Multifactor models and the APT: Evidence from a broad cross-section of stock returns

Ilan Cooper[1] Paulo Maio[2] Dennis Philip[3]

This version: June 2016[4]

[1]Norwegian Business School (BI), Department of Finance. E-mail: ilan.cooper@bi.no
[2]Hanken School of Economics, Department of Finance and Statistics. E-mail: paulofmaio@gmail.com
[3]Durham University Business School. E-mail: dennis.philip@durham.ac.uk
[4]We are grateful to Kenneth French, Robert Stambaugh, and Lu Zhang for providing stock market data. Any remaining errors are our own.

Abstract

We seek to describe the broad cross-section of average stock returns. We follow the APT literature and estimate the common factor structure among a large cross-section containing 278 decile portfolios associated with 28 market anomalies. Our statistical model contains seven common factors (with an economic meaning) and prices well both the original portfolio returns and an efficient combination of these portfolios. This model clearly outperforms the empirical workhorses in the literature when it comes to pricing this broad cross-section. Augmenting the empirical models with new factor-mimicking portfolios, based on APT principles, significantly improves their performance. Our results shed light on the number of factors necessary to describe expected stock returns. Moreover, we show that there is significant room for improving the existing multifactor models in terms of explaining the large cross-section of stock returns (and in a way that is consistent with the APT).

Keywords: asset pricing; linear multifactor models; APT; equity risk factors; stock market anomalies; cross-section of stock returns; principal components

JEL classification: G10; G12

1 Introduction

For many years, price momentum (Jegadeesh and Titman (1993), Fama and French (1996)) and the value premium (Basu (1983), Rosenberg, Reid, and Lanstein (1985), and Lakonishok, Shleifer, and Vishny (1994)) have been the traditional market anomalies, and hence the focus of attention for newly proposed asset pricing models.[1] However, recent years have noticed an explosion of new CAPM anomalies, which correspond to new patterns in cross-sectional equity risk premia left unexplained by the baseline CAPM of Sharpe (1964) and Lintner (1965). Specifically, Hou, Xue, and Zhang (2015b) examine in total around 80 anomalies covering different broad categories: momentum, value-growth, investment, profitability, intangibles, and trading frictions. Yet, they find that nearly one-half of these anomalies (including those related to trading frictions) are not statistically significant and end up testing their four-factor model over 35 portfolio sorts. Among the most prominent new patterns in cross-sectional risk premia are a number of investment- and profitability-based anomalies. The investment anomaly can be broadly classified as a pattern in which stocks of firms that invest more exhibit lower average returns than the stocks of firms that invest less.[2] The profitability-based anomalies refer to the evidence indicating that more profitable firms earn higher average returns than less profitable firms.[3] Other pervasive CAPM anomalies include earnings surprise (Foster, Olsen, and Shevlin (1984)), industry momentum (Moskowitz and Grinblatt (1999)), equity duration (Dechow, Sloan, and Soliman (2004)), operating leverage (Novy-Marx (2011)), organizational capital-to-assets (Eisfeldt and Papanikolaou (2013)), and advertisement expense-to-market (Chan, Lakonishok, and Sougiannis (2001)).

The traditional workhorse in the empirical asset pricing literature—the three-factor model of Fama and French (1993, 1996) (FF3 henceforth)—fails to explain the new market anomalies (see, for example, Fama and French (2015), Hou, Xue, and Zhang (2015a, 2015b), and Maio (2016a, 2016b)). Moreover, the four-factor model of Carhart (1997) (C4) does a good

[1]The value premium refers to the evidence showing that value stocks (stocks with high equity valuation ratios like book-to-market, earnings-to-price, or cash flow-to-price) outperform growth stocks (low valuation ratios). On the other, price momentum corresponds to a cross-sectional pattern where stocks with high prior short-term returns outperform stocks with low prior returns.

[2]The variables that represent corporate investment can be investment-to-assets (Cooper, Gulen, and Schill (2008)), abnormal corporate investment (Titman, Wei, and Xie (2004)), investment growth (Xing (2008)), changes in property, plant, equipment, and inventories scaled by assets (Lyandres, Sun, and Zhang (2008)), inventory growth (Belo and Lin (2011)), composite issuance (Daniel and Titman (2006)), net stock issues (Pontiff and Woodgate (2008)), and different measures of accruals (Sloan (1996), Richardson et al. (2005), and Hafzalla, Lundholm, and Van Winkle (2011)).

[3]The profitability measures that have been employed include return on equity (Haugen and Baker (1996)), return on assets (Balakrishnan, Bartov, and Faurel (2010)), gross profits-to-assets (Novy-Marx (2013)), revenue surprise (Jegadeesh and Livnat (2006)), number of consecutive quarters with earnings increases (Barth, Elliott, and Finn (1999)), and failure probability (Campbell, Hilscher, and Szilagyi (2008)).

job in capturing price momentum, but also struggles in terms of explaining some of the profitability- and investment-based anomalies (see Hou, Xue, and Zhang (2015b) and Maio (2016a, 2016b) for details). In response to this gap, we have witnessed the emergence of new multifactor models containing (different versions of) investment and profitability factors, in particular the five-factor model of Fama and French (2015, 2016) (FF5) and the four-factor model of Hou, Xue, and Zhang (2015a, 2015b) (HXZ4). However, several dimensions of the broad cross-section of stock returns are still not explained by the new factor models. In particular, the five-factor model does not account for momentum (including industry momentum), while both of these models do not capture several profitability and investment-based (in particular, several forms of accruals) anomalies (see Hou, Xue, and Zhang (2015a, 2015b), Fama and French (2016), and Maio (2016a, 2016b) for details on the performance of those models for the broad cross-section).

Following such evidence, several questions naturally emerge in the empirical asset pricing literature: How many factors do we need, and what are these factors, to describe well the broad cross-section of stock returns?[4] To which dimensions of the cross-section of stock returns are these factors more correlated? To what extent (and how) can we improve the current multifactor models proposed in the literature in order to achieve a better description of large-scale cross-sectional risk premia? This paper attempts at providing answers to these questions. In order to achieve this goal, we adopt the general framework of the Arbitrage Pricing Theory (APT) of Ross (1976). According to the APT, variables that provide a fairly good description of the time-series variation in average stock returns should represent risk factors that help to price those same assets. Thus, the APT is an obvious asset pricing framework to study the large cross-section of stock returns since several of the most successful multifactor models in the literature, as those mentioned above, contain factors that are (nearly) mechanically correlated with the testing portfolios.[5] Moreover, the APT is less demanding than other asset pricing frameworks (like the ICAPM of Merton (1973)) in the sense that it relies on relative asset pricing, specifically, given the common sources of systematic risk (factors) what should be the correct discount rates for equity portfolios.

We follow part of the relatively small empirical APT literature in terms of estimating common stock return factors by applying asymptotical principal components analysis (APCA) to a large cross-section of stock returns (e.g., Connor and Korajczyk (1986, 1988) and Goyal,

[4]Cochrane (2011) asks similar questions: "Can we again account for N dimensions of expected returns with $K < N$ factor exposures?".

[5]This is the case of the value-growth factor (HML) in relation to portfolios sorted on valuation ratios, the momentum factor (UMD) against momentum portfolios, and the investment and profitability factors used in Fama and French (2015, 2016) and Hou, Xue, and Zhang (2015a, 2015b) in relation to portfolios sorted on these two variables.

Pérignon, and Villa (2008)). We employ a total of 28 anomalies or portfolio sorts, which represent a subset of the anomalies considered in Hou, Xue, and Zhang (2015a, 2015b) for a total of 278 decile portfolios. The estimation results show that there are seven common factors that are statistically significant over our sample period (1972 to 2013). These seven factors cumulatively explain around 91% of the cross-sectional variations in the 278 portfolio returns. The first common factor basically captures the average anomaly and thus resembles a market factor. The other six factors capture different dimensions of the large cross-section of market anomalies. In particular, the second, third, and four factors are strongly correlated with value-growth, investment, profitability, and momentum-based anomalies. This is consistent with the role of the seven-factor model in terms of describing well this cross-section of 278 equity portfolios. This statistical model is thus a benchmark for this specific cross-section of stock returns, against which the existent models are compared.

We conduct cross-sectional asset pricing tests of our APT model by using the 278 equity portfolios as testing assets. The results confirm that the seven-factor model explains about 60% of the cross-sectional variation in the risk premia associated with the 278 portfolios. Moreover, most factor risk price estimates are statistically significant. Across categories of anomalies, the APT does a better job in pricing value-growth and intangibles, compared to the group of investment-based anomalies. Moreover, the model prices perfectly an efficient combination of the original portfolios as indicated by the GLS cross-sectional R^2 estimates around 100%. This result confirms that the statistical model is a successful APT.

Next, we compare our APT model to some of most popular multifactor models existent in the literature in terms of pricing the 278 portfolios. The models include the already mentioned FF3, C4, HXZ4, FF5, in addition to a restricted version of FF5 that excludes HML (FF4), and the four-factor model of Pástor and Stambaugh (2003) (which includes a stock liquidity factor). The results show that only C4 and HXZ4 offer an economically significant explanatory power for the broad cross-section of stock returns, while the fit of both FF5 and FF4 is quite small. Moreover, the performance of all the six empirical factor models clearly lags behind the fit of the seven-factor APT, suggesting that these models have a large room for improvement in terms of describing large-scale cross-sectional risk premia.

In light of such evidence, we define and estimate new empirical multifactor models to better describe the broad cross-section of anomalies. All these models contain seven factors, to be consistent with our benchmark APT, and represent augmented versions of C4, HXZ4, FF5, and FF4, the best performing empirical models. The new factors in each of these models represent factor-mimicking portfolios (spreads among extreme portfolio deciles) associated with selected anomalies. These anomalies are those for which the original factors in each model do a worse job in terms of describing the time-series variation in the corresponding decile portfolio returns. Thus, our criteria for selecting the new factors relies on the APT restriction that the risk factors should explain well the time-series variation in the returns of the testing assets. The results show that adding the new factors improves all four empirical models, and helps especially the performance of both FF5 and FF4 in terms of explaining the large cross-section of stock returns. Moreover, the augmented models do a very good job in explaining an efficient combination of the original portfolios, thus, showing that they represent valid APTs. Therefore, the performance of the augmented empirical models is quite similar to that of our benchmark APT. Overall, our results indicate that there is a significant room for improving the existing empirical multifactor models in terms of explaining the large cross-section of stock returns in a way that is consistent with the APT.

