma}_\alpha$ coefficients. The term $\hat{\Gamma}_{\alpha,l}$ denotes the estimated coefficient of the l-th characteristic for the alpha.

Appendix B. Empirical results

This appendix replicates Table 1 in KPS and extends it by including results for the alpha-only IPCA (i.e., $K = 0$).

		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 B.27: 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., $K = 0$).