The paper proceeds as follows. Section 2 discusses the derivation and empirical implications of the APT, while in Section 3, we estimate our benchmark statistical model. Section 4 presents the asset pricing tests for our benchmark APT, whereas in Section 5 we conduct a comparison with existing multifactor models. Section 6 concludes.

2 Theoretical background

In this section, we provide a simple derivation of the Arbitrage Pricing Theory (APT) model of Ross (1976), which follows closely the exposition in Cochrane (2005), Chapter 9.

Consider the following time-series regression for an arbitrary risky asset $i = 1, ..., N$,

$$R_{i,t+1} = \alpha_i + \beta_{i,1}\widetilde{f}_{1,t+1} + ... + \beta_{i,K}\widetilde{f}_{K,t+1} + \varepsilon_{i,t+1}, \tag{1}$$

where $R_{i,t+1}$ denotes the gross return on asset i, $\varepsilon_{i,t+1}$ is the idiosyncratic return, and $\widetilde{f}_{j,t+1} \equiv f_{j,t+1} - \mathrm{E}(f_{j,t+1}), j = 1, ..., K$ represents each of the demeaned common K factors. Since the factors are demeaned, it follows that $\mathrm{E}(R_{i,t+1}) = \alpha_i$.

Assume that there is a stochastic discount factor (SDF), M_{t+1}, that prices assets in this economy. By multiplying both sides of the regression above by M_{t+1}, taking unconditional expectations, and using both $\mathrm{E}(M_{t+1}R_{i,t+1}) = 1$ and $\mathrm{E}(R_{i,t+1}) = \alpha_i$, we obtain:

$$\mathrm{E}(R_{i,t+1}) - \frac{1}{\mathrm{E}(M_{t+1})} = -\beta_{i,1}\frac{\mathrm{E}(M_{t+1}\widetilde{f}_{1,t+1})}{\mathrm{E}(M_{t+1})} - ... - \beta_{i,K}\frac{\mathrm{E}(M_{t+1}\widetilde{f}_{K,t+1})}{\mathrm{E}(M_{t+1})} - \frac{\mathrm{E}(M_{t+1}\varepsilon_{i,t+1})}{\mathrm{E}(M_{t+1})}. \tag{2}$$

In the derivation of the equation above we are simply using the law of one price by forcing both sides of the regression to have the same price. In other words, the returns associated with the risky asset i and the replicating portfolio have the same price.

Now assume that idiosyncratic risk is small, $\mathrm{Var}(\varepsilon_{i,t+1}) \simeq 0$. This implies that $\mathrm{E}(M_{t+1}\varepsilon_{i,t+1}) =$

$\mathrm{Cov}(M_{t+1}, \varepsilon_{i,t+1}) \simeq 0$, since in the limit a very small value of $\mathrm{Var}(\varepsilon_{i,t+1})$ means that $\varepsilon_{i,t+1}$ is not a random variable. This in turn implies:

$$\mathrm{E}(R_{i,t+1}) - \frac{1}{\mathrm{E}(M_{t+1})} \simeq -\beta_{i,1}\frac{\mathrm{E}(M_{t+1}\widetilde{f}_{1,t+1})}{\mathrm{E}(M_{t+1})} - ... - \beta_{i,K}\frac{\mathrm{E}(M_{t+1}\widetilde{f}_{K,t+1})}{\mathrm{E}(M_{t+1})}. \tag{3}$$

Assume also that there is a risk-free asset with gross return given by $R_{f,t+1}$, where $\mathrm{E}(R_{f,t+1}) = 1/\mathrm{E}(M_{t+1})$. This leads to the following expected return-beta equation,

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) \simeq \beta_{i,1}\lambda_1 + ... + \beta_{i,K}\lambda_K, \tag{4}$$

where

$$\lambda_j = -\mathrm{E}(R_{f,t+1})\,\mathrm{E}(M_{t+1}\widetilde{f}_{j,t+1}), j = 1, ..., K, \tag{5}$$

represents the risk price for factor j. Hence, under the statistical model above, by assuming low idiosyncratic risk and using non-arbitrage (or more specifically, the law of one price), we obtain a linear asset pricing model where the common factors represent risk factors that price risky assets.

If we assume that all the factors are excess returns, $\mathrm{E}(M_{t+1}f_{j,t+1}) = 0$, the risk price for factor j simplifies to

$$\begin{aligned} \lambda_j &= -\mathrm{E}(R_{f,t+1})\,\mathrm{E}[M_{t+1}(f_{j,t+1} - \mathrm{E}(f_{j,t+1}))] \\ &= -\mathrm{E}(R_{f,t+1})[\mathrm{E}(M_{t+1}f_{j,t+1}) - \mathrm{E}(M_{t+1})\,\mathrm{E}(f_{j,t+1})] \\ &= \mathrm{E}(R_{f,t+1})\,\mathrm{E}(M_{t+1})\,\mathrm{E}(f_{j,t+1}) \\ &= \mathrm{E}(f_{j,t+1}) \end{aligned} \tag{6}$$

that is, the risk price for each factor equals the corresponding factor mean.

Some observations about the empirical implications of the APT are in order. First, as shown above, one of the major assumptions of the APT is that each risky assets follows a factor structure (linear regression), and that the amount of idiosyncratic risk is small.[6] This implies that the coefficient of determination (R^2) of the regression above should be large:[7]

$$R^2 = 1 - \frac{\mathrm{Var}(\varepsilon_{i,t+1})}{\mathrm{Var}(R_{i,t+1})}. \tag{7}$$

[6]This leads some people to classify the APT as a statistical model. Yet, as shown in this simple derivation, there is economic content in the form of absence of arbitrage opportunities.

[7]Several APT frameworks assume that idiosyncratic returns are orthogonal across assets, $\mathrm{Cov}(\varepsilon_{i,t+1}, \varepsilon_{l,t+1}) = 0, i \neq l$. Yet, as shown in this section, this assumption is not necessary to obtain the approximate beta model (see also Chamberlain and Rothschild (1983)).

Hence, a necessary condition for a given model to be (approximately) a valid APT model is that the time-series regressions of asset returns on the common factors produce large R^2 estimates. Although the APT framework applies to any risky asset, the restriction of a high R^2 limits in practice its application to equity portfolios as testing assets. The reason is that individual stocks typically have large idiosyncratic risk, in contrast to equity portfolios, and thus quite low R^2 estimates.

Second, the exposition above shows that the factors in the APT can be either traded or non-traded. Yet, given the restriction of a large fit in the time-series regressions it follows that plausible empirical applications of the APT should contain factors that represent excess stock returns (zero-cost portfolios). The reason is that those factors typically are more correlated with the testing assets (equity portfolios) than non-traded factors like macro variables (e.g., CPI inflation, industrial production growth, bond yields, short-term interest rates). Actually, in some cases these large correlations are (nearly) mechanical (like the case of HML against portfolios sorted on the book-to-market ratio or the case of UMD in relation to momentum portfolios). Furthermore, if the factors are excess returns, the risk price estimates cannot be freely estimated by a cross-sectional regression and should equal the corresponding factor means.

Third, the APT is mainly about relative asset pricing: given the factors, what should be the correct prices (i.e., expected returns) of the other assets in the economy. Yet, the APT does not provide an economic explanation for the risk premium associated with each original source of systematic risk (the factors).[8]

3 Common factors

In this section, we estimate the common factors that summarize the information from the broad cross-section of stock returns.

3.1 Data

The portfolio return data used in the estimation of the common factors are associated with the most relevant market or CAPM anomalies, which represent patterns in cross-sectional stock returns that are not explained by the baseline CAPM. We employ a total of 28 anomalies or portfolio sorts, which represent a subset of the anomalies considered in Hou, Xue, and

[8] Alternative asset pricing frameworks, which provide a theory of the factor risk premiums, include the Consumption CAPM (Breeden (1979)), the Intertemporal CAPM (Merton (1973)), and the baseline CAPM (Sharpe (1964) and Lintner (1965)).

Zhang (2015a, 2015b) for a total of 278 portfolios. Table 1 contains the list and description of the anomalies included in this paper. Following Hou, Xue, and Zhang (2015b), these anomalies can be generically classified in strategies related to value-growth (BM, DUR, CFP, EP, and REV), momentum (MOM, SUE, ABR, IM, and ABR*), investment (IA, ACI, NSI, CEI, PIA, IG, IVC, IVG, NOA, OA, POA, and PTA), profitability (ROE, GPA, NEI, and RS), and intangibles (OCA and OL). The portfolio returns are value-weighted and all the groups include decile portfolios, except IM and NEI with nine portfolios each. In comparison to the portfolio groups employed in Hou, Xue, and Zhang (2015b), we do not consider the return on assets (Balakrishnan, Bartov, and Faurel (2010)) deciles because they are strongly correlated with the return on equity deciles (ROE). Moreover, we use only one measure of price momentum (MOM) and earnings surprise (SUE) since the other measures used in Hou, Xue, and Zhang (2015b) are strongly correlated with either MOM or SUE.[9] We also exclude all portfolio sorts used in Hou, Xue, and Zhang (2015b) with data that starts after 1972:01.[10] All the portfolio return data correspond to Hou, Xue, and Zhang (2015b) and were obtained from Lu Zhang. To construct portfolio excess returns, we subtract the one-month Treasury bill rate available from Kenneth French's website. The sample period is 1972:01 to 2013:12.

Table 2 presents the descriptive statistics for high-minus-low spreads in returns between the last and first decile among each portfolio class. The anomaly with the largest spread in average returns is price momentum (MOM) with a premium above 1% per month. The return spreads associated with book-to-market (BM), ABR (abnormal one-month returns after earnings announcements), ROE, and net stock issues (NSI) are also quite pervasive with (absolute) means around 0.70% per month. The anomalies with lower average returns are ABR* (abnormal six-month returns after earnings announcements), revenue surprises (RS), abnormal corporate investment (ACI), and operating leverage (OL), all with average return spreads around or below 0.30% (in absolute value). Price momentum is by far the anomaly with the most volatile spread in returns (standard deviation above 7% per month), followed by ROE, return reversal (REV), and industry momentum (IM), all three spreads with volatilities above 5%. On the other extreme, investment growth (IG) and numbers of consecutive quarters with earnings increases (NEI) show the least volatile return spreads (below 3% per month).

[9] We exclude anomalies for which the corresponding spreads high-minus-low in returns have correlations above 90% (in magnitude) relative to other anomalies.

[10] This includes, for example, portfolios sorted on revision in analysts' earnings forecasts (Chan, Jegadeesh, and Lakonishok (1996)), advertisement expense-to-market (Chan, Lakonishok, and Sougiannis (2001)), R&D-to-market (Chan, Lakonishok, and Sougiannis (2001)), failure probability (Campbell, Hilscher, and Szilagyi (2008)), and systematic volatility (Ang et al. (2006)).

3.2 Factors estimation

To estimate the common stock return factors, we use the approximate factor model framework developed by Connor and Korajczyk (1986, 1988) and widely implemented to capture the pervasive cross-correlations present in large macroeconomic or financial panels (see Stock and Watson (2002a, 2002b), Ludvigson and Ng (2007, 2009, 2010), Goyal, Pérignon, and Villa (2008), Maio and Philip (2015), among others).

Consider that equity portfolio returns are driven by a finite number of r static unobservable factors,

$$R_{it} = \mathbf{f}_t' \boldsymbol{\theta}_i + \varepsilon_{it}, \tag{8}$$

where R_{it} is the portfolio $(i = 1, ..., N)$ return at time $t(= 1, ..., T)$; $\mathbf{F}_t$ is the r-dimensional vector of latent common factors for all returns at t; $\boldsymbol{\theta}_i$ is the r-dimensional vector of factor loadings for the return on asset i; and ε_{it} stands for the idiosyncratic *i.i.d.* errors, which are allowed to have limited correlation among returns.

This model captures the main sources of variations and covariations among the N portfolio returns with a set of r common factors $(r << N)$. The framework is estimated using asymptotic principal components procedure, which involves an eigen decomposition of the sample covariance matrix. The estimated $(T \times r)$ factors matrix $\hat{\mathbf{F}}$ is equal to $\sqrt{T}$ multiplied by the r eigenvectors corresponding to the first r largest eigenvalues of the $T \times T$ matrix, $\mathbf{RR}'/(NT)$, where $\mathbf{R}$ is a $(T \times N)$ return data matrix. The normalization $\hat{\mathbf{F}}'\hat{\mathbf{F}} = \mathbf{I}_r$ is imposed, where $\mathbf{I}_r$ is the $r-$dimensional identity matrix, since $\mathbf{F}$ and the factor loadings matrix are not separately identifiable. The factor loadings matrix can be obtained as $\hat{\boldsymbol{\Theta}} = \mathbf{R}'\hat{\mathbf{F}}/T$. For a large number of return time series this methodology can effectively distinguish noise from signal and summarize information into a small number of estimated common factors.

To determine the value of r, which is the number of statistically significant common factors, we use the IC_2 information criterion suggested by Bai and Ng (2002). We minimize over r the following criterion,

$$\ln(V_r) + r\left(\frac{N+T}{NT}\right)\ln(\min\{N, T\}), \tag{9}$$

where $V_r = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(R_{it} - \sum_{j=1}^{r} \hat{F}_{jt}\hat{\theta}_{ji} \right)^2$, with $\hat{F}_{jt}$ denoting the j^{th} factor estimate at time t. We consider a maximum set of 20 factors when estimating the optimal r.

The estimation results show that there are seven common return factors that are statistically significant over our sample period.[11] Table 3 reports summary statistics of the seven

[11] For robustness, we also implement the tuning-stability checkup procedure proposed by Alessi, Barigozzi,

estimated factors. We can see that none of the factors is persistent as shown by the first-order autocorrelation coefficients around or below 0.10. This stems from the fact that stock returns do not usually exhibit significant serial correlation. The seven factors cumulatively explain around 91% of the total variations in the 278 portfolio returns, with the first factor explaining the largest proportion of the cross-sectional variation in returns (around 85%).[12]

As stressed in the last section, a necessary condition for a factor model to represent a valid APT is that the time-series regressions of the testing returns on the risk factors have a large fit in the form of large R^2 estimates. To confirm this proposition, we run multiple time-series regressions for each of the 278 excess portfolio return on the seven common factors,

$$R_{i,t+1}-R_{f,t+1}=\delta_i+\beta_{i,1}F_{1,t+1}+\beta_{i,2}F_{2,t+1}+\beta_{i,3}F_{3,t+1}+\beta_{i,4}F_{4,t+1}+\beta_{i,5}F_{5,t+1}+\beta_{i,6}F_{6,t+1}+\beta_{i,7}F_{7,t+1}+\varepsilon_{i,t+1}. \tag{10}$$

In Table 3, we report the average R^2 estimates (across the 278 regressions) for a regression on the seven factors and also for regressions on subsets of factors. We can see that the average fit of the regression including all seven factors is large (90%). Moreover, the first four factors contribute the most for this explanatory power (89%). As a reference point, untabulated results indicate that the CAPM produces an average R^2 of 84% across the 278 portfolios. This is similar to the average explanatory ratio associated with the first APCA factor.

To understand the correlations of the estimated common factors with the raw portfolio returns, we conduct single regressions of the 28 return spreads indicated above on each of the factors,

$$R_{l,10,t+1}-R_{l,1,t+1}=\delta_{l,j}+\beta_{l,j}F_{j,t+1}+\varepsilon_{l,t+1}, j=1,...,7, \tag{11}$$

where $R_{l,10,t+1}-R_{l,1,t+1}$ denotes the spread high-minus-low associated with anomaly $l, l=1,...,28$.

Table 4 presents the R^2 estimates associated with these single regressions. These estimates represent the square of the simple pairwise correlations between the returns and factors. We can see that the first factor has small correlations across most of the 28 anomalies. This is consistent with the evidence provided below that this factor is mainly correlated with the usual stock market factor. The main outlier occurs for the CEI spread with a R^2

and Capasso (2010) on the IC_2 criteria, as it is shown to outperform the original estimators in finite samples. The tests results indicate the presence of eight significant common return factors. Yet, since the last estimated factors explain a relatively small variation of the cross-section of stock returns, as shown below, we opted to work with seven factors.

[12]In related work, Clarke (2016) extracts three common factors from the cross-sectional of stock returns. However, the procedure in that paper differs significantly from the one used here. First, the principal components are estimated from 25 portfolios sorted on expected returns rather than the original anomaly variables. Second, he uses seven anomaly variables in the estimation of individual expected stock returns, which represents a significantly smaller cross-section than the 28 anomalies used in our study.

of 24%. On the other hand, F_2 is strongly correlated with the value-growth anomalies as indicated by the R^2 estimates above 40% for the spreads associated with BM, CFP, EP, and DUR. This factor is also significantly correlated with several investment (IA, CEI, and NSI) and accruals (POA and PTA) anomalies. This indicates that there is some degree of comovement among the value- and investment-based anomalies.

The third factor is heavily correlated with the profitability-based anomalies as shown by the R^2 estimates between 21% (GPA) and 62% (ROE). Moreover, this factor is also correlated with the BM and REV return spreads (R^2 close to 30%). The fourth factor mainly captures price momentum as indicated by the explanatory ratios around 60% for both the MOM and IM spreads. F_5 is mainly correlated with OL (R^2 of 35%) and OA (23%), while the sixth factor loads more significantly on the PIA spread (30%). The R^2 estimates associated with the seventh factor are significantly smaller in comparison with the other factors, with the largest values occurring for the NOA and OL return spreads (8%). Hence, the last three factors mainly capture operating leverage, some forms of accruals, and PIA. In sum, the results from Table 4 suggest that to a large degree the seven common factors capture different dimensions of the large cross-section of market anomalies. This is consistent with the role of the seven-factor model in terms of describing well this cross-section of 278 equity portfolios.

4 Asset pricing tests

In this section, we estimate our APT model for the broad cross-section of stock returns.

4.1 Methodology

To test our APT for the broad cross-section of stock returns, we use the two-step time-series/cross-sectional regression procedure employed in Black, Jensen, and Scholes (1972), Jagannathan and Wang (1998), Cochrane (2005), Brennan, Wang, and Xia (2004), Campbell and Vuolteenaho (2004), and Maio and Santa-Clara (2016), among others.[13] In the first step, the factor betas are estimated from the time-series multivariate regressions for each of the testing portfolios,

$$\mathbf{r}_{t+1} = \boldsymbol{\beta}\mathbf{f}_{t+1} + \boldsymbol{\varepsilon}_{t+1}, \tag{12}$$

[13]Since the common factors estimated by APCA do not represent excess stock returns we can not use time-series regressions in order to test the asset pricing restrictions (see, for example, Fama and French (1993, 1996) and Cochrane (2005), Chapter 12).

where $\mathbf{r}_{t+1}$ is a vector of excess portfolio returns, $\boldsymbol{\beta}(N \times K)$ is a matrix of K factor loadings for the N test assets, $\mathbf{f}_{t+1}(K \times 1)$ is a vector of factor realizations, and $\boldsymbol{\varepsilon}_{t+1}(N \times 1)$ is the vector of return disturbances.

In the second step, the K-factor APT is estimated by an OLS cross-sectional regression,

$$\overline{\mathbf{r}} = \boldsymbol{\beta}\boldsymbol{\lambda} + \boldsymbol{\alpha}, \tag{13}$$

where $\overline{\mathbf{r}}(N \times 1)$ is a vector of average excess returns, $\boldsymbol{\lambda}(K \times 1)$ is a vector of risk prices, and $\boldsymbol{\alpha}(N \times 1)$ is the vector of pricing errors.

The t-statistics associated with the factor risk price estimates are based on Shanken's standard errors (Shanken (1992)), which incorporate a correction for the estimation error in the factor loadings. We do not include an intercept in the cross-sectional regression, since we want to impose the economic restrictions associated with each factor model. If the APT is correctly specified, the intercept in the cross-sectional regression should be equal to zero.[14]

To gauge the fit of each model, we compute the cross-sectional OLS coefficient of determination,

$$R^2_{OLS} = 1 - \frac{\mathrm{Var}_N(\hat{\alpha}_i)}{\mathrm{Var}_N(\overline{R_i - R_f})}, \tag{14}$$

where $\mathrm{Var}_N(\cdot)$ stands for the cross-sectional variance. R^2_{OLS} represents the fraction of the cross-sectional variance of average excess returns on the testing assets explained by the factor loadings associated with a given model. Since we do not include an intercept in the cross-sectional regression, this R^2 measure can assume negative values. A negative estimate means that the regression including the factor loadings (associated with a given model) as regressors does worse than a simple regression containing just an intercept (see Campbell and Vuolteenaho (2004)).[15]

An alternative empirical method to test the factor models consists of estimating a GLS cross-sectional regression in the second step,

$$\boldsymbol{\Sigma}^{-\frac{1}{2}}\overline{\mathbf{r}} = \left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\beta}\right)\boldsymbol{\lambda} + \boldsymbol{\alpha}, \tag{15}$$

[14] Another reason for not including the intercept in the cross-sectional regressions is that often the factor loadings associated with equity portfolios are very close to each other, creating a multicollinearity problem in the cross-sectional regression (see, for example, Jagannathan and Wang (2007)). Moreover, excluding the intercept from the cross-sectional regression enables consistency with the time-series regression approach, which applies to models where all factors are excess returns. Specifically, the time-series intercepts (alphas) correspond to the pricing errors from an implied cross-sectional regression without intercept and where the risk price estimates are equal to the factor means (see Cochrane (2005), Kan, Robotti, and Shanken (2013), and Maio (2016a)).

[15] Similar R^2 measures are used in Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Campbell and Vuolteenaho (2004), Yogo (2006), Maio and Santa-Clara (2012, 2016), Maio (2013), Lioui and Maio (2014), among others.

where $\mathbf{\Sigma} \equiv \mathrm{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t')$ denotes the variance-covariance matrix associated with the residuals from the time-series regressions (see Cochrane (2005), Shanken and Zhou (2007), Lewellen, Nagel, and Shanken (2010), among others). Under this approach, the testing assets with a lower variance of the time-series residuals receive more weight in the cross-sectional regression. Thus, unlike the OLS regression, the GLS regression attempts to price as best as possible an efficient (minimum variance) combination of the testing portfolios rather than the original returns. This means that we are no longer pricing the original portfolios, which frequently have an economic interest attached as is the case in this paper (see Cochrane (2005), Chapter 12 and Ludvigson (2013) for a detailed discussion).

To assess the fit of each model for the repackaged portfolios, we compute the cross-sectional GLS coefficient of determination,

$$R_{GLS}^2 = 1 - \frac{\hat{\boldsymbol{\alpha}}' \mathbf{\Sigma}^{-1} \hat{\boldsymbol{\alpha}}}{\overline{\mathbf{r}^*}' \mathbf{\Sigma}^{-1} \overline{\mathbf{r}^*}}, \tag{16}$$

where $\overline{\mathbf{r}^*}$ denotes the $N \times 1$ vector of (cross-sectionally) demeaned average excess returns. This metric measures the fraction of the cross-sectional variation in risk premia among the "transformed" portfolios explained by the factors associated with a given model. Thus, a high value of R_{GLS}^2 means that a combination of the K factors is close to the mean-variance frontier constructed from the testing portfolios (see Kandel and Stambaugh (1995), Cochrane (2005), and Lewellen, Nagel, and Shanken (2010)).

4.2 Results

The results for the asset pricing tests of the seven-factor model, as well as the nested models, based on the OLS cross-sectional regression approach are displayed in Table 5. The testing assets are the 278 equity portfolios. The seven-factor APT model explains about 60% of the cross-sectional variation in the equity risk premia among the 278 portfolios. This represents a large fit given the large dimension of the cross-section and the high number of anomalies considered. Regarding the nested models, we can see that the fit increases almost monotonically as we add one factor consecutively: the R^2 estimates increase from -46% for the model including only the first APCA factor (F_1) to 55% for the model containing the first six common factors. Since the first factor represents basically a market factor (as discussed below), these results are consistent with previous evidence showing that the baseline CAPM has a negative fit when it comes to explain these market anomalies.[16] In fact, untabulated results

[16]When tested on portfolios like those used in this paper, the CAPM typically produces negative R_{OLS}^2 estimates (see Campbell and Vuolteenaho (2004), Yogo (2006), Maio and Santa-Clara (2012) Maio (2013, 2016), among others). This means that the model performs worse than a trivial model that predicts constant

show that the CAPM produces an explanatory ratio of -49%, which represents basically the same fit as that produced by the one-factor APT.

Turning to the risk price estimates, it turns out that the estimates for λ_1, λ_2, λ_4, and λ_6 are consistently negative, while the estimates for λ_3, λ_5, and λ_7 are always positive. Most of the risk price estimates are strongly significant (1% level). The exceptions are λ_3, where there is significance at the 5% or 10% level, and λ_5 in which case the estimates are not significant at the 10% level. This suggests that F_5 is less relevant than the other factors in terms of explaining the broad cross-section of expected stock returns.

In Table 6, we present the OLS risk price estimates when the seven-factor APT is tested on each of five groups of anomalies considered in Section 3: value-growth, momentum, profitability, investment, and intangibles. The objective is to assess the relative explanatory power of the APT for different types of anomalies. The results show that the fit of the APT is large across the five groups of anomalies: the explanatory ratios vary between 62% (investment anomalies) and 89% (value-growth). Thus, the model does relatively worse in explaining the investment-based anomalies, however, this group has significantly more anomalies than the other groups (see Table 1) and thus represents a bigger hurdle for the APT. The risk price estimates for the seven factors have generally the same sign as in the joint estimation with the 278 portfolios. The few exceptions are λ_5 in the estimation with the investment anomalies and λ_7 in the estimation with intangibles, which change sign relative to the benchmark case. Yet, in both cases there is no statistical significance for those estimates. Moreover, the statistical significance of the risk price estimates tends to be weaker than in the full-sample regression as a result of the lower statistical power in the sub-sample tests (fewer cross-sectional observations).

Next, we focus on the asset pricing tests based on the GLS cross-sectional regression approach. Table 7 presents the results for the full cross-section of 28 market anomalies. We can see that the seven-factor model, and the nested specifications, have a large fit for the transformed portfolios. Indeed, apart from the one-factor APT, which has a negative explanatory ratio, all the remaining six specifications produce R^2_{GLS} estimates around 100%. This indicates that different combinations of the common estimated factors are approximately mean-variance efficient.

The risk price estimates have the same signs as the OLS counterparts, and those estimates are in most cases significant at the 1% level. The exceptions are λ_3 (significant at the 10% level) and λ_5 (highly insignificant), which goes in line with the results from the OLS cross-sectional regressions. The large fit of the APT for the transformed portfolios is not totally surprising. As discussed above, the GLS cross-sectional regression attempts to price as best

average returns in the cross-section of equity portfolios.

as possible testing portfolios that have low idiosyncratic risk. Under the APT framework, the common risk factors determine an approximate factor structure for (at least some of the) testing returns, that is, the amount of idiosyncratic risk is quite small for those portfolios. This implies that the APT model will do a very good job in pricing testing assets with low idiosyncratic risk, and the GLS regression assigns a larger weight for those portfolios, which receive (nearly) zero pricing errors (thus producing very large R^2_{GLS} estimates). Hence, another sign of a successful APT is to achieve a large R^2_{GLS} value.

The GLS risk price estimates associated with the seven-factor model across the groups of anomalies are presented in Table 8. The GLS explanatory ratios assume high values, varying between 66% (intangibles) and 91% (value-growth). Thus, the model does a better job in explaining an efficient transformation of the value-related anomalies in comparison to other portfolios. The fact that these estimates are below one stems from the fact that the factors were constructed from the full set of 28 anomalies rather than estimating a different common set of factors for each group of anomalies. The R^2_{GLS} estimates tend to be slightly higher than the OLS explanatory ratios across most categories. The exceptions are for the profitability and intangibles groups, yet the difference in fit is not large.

Overall, the results of this section indicate that the seven-factor model, estimated by APCA, does a good job in describing the cross-section of 278 equity portfolios considered in this study.

5 Relation to multifactor models

In this section, we compare our APT model to some of most popular multifactor models existent in the literature. Following Section 2, we restrict the analysis to models where all the factors represent excess stock returns (zero-cost portfolios). Since the theoretical background of these models is not totally clear, we designate these factor models by "empirical" models.[17]

[17]Specifically, the models proposed by Fama and French (2015, 2016) and Hou, Xue, and Zhang (2015a, 2015b) both contain profitability and investment risk factors. However, while Fama and French (2015) motivate their five-factor model based on the present-value valuation model of Miller and Modigliani (1961), it turns out that Hou, Xue, and Zhang (2015b) rely on the q-theory of investment. On the other hand, Maio and Santa-Clara (2012) and Cooper and Maio (2016) provide evidence that several of the factors included in the models analyzed in this paper are consistent with the Merton's ICAPM (Merton (1973)).

5.1 Factor models

We employ five multifactor models widely used in the cross-sectional asset pricing literature. The first model is the Fama and French (1993, 1996) three-factor model (FF3 henceforth),

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{SMB}\beta_{i,SMB} + \lambda_{HML}\beta_{i,HML}, \tag{17}$$

where λ_M, λ_{SMB}, and λ_{HML} denote the risk prices corresponding to the market, size (SMB), and value (HML) factors, respectively.

The second model is the four-factor model of Carhart (1997) (C4), which incorporates a momentum factor (UMD, up-minus-down short-term past returns) to the FF3 model:

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{SMB}\beta_{i,SMB} + \lambda_{HML}\beta_{i,HML} + \lambda_{UMD}\beta_{i,UMD}. \tag{18}$$

Next, we estimate the four-factor model of Pástor and Stambaugh (2003) (PS4), which replaces UMD by a stock liquidity factor (LIQ):

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{SMB}\beta_{i,SMB} + \lambda_{HML}\beta_{i,HML} + \lambda_{LIQ}\beta_{i,LIQ}. \tag{19}$$

The fourth model is the four-factor model proposed by Hou, Xue, and Zhang (2015a, 2015b) (HXZ4). This model adds an investment factor (IA, low-minus-high investment-to-assets ratio) and a profitability factor (ROE, high-minus-low return on equity) to the usual market and size (ME) factors:[18]

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{ME}\beta_{i,ME} + \lambda_{IA}\beta_{i,IA} + \lambda_{ROE}\beta_{i,ROE}. \tag{20}$$

The fifth model is the five-factor model of Fama and French (2015, 2016) (FF5), which adds an investment (CMA, low-minus-high asset growth) and a profitability (RMW, high-minus-low operating profitability) factor to the FF3 model:

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{SMB^*}\beta_{i,SMB^*} + \lambda_{HML^*}\beta_{i,HML^*} + \lambda_{CMA}\beta_{i,CMA} + \lambda_{RMW}\beta_{i,RMW}. \tag{21}$$

Both CMA and RMW are constructed in a different way than the corresponding investment and profitability factors from Hou, Xue, and Zhang (2015b).[19] In addition, both SMB^* and HML^* are constructed from different portfolio sorts than the original size and value factors

[18] The size factor employed in Hou, Xue, and Zhang (2015b) is constructed in a slightly different way than SMB.

[19] See Fama and French (2015) and Hou, Xue, and Zhang (2015b) for details.

of Fama and French (1993).[20]

Finally, we estimate a restricted version of FF5 that excludes the HML^* factor (FF4):

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{SMB^*}\beta_{i,SMB^*} + \lambda_{CMA}\beta_{i,CMA} + \lambda_{RMW}\beta_{i,RMW}. \tag{22}$$

This follows previous evidence showing that the value factor is redundant within the FF5 model (see Fama and French (2015, 2016) and Hou, Xue, and Zhang (2015a)).

5.2 Data

The data on the equity factors, RM, SMB/SMB^*, HML/HML^*, UMD, CMA, and RMW are obtained from Kenneth French's data library, while LIQ is retrieved from Robert Stambaugh's webpage. The data associated with ME, IA, and ROE are obtained from Lu Zhang. The descriptive statistics for the equity factors are presented in Table 9. The factor with the largest mean is UMD (0.71% per month), followed by ROE and RM, both with means above 0.50% per month. The factors with the lowest average are SMB and SMB^* (around 0.20% per month), which confirms previous evidence that the size premium has declined in the last decades. The most volatile factors are the equity premium and the momentum factor, with standard deviations above 4.5% per month. On the other hand, the investment-based factors (IA and CMA) are the least volatile, with standard deviations slightly above 1.8% per month. All the factors have very low serial correlation, as shown by the small first-order autoregressive coefficients (below 20% in all cases).

Table 10 displays the pairwise correlations among the different equity factors. The different versions of the size (SMB, SMB^*, and ME) and value (HML and HML^*) factors are strongly correlated as indicated by the correlations quite close to one. We can also observe a similar pattern for the asset growth factors (IA and CMA), as shown by the correlation of 0.90. On the other hand, the two profitability factors (ROE and RMW) show a smaller correlation (0.67). Both investment factors are positively correlated with the value factors (estimates around 0.70). On the other hand, ROE is positively correlated with UMD (0.50), but the same does not occur with RMW. Hence, these results suggest that the two profitability factors measure different types of risks.

5.3 Correlation with APCA factors

Table 11 shows the correlations of each of the seven APCA factors against the 12 equity factors presented above. Since the APCA factors represent a rotation of the original portfolio

[20] See Fama and French (2015) for details.

returns, the signs of the correlations are irrelevant and only the magnitudes matter. As expected, F_1 has a large correlation with the aggregate equity premium (-0.99), thus the first principal component represents mainly a market factor. The second factor shows the largest correlations with the value (-0.79) and investment (above 0.60 in magnitude) factors. F_3 is strongly correlated with the three size factors and the two profitability factors with correlations above 0.50 (in magnitude) in all cases. The fourth factor is mainly correlated with the momentum factor as indicated by the correlation of -0.80. F_5 is especially correlated with the three size factors, while the sixth factor has a larger comovement with the two investment factors. The correlations between F_7 and the equity factors are smaller than for the first six APCA factors, with the largest comovement occurring with the two value factors (around 0.20).

These results are largely in line with the evidence from Table 4 above. Specifically, the second APT factor is a mix of value and investment factors, while F_3 is a mix of size and profitability factors. On the other hand, the fourth factor basically captures momentum, while F_6 is correlated with some investment-based strategies. Furthermore, the fact that both F_3 and F_5 are correlated with the size factors suggests that size plays an important role within several of these market anomalies. This is consistent with the evidence provided in Fama and French (2008) showing that several market anomalies are more important among small stocks.

To further understand these correlations, we regress the empirical factors on our seven estimated APT factors. Specifically, in the case of HML we run the following multivariate regression,

$$\begin{aligned} HML_{t+1} = \delta_{HML} + \beta_{HML,1}F_{1,t+1} + \beta_{HML,2}F_{2,t+1} + \beta_{HML,3}F_{3,t+1} + \beta_{HML,4}F_{4,t+1} + \beta_{HML,5}F_{5,t+1} \\ +\beta_{HML,6}F_{6,t+1} + \beta_{HML,7}F_{7,t+1} + \varepsilon_{HML,t+1}, \end{aligned} \tag{23}$$

and similarly for the remaining factors.

The results for the multiple regressions are reported in Table 12. We can see that the seven APT factors explain a large fraction of the time-series variation in each of the empirical factors. The main outlier is the liquidity factor, with a R^2 of only 5%, thus showing that this factor does a poor job in helping capturing the broad cross-section of stock returns. Apart from this isolated case, the R^2 estimates vary between 65% (regression for RMW) and 99% (regression for RM). The large fit in the regression for the market factor comes mainly from the large correlation with F_1, as shown above. Apart from the market factor, the two value factors as well as the momentum factor show the largest explanatory power from the estimated factors with explanatory ratios above 80%. Comparing among related equity factors, we can see that the seven APCA factors are more correlated with ROE than RMW (R^2 of 74% versus 65%), while the regression associated with CMA shows a slightly larger fit than the model corresponding to IA (75% versus 70%).

Apart from the already mentioned case of LIQ, most of the slopes associated with the APT factors are statistically significant at the 5% level. Among the exceptions, the loadings associated with F_4, F_5, and F_6 are not significant in the regressions for both HML and HML^* and the same happens to F_3, F_4, and F_7 in the regression for IA. Moreover, the coefficients associated with F_7 are not significant in the regressions for RMW and CMA, confirming the weaker average correlation of the seventh APT factor with the empirical factors, as shown in Table 11. Overall, the results from the multiple regressions show that there is a large degree of comovement between the empirical and the APT factors.

5.4 Asset pricing tests

We test the multifactor models presented above for the broad cross-section of stock returns. We use the same empirical approach as for the APT model estimated in the last section. Furthermore, and following Maio (2016a) (see also Cochrane (2005)), we also compute the "constrained" cross-sectional R^2 for the empirical models:

$$R_C^2 = 1 - \frac{\mathrm{Var}_N(\hat{\alpha}_{i,C})}{\mathrm{Var}_N(\overline{R_i - R_f})}. \tag{24}$$

This metric is similar to R_{OLS}^2, but it is based on the pricing errors ($\hat{\alpha}_{i,C}$) from a "constrained regression" that restricts the risk price estimates to be equal to the respective factor means. Indeed, when all the factors in a specific model represent excess stock returns (as is the case with the models in this section), the factor risk price estimates should be equal to the respective factor means.

For example, in the case of C4, these pricing errors are obtained from the following equation,

$$\overline{R_i - R_f} = \overline{RM}\beta_{i,M} + \overline{SMB}\beta_{i,SMB} + \overline{HML}\beta_{i,HML} + \overline{UMD}\beta_{i,UMD} + \alpha_{i,C}, \tag{25}$$

where $\overline{RM}$, $\overline{SMB}$, $\overline{HML}$, and $\overline{UMD}$ denote the sample means of the market, size, value, and momentum factors, respectively.

The OLS risk price estimates for the empirical models are presented in Table 13. The results show that the best performing model for the broad cross-section of stock returns is C4 with R_{OLS}^2 and R_C^2 estimates of 49% and 44%, respectively. The HXZ4 model also produces a positive explanatory power as indicated by the R_{OLS}^2 and R_C^2 estimates of 42% and 29%, respectively. The fact that R^2_C is lower than the OLS counterpart by more than 10 points reveals that some of the OLS factor risk price estimates associated with HXZ4 significantly differ from the correct estimates (factor means).

In comparison with these two four-factor models both FF5 and FF4 register a rather modest fit for the cross-sectional equity premia as indicated by the R^2_C estimates of 11% and 5%, respectively. The corresponding OLS explanatory ratios are around 20% for both models, thus indicating that some of the OLS risk price estimates in these two models are incorrect. On the other hand, both FF3 and PS4 register an even weaker performance as indicated by the negative values of R^2_C, which means that these two models perform worse than a trivial model containing just an intercept. The results from Table 13 also show that the performance of the empirical models clearly lags behind the fit of the seven-factor APT estimated in the last section. This suggests that these models still have room for improvement in terms of describing this cross-section of stock returns.

The results associated with the GLS cross-sectional regressions are presented in Table 14. These results show a different picture than the OLS counterparts. The best performing model is still C4 by a good margin with a R^2_{GLS} of 39%. However, the fit of the remaining five models is relatively even with explanatory ratios varying between 16% (HXZ4) and 22% (FF5). Hence, most factor models have an approximate performance when it comes to price an efficient combination of the testing portfolios. These results also illustrate that the performance of asset pricing models can change widely in OLS versus GLS cross-sectional approaches to estimate factor risk premia. On the other hand, in contrast to our benchmark APT, all R^2_{GLS} estimates are clearly below one. This suggests that there is substantial room for improvement for these empirical models to become successful APT applications.

5.5 Improving the factor models

The results from the last subsection suggest that there is a significant margin for improvement of the empirical models in terms of describing the broad cross-section of stock returns and representing valid APT candidates. In light of this evidence, in this subsection we augment these models with new factors to improve the fit for the broad cross-section. To maintain focus, we restrict the analysis in this subsection to C4, HXZ4, FF5, and FF4.

The choice of the new factors relies on APT arguments. Specifically, we use factors related to anomalies that have a relatively weak correlation with the existing empirical factors. For this goal, we run time-series regressions of the return spreads associated with each of the 28 market anomalies on each of the four factor models. The R^2 estimates of these multiple regressions are presented in Table 15. Then we pick the anomalies with lowest R^2 for each model subject to the constraint that the anomalies are associated with different categories (as indicated in Table 1) to avoid excessive overlapping among the factors on a given model. Each of the new models contains seven factors to be consistent with our benchmark APT.

Hence, we estimate the follow augmented version of C4 (C4*), model,

$$\begin{aligned}\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = {} & \lambda_M \beta_{i,M} + \lambda_{SMB} \beta_{i,SMB} + \lambda_{HML} \beta_{i,HML} + \lambda_{UMD} \beta_{i,UMD} \\ & + \lambda_{GPA} \beta_{i,GPA} + \lambda_{OL} \beta_{i,OL} + \lambda_{OA} \beta_{i,OA}, \end{aligned} \tag{26}$$

where λ_{GPA}, λ_{OL}, and λ_{OA} represent the risk prices associated with new factors corresponding to the gross profits-to-assets (GPA), operating leverage (OL), and operating accruals (OA) anomalies.

Similarly, the augmented HXZ4 (HXZ4*) is given by

$$\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = \lambda_M \beta_{i,M} + \lambda_{ME} \beta_{i,ME} + \lambda_{IA} \beta_{i,IA} + \lambda_{ROE} \beta_{i,ROE} + \lambda_{ABR} \beta_{i,ABR} + \lambda_{OL} \beta_{i,OL} + \lambda_{NOA} \beta_{i,NOA}, \tag{27}$$

where λ_{ABR} and λ_{NOA} represent the risk prices associated with new factors corresponding to abnormal stock returns around earnings announcements (ABR) and net operating assets (NOA).

Next, the augmented FF5 (FF5*) incorporates a factor related to abnormal corporate investment (ACI) (in addition to ABR):

$$\begin{aligned}\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = {} & \lambda_M \beta_{i,M} + \lambda_{SMB^*} \beta_{i,SMB^*} + \lambda_{HML^*} \beta_{i,HML^*} + \lambda_{CMA} \beta_{i,CMA} \\ & + \lambda_{RMW} \beta_{i,RMW} + \lambda_{ABR} \beta_{i,ABR} + \lambda_{ACI} \beta_{i,ACI}. \end{aligned} \tag{28}$$

Finally, we add factors related to NOA, revenue surprise (RS), and ABR to the FF4 model (FF4*):

$$\begin{aligned}\mathrm{E}(R_{i,t+1} - R_{f,t+1}) = {} & \lambda_M \beta_{i,M} + \lambda_{SMB^*} \beta_{i,SMB^*} + \lambda_{CMA} \beta_{i,CMA} + \lambda_{RMW} \beta_{i,RMW} \\ & + \lambda_{ABR} \beta_{i,ABR} + \lambda_{NOA} \beta_{i,NOA} + \lambda_{RS} \beta_{i,RS}. \end{aligned} \tag{29}$$

The GPA, OL, ABR, and RS factors represent the return spreads high-minus-low associated with the corresponding decile portfolios, while the OA, NOA, and ACI represent low-minus-high return spreads for the corresponding anomalies.

The results for the OLS risk price estimates associated with the augmented models are shown in Table 16. We can see that the performance of all models rises significantly by adding the new factors, with R^2_C estimates that vary between 43% (HXZ4*) and 57% (C4*). The increase in fit is especially notable in the cases of FF5 and FF4 (around 40 percentage points).

Furthermore, the best performing of the augmented models is C4*, thus preserving the dominance of C4 in the case of the original models. However, while HXZ4 clearly outperforms both FF5 and FF4, it turns out that HXZ4*, FF5*, and FF4* have a similar performance, with the later showing a marginal dominance. We can also see that the performance of C4* is only marginally smaller than the seven-factor APT (57% versus 66%). Turning to the risk price estimates, we can see that the risk price estimates associated with the new factors are statistically significant in most cases. The exceptions are λ_{GPA} (within C4*) and λ_{RS} (within FF4*) while both λ_{OL} and λ_{NOA} in HXZ4* are marginally significant (at the 10% level).

The results for the GLS risk premia estimates are displayed in Table 17. We can see that all four models produce R^2_{GLS} estimates around one, which indicates a sharp increase in fit relative to the original models. Hence, the new factors do a very good job in pricing the transformed portfolios. These results are not totally surprising since the new factors are highly correlated with some of the testing portfolios, thus the GLS cross-sectional regression will assign a large weight to those same portfolios. In other words, the large R^2_{GLS} values show that the augmented factor models represent successful APT applications, similarly to the seven-factor models, and unlike the original factor models.

Overall, these results show that adding the new factors improves all models, and helps especially the performance of both FF5 and FF4 in terms of explaining the large cross-section of stock returns. Moreover, the augmented models do a very good job in explaining an efficient combination of the original portfolios. Therefore, there is a significant room for improving the existing empirical multifactor models in terms of explaining the large cross-section of stock returns in a way that is consistent with the APT.

6 Conclusion

Recent years have noticed an explosion of new market anomalies, and some of the most prominent multifactor models in the empirical asset pricing literature fail to describe well the large cross-section of stock returns. Following such evidence, several questions naturally emerge in the literature: How many factors do we need, and what are these factors, to describe well the broad cross-section of stock returns? To which dimensions of the cross-section of stock returns are these factors more correlated? To what extent (and how) can we improve the current multifactor models proposed in the literature in order to achieve a better description of large-scale cross-sectional risk premia? This paper attempts at providing answers to these questions. In order to achieve this goal, we adopt the general framework of the Arbitrage Pricing Theory (APT) of Ross (1976).

We follow part of the relatively small empirical APT literature in terms of estimating common stock return factors by applying asymptotical principal components analysis (APCA) to a large cross-section of stock returns. We employ a total of 28 anomalies or portfolio sorts, which represent a subset of the anomalies considered in Hou, Xue, and Zhang (2015a, 2015b) for a total of 278 decile portfolios. The estimation results show that there are seven common factors that are statistically significant over our sample period (1972 to 2013). These seven factors cumulatively explain around 91% of the cross-sectional variations in the 278 portfolio returns. The first common factor basically captures the average anomaly and thus resembles a market factor. The other six factors capture different dimensions of the large cross-section of market anomalies. In particular, the second, third, and four factors are strongly correlated with value-growth, investment, profitability, and momentum-based anomalies.

We conduct cross-sectional asset pricing tests of our APT model by using the 278 equity portfolios as testing assets. The results confirm that the seven-factor model explains about 60% of the cross-sectional variation in the risk premia associated with the 278 portfolios. Moreover, most factor risk price estimates are statistically significant. Across categories of anomalies, the APT does a better job in pricing value-growth and intangibles, compared to the group of investment-based anomalies. Moreover, the model prices perfectly an efficient combination of the original portfolios as indicated by the GLS cross-sectional R^2 estimates around 100%. This result confirms that the statistical model is a successful APT.

Next, we compare our APT model to some of most popular multifactor models existent in the literature in terms of pricing the 278 portfolios. The results show that only the four-factor models of Carhart (1997) and Hou, Xue, and Zhang (2015a, 2015b) offer an economically significant explanatory power for the broad cross-section of stock returns, while the fit of the five-factor model of Fama-French (2015, 2016) (and a restricted four-factor version that excludes HML) is quite small. Moreover, the performance of all the empirical factor models clearly lags behind the fit of the seven-factor APT, suggesting that these models have a large room for improvement in terms of describing large-scale cross-sectional risk premia.

In light of such evidence, we define and estimate new empirical multifactor models to better describe the broad cross-section of anomalies. All these models contain seven factors, to be consistent with our benchmark APT. The new factors in each of these models represent factor-mimicking portfolios (spreads among extreme portfolio deciles) associated with selected anomalies. The results show that adding the new factors improves all four empirical models, and helps especially the performance of the five- and four-factor models of Fama-French (2015, 2016) in terms of explaining the large cross-section of stock returns. Moreover, the augmented models do a very good job in explaining an efficient combination of the original portfolios, thus, showing that they represent valid APTs. Therefore, the performance of the augmented empirical models is quite similar to that of our benchmark APT. Overall, our results indicate that there is a significant room for improving the existing empirical multifactor models in terms of explaining the large cross-section of stock returns in a way that is consistent with the APT.

References

Alessi, L., M. Barigozzi, and M. Capasso, 2010, Improved penalization for determining the number of factors in approximate factor models, Statistics and Probability Letters 80, 1806–1813.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang, 2006, The cross-section of volatility and expected returns, Journal of Finance 61, 259–300.

Bai, J., and S. Ng, 2002, Determining the number of factors in approximate factor models, Econometrica 70, 191–221.

Balakrishnan, K., E. Bartov, and L. Faurel, 2010, Post loss/profit announcement drift, Journal of Accounting and Economics 50, 20–41.

Barth, M., J. Elliott, and M. Finn, 1999, Market rewards associated with patterns of increasing earnings, Journal of Accounting Research 37, 387–413.

Basu, S., 1983, The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence, Journal of Financial Economics 12, 129–156.

Belo, F., and X. Lin, 2011, The inventory growth spread, Review of Financial Studies 25, 278–313.

Black, F., M. Jensen, and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, Studies in the theory of capital markets, e.d. by Michael Jensen, Praeger, New York.

Breeden, D., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265–296.

Brennan, M., A. Wang, and Yihong Xia, 2004, Estimation and test of a simple model of intertemporal capital asset pricing, Journal of Finance 59, 1743–1775.

Campbell, J., J. Hilscher, and J. Szilagyi, 2008, In search of distress risk, Journal of Finance 63, 2899–2939.

Campbell, J., and T. Vuolteenaho, 2004, Bad beta, good beta, American Economic Review 94, 1249–1275.

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

Chamberlain, G., and M. Rothschild, 1983, Arbitrage and mean variance analysis on large asset markets, Econometrica 51, 1281–1304.

Chan, L., N. Jegadeesh, and J. Lakonishok, 1996, Momentum strategies, Journal of Finance 51, 1681–1713.

Chan, L., J. Lakonishok, and T. Sougiannis, 2001, The stock market valuation of research and development expenditures, Journal of Finance 56, 2431–2456.

Clarke, C., 2016, The level, slope and curve factor model for stocks, Working paper, University of Connecticut.

Cochrane, J., 2005, Asset pricing (revised edition), Princeton University Press.

Cochrane, J., 2011, Presidential address: Discount rates, Journal of Finance 66, 1047–1108.

Connor, G., and R. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: a new framework for analysis, Journal of Financial Economics 15, 373–394.

Connor, G., and R. Korajczyk, 1988, Risk and return in an equilibrium APT: application of a new test methodology, Journal of Financial Economics 21, 255–289.

Cooper, M., H. Gulen, and M. Schill, 2008, Asset growth and the cross-section of stock returns, Journal of Finance 63, 1609–1651.

Cooper, I., and P. Maio, 2016, Equity risk factors and the Intertemporal CAPM, Working paper, Hanken School of Economics.

Daniel, K., and S. Titman, 2006, Market raections to tangible and intangible information, Journal of Finance 61, 1605–1643.

De Bondt, W., and R. Thaler, 1985, Does the stock market overreact? Journal of Finance 40, 793–805.

Dechow, P., R. Sloan, and M. Soliman, 2004, Implied equity duration: A new measure of equity risk, Review of Accounting Studies 9, 197–228.

Eisfeldt, A., and D. Papanikolaou, 2013, Organizational capital and the cross-section of expected returns, Journal of Finance 68, 1365–1406.

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

Fama, E., and K. French, 1996, Multifactor explanations of asset pricing anomalies, Journal of Finance 51, 55–84.

Fama, E., and K. French, 2008, Dissecting anomalies, Journal of Finance 63, 1653–1678.

Fama, E., and K. French, 2015, A five-factor asset pricing model, Journal of Financial Economics 116, 1–22.

Fama, E., and K. French, 2016, Dissecting anomalies with a five-factor model, Review of Financial Studies 29, 69–103.

Foster, G., C. Olsen, and T. Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, Accounting Review 59, 574–603.

Goyal, A., C. Pérignon, and C. Villa, 2008, How common are common return factors across the NYSE and Nasdaq? Journal of Financial Economics 90, 252–271.

Hafzalla, N., R. Lundholm, and E. Van Winkle, 2011, Percent accruals, Accounting Review 86, 209–236.

Haugen, R., and N. Baker, 1996, Commonality in the determinants of expected stock returns, Journal of Financial Economics 41, 401–439.

Hirshleifer, D., K. Hou, S. Teoh, and Y. Zhang, 2004, Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.

Hou, K., C. Xue, and L. Zhang, 2015a, A comparison of new factor models, Working paper, Ohio State University.

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

Jagannathan, R., and Y. Wang, 2007, Lazy investors, discretionary consumption, and the cross-section of stock returns, Journal of Finance 62, 1623–1661.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, Journal of Finance 51, 3–53.

Jagannathan, R., and Z. Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regressions, Journal of Finance 53, 1285–1309.

Jegadeesh, N., and J. Livnat, 2006, Revenue surprises and stock returns, Journal of Accounting and Economics 41, 147–171.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, Journal of Finance 48, 65–91.

Kan, R., C. Robotti, and J. Shanken, 2013, Pricing model performance and the two-pass cross-sectional regression methodology, Journal of Finance 68, 2617–2649.

Kandel, S., and R. Stambaugh, 1995, Portfolio inefficiency and the cross section of expected returns, Journal of Finance 50, 157–184.

Lakonishok, J., A. Shleifer, and R. Vishny, 1994, Contrarian investment, extrapolation, and risk, Journal of Finance 49, 1541–1578.

Lettau, M., and S. Ludvigson, 2001, Resurrecting the (C)CAPM: A cross sectional test when risk premia are time-varying, Journal of Political Economy 109, 1238–1287.

Lewellen, J., S. Nagel, and J. Shanken, 2010, A skeptical appraisal of asset-pricing tests, Journal of Financial Economics 96, 175–194.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13–37.

Lioui, A., and P. Maio, 2014, Interest rate risk and the cross-section of stock returns, Journal of Financial and Quantitative Analysis 49, 483–511.

Ludvigson, S., 2013, Advances in consumption-based asset pricing: empirical tests, Handbook of the Economics of Finance, 2013, e.d. by G. Constantinides, M. Harris, and R. Stulz, 799–906, Elsevier.

Ludvigson, S., and S. Ng, 2007, The empirical risk-return relation: A factor analysis approach, Journal of Financial Economics 83, 171–222.

Ludvigson, S., and S. Ng, 2009, Macro factors in bond risk premia, Review of Financial Studies 22, 5027–5067.

Ludvigson, S., and S. Ng, 2010, A factor analysis of bond risk premia, Handbook of Empirical Economics and Finance, e.d. by Aman Ulah and David A. Giles, 313–372, Chapman and Hall.

Lyandres, E., L. Sun, and L. Zhang, 2008, The new issues puzzle: Testing the investment-based explanation, Review of Financial Studies 21, 2825–2855.

Maio, P., 2013, Intertemporal CAPM with conditioning variables, Management Science 59, 122–141.

Maio, P., 2016a, Do equity-based factors outperform other risk factors? Evidence from a large cross-section of stock returns, Working paper, Hanken School of Economics.

Maio, P., 2016b, New evidence on conditional factor models, Working paper, Hanken School of Economics.

Maio, P., and D. Philip, 2015, Macro variables and the components of stock returns, Journal of Empirical Finance 33, 287–308.

Maio, P., and P. Santa-Clara, 2012, Multifactor models and their consistency with the ICAPM, Journal of Financial Economics 106, 586–613.

Maio, P., and P. Santa-Clara, 2016, Short-term interest rates and stock market anomalies, Working paper, Hanken School of Economics.

Merton, R., 1973, An intertemporal capital asset pricing model, Econometrica 41, 867–887.

Miller, M., and F. Modigliani, 1961, Dividend policy, growth, and the valuation of shares, Journal of Business 34, 411–433.

Moskowitz, T., and M. Grinblatt, 1999, Do industries explain momentum? Journal of Finance 54, 1249–1290.

Novy-Marx, R., 2011, Operating leverage, Review of Finance 15, 103–134.

Novy-Marx, R., 2013, The other side of value: The gross profitability premium, Journal of Financial Economics 108, 1–28.

Pástor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, Journal of Political Economy 111, 642–684.

Pontiff, J., and A. Woodgate, 2008, Share issuance and cross-sectional returns, Journal of Finance 63, 921–945.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna, 2005, Accrual reliability, earnings persistence and stock prices, Journal of Accounting and Economics 39, 437–485.

Rosenberg, B., K. Reid, and R. Lanstein, 1985, Persuasive evidence of market inefficiency, Journal of Portfolio Management 11, 9–17.

Ross, S., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341–360.

Shanken, J., 1992, On the estimation of beta pricing models, Review of Financial Studies 5, 1–34.

Shanken, J., and G. Zhou, 2007, Estimating and testing beta pricing models: Alternative methods and their performance in simulations, Journal of Financial Economics 84, 40–86.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.

Sloan, R., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? Accounting Review 71, 289–315.

Stock, J., and M. Watson, 2002a, Forecasting using principal components from a large number of predictors, Journal of the American Statistical Association 97, 1167–1179.

Stock, J., and M. Watson, 2002b, Macroeconomic forecasting using diffusion indexes, Journal of Business and Economic Statistics 20, 147–162.

Thomas, J., and H. Zhang, 2002, Inventory changes and future returns, Review of Accounting Studies 7, 163–187.

Titman, S., K. Wei, and F. Xie, 2004, Capital investments and stock returns, Journal of Financial and Quantitative Analysis 39, 677–700.

Xing, Y., 2008, Interpreting the value effect through the Q-theory: An empirical investigation, Review of Financial Studies 21, 1767–1795.

Yogo, M., 2006, A consumption-based explanation of expected stock returns, Journal of Finance 61, 539–580.

Table 1: List of portfolio sorts

This table lists the 28 alternative anomalies or portfolio sorts employed in the empirical analysis. "Category" refers to the broad classification employed by Hou, Xue, and Zhang (2015b), and "#" represents the number of portfolios in each group. "Reference" shows the paper that represents the source of the anomaly.

Symbol	Anomaly	Category	#	Reference
BM	Book-to-market equity	Value-growth	10	Rosenberg, Reid, and Lanstein (1985)
MOM	Price momentum (11-month prior returns)	Momentum	10	Fama and French (1996)
IA	Investment-to-assets	Investment	10	Cooper, Gulen, and Schill (2008)
ROE	Return on equity	Profitability	10	Haugen and Baker (1996)
SUE	Earnings surprise (1-month holding period)	Momentum	10	Foster, Olsen, and Shevlin (1984)
GPA	Gross profits-to-assets	Profitability	10	Novy-Marx (2013)
NSI	Net stock issues	Investment	10	Pontiff and Woodgate (2008)
OCA	Organizational capital-to-assets	Intangibles	10	Eisfeldt and Papanikolaou (2013)
OL	Operating leverage	Intangibles	10	Novy-Marx (2011)
ABR	Cumulative abnormal stock returns around earnings announcements (1-month)	Momentum	10	Chan, Jegadeesh, and Lakonishok (1996)
CEI	Composite issuance	Investment	10	Daniel and Titman (2006)
PIA	Changes in property, plant, and equipment scaled by assets	Investment	10	Lyandres, Sun, and Zhang (2008)
DUR	Equity duration	Value-growth	10	Dechow, Sloan, and Soliman (2004)
IG	Investment growth	Investment	10	Xing (2008)
IVC	Inventory changes	Investment	10	Thomas and Zhang (2002)
IVG	Inventory growth	Investment	10	Belo and Lin (2011)
NOA	Net operating assets	Investment	10	Hirshleifer et al. (2004)
OA	Operating accruals	Investment	10	Sloan (1996)
POA	Percent operating accruals	Investment	10	Hafzalla, Lundholm, and Van Winkle (2011)
PTA	Percent total accruals	Investment	10	Hafzalla, Lundholm, and Van Winkle (2011)
IM	Industry momentum	Momentum	9	Moskowitz and Grinblatt (1999)
NEI	Number of consecutive quarters with earnings increases	Profitability	9	Barth, Elliott, and Finn (1999)
ABR*	Cumulative abnormal stock returns around earnings announcements (6-month)	Momentum	10	Chan, Jegadeesh, and Lakonishok (1996)
CFP	Cash flow-to-price	Value-growth	10	Lakonishok, Shleifer, and Vishny (1994)
RS	Revenue surprise	Profitability	10	Jegadeesh and Livnat (2006)
EP	Earnings-to-price	Value-growth	10	Basu (1983)
REV	Return reversal	Value-growth	10	De Bondt and Thaler (1985)
ACI	Abnormal corporate investment	Investment	10	Titman, Wei, and Xie (2004)

Table 2: Descriptive statistics for spreads in returns

This table reports descriptive statistics for the "high-minus-low" spreads in returns associated with different portfolio classes. See Table 1 for a description of the different portfolio sorts. The sample is 1972:01–2013:12. ϕ designates the first-order autocorrelation coefficient.

	Mean (%)	Stdev. (%)	Min. (%)	Max. (%)	ϕ
BM	0.69	4.86	−14.18	20.45	0.11
DUR	−0.52	4.34	−21.38	15.77	0.09
CFP	0.49	4.66	−18.95	16.26	0.02
EP	0.58	4.83	−15.47	22.53	0.02
REV	−0.41	5.21	−32.99	18.08	0.06
MOM	1.17	7.21	−61.35	26.30	0.05
SUE	0.44	3.05	−14.27	12.09	−0.00
ABR	0.73	3.17	−15.80	15.32	−0.10
IM	0.54	5.09	−33.33	20.27	0.05
ABR*	0.30	2.08	−10.45	9.86	−0.01
ROE	0.75	5.28	−26.37	29.30	0.16
GPA	0.34	3.36	−13.55	12.35	0.04
NEI	0.36	2.79	−12.10	12.21	0.00
RS	0.30	3.46	−12.85	20.08	0.07
IA	−0.42	3.62	−14.39	11.83	0.04
ACI	−0.26	3.21	−18.13	13.93	0.13
NSI	−0.69	3.28	−20.47	12.88	0.10
CEI	−0.55	4.06	−16.34	17.94	0.06
PIA	−0.49	3.00	−10.37	8.60	0.08
IG	−0.38	2.83	−12.81	9.67	0.07
IVC	−0.43	3.19	−12.21	11.64	0.06
IVG	−0.36	3.15	−9.69	12.04	0.07
NOA	−0.39	3.11	−14.26	13.45	0.02
OA	−0.27	3.10	−10.39	12.81	−0.01
POA	−0.43	3.12	−11.84	19.87	0.06
PTA	−0.40	3.38	−11.25	19.13	0.01
OCA	0.55	3.13	−13.68	13.60	−0.02
OL	0.39	3.86	−10.34	17.37	0.11

Table 3: Descriptive statistics for common factors

This table reports descriptive statistics for the common factors ($F_j, j = 1, ..., 7$) estimated from 278 equity portfolios. ϕ designates the first-order autocorrelation coefficient. R_j^2 represents the cumulative proportion of the cross-sectional variance in the raw portfolio returns explained by the factors F_1 to F_j. $\overline{R^2}$ denotes the average R^2 among time-series regressions of the 278 portfolio returns on an increasing number of factors as regressors, for example the column labeled F_3 corresponds to a regression that contains F_1, F_2, and F_3 as regressors. The sample is 1972:01–2013:12.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7
ϕ	0.06	0.06	0.04	0.03	0.05	0.07	0.11
R_j^2	0.85	0.87	0.88	0.89	0.90	0.90	0.91
$\overline{R^2}$	0.85	0.87	0.88	0.89	0.89	0.90	0.90

Table 4: Anomalies and common factors

This table reports R^2 estimates from single regressions of return spreads onto the estimated common factors (F_j). The "high-minus-low" spreads in returns are associated with 28 market anomalies. See Table 1 for a description of the different portfolio sorts. The sample is 1972:01–2013:12.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7
BM	0.01	0.46	0.26	0.02	0.00	0.02	0.01
DUR	0.02	0.41	0.15	0.01	0.01	0.02	0.03
CFP	0.05	0.51	0.11	0.04	0.00	0.04	0.04
EP	0.03	0.52	0.06	0.01	0.03	0.04	0.06
REV	0.00	0.12	0.29	0.00	0.02	0.19	0.02
MOM	0.02	0.01	0.16	0.64	0.01	0.01	0.01
SUE	0.03	0.01	0.07	0.11	0.01	0.03	0.02
ABR	0.01	0.02	0.02	0.13	0.00	0.01	0.01
IM	0.01	0.01	0.08	0.60	0.01	0.00	0.01
ABR*	0.00	0.07	0.01	0.21	0.00	0.01	0.04
ROE	0.08	0.02	0.62	0.00	0.02	0.02	0.00
GPA	0.00	0.06	0.21	0.03	0.12	0.01	0.01
NEI	0.00	0.12	0.40	0.03	0.04	0.00	0.00
RS	0.00	0.14	0.33	0.01	0.02	0.00	0.00
IA	0.05	0.30	0.01	0.00	0.01	0.18	0.00
ACI	0.00	0.03	0.00	0.07	0.02	0.02	0.00
NSI	0.06	0.20	0.10	0.00	0.00	0.01	0.03
CEI	0.24	0.35	0.04	0.00	0.02	0.00	0.02
PIA	0.04	0.06	0.01	0.01	0.05	0.30	0.01
IG	0.02	0.13	0.00	0.01	0.00	0.14	0.00
IVC	0.03	0.03	0.03	0.04	0.12	0.17	0.01
IVG	0.04	0.16	0.00	0.02	0.10	0.08	0.00
NOA	0.00	0.09	0.03	0.01	0.01	0.12	0.08
OA	0.01	0.02	0.00	0.04	0.23	0.00	0.00
POA	0.04	0.23	0.00	0.01	0.08	0.04	0.01
PTA	0.08	0.22	0.02	0.01	0.02	0.09	0.02
OCA	0.07	0.04	0.11	0.04	0.10	0.02	0.00
OL	0.01	0.01	0.05	0.02	0.35	0.04	0.08

Table 5: APT model: OLS factor risk premia estimates

This table reports the factor risk price estimates for the APT seven-factor model (and corresponding nested models). The common factors are estimated from APCA applied to 278 equity portfolios. The empirical method is the two-step regression approach where the second step consists of an OLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_j denotes the risk price estimate (in %) for the jth common factor (F_j). Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{OLS} denotes the cross-sectional OLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	R^2_{OLS}
1	−12.42							−0.46
	(**−2.78**)							
2	−12.77	−15.02						0.02
	(**−2.86**)	(**−3.36**)						
3	−12.98	−14.76	9.13					0.16
	(**−2.91**)	(**−3.31**)	(<u>2.04</u>)					
4	−13.02	−15.23	7.89	−17.32				0.46
	(**−2.92**)	(**−3.41**)	(1.77)	(**−3.87**)				
5	−13.02	−15.23	7.90	−17.30	0.74			0.46
	(**−2.92**)	(**−3.41**)	(1.77)	(**−3.87**)	(0.17)			
6	−13.04	−14.97	8.22	−16.92	0.74	−13.31		0.55
	(**−2.92**)	(**−3.36**)	(1.84)	(**−3.78**)	(0.17)	(**−2.97**)		
7	−13.11	−14.77	8.07	−16.56	0.73	−13.01	14.32	0.62
	(**−2.94**)	(**−3.31**)	(1.81)	(**−3.70**)	(0.16)	(**−2.91**)	(**3.20**)	

Table 6: APT model: OLS factor risk premia estimates across categories

This table reports the factor risk price estimates for the APT seven-factor model. The common factors are estimated from APCA applied to 278 equity portfolios. The empirical method is the two-step regression approach where the second step consists of an OLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets are combinations of 28 different portfolios sorts that correspond to the categories defined in Table 1. For example, the value-growth category contains the BM, DUR, CFP, EP, and REV deciles. See Table 1 for a description of the different portfolio sorts and categories. λ_j denotes the risk price estimate (in %) for the jth common factor (F_j). Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{OLS} denotes the cross-sectional OLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	R^2_{OLS}
Panel A: Value-growth							
−13.21	−13.00	2.30	−23.51	3.65	−8.81	14.45	0.89
(−**2.95**)	(−**2.72**)	(0.43)	(−**2.86**)	(0.47)	(−1.53)	(<u>2.29</u>)	
Panel B: Momentum							
−12.68	−4.75	10.49	−9.33	6.04	−25.55	11.38	0.74
(−**2.84**)	(−0.87)	(<u>2.04</u>)	(−1.85)	(1.02)	(−**2.67**)	(1.48)	
Panel C: Profitability							
−12.88	−12.92	18.99	−9.08	1.69	−8.90	4.92	0.75
(−**2.88**)	(<u>−2.23</u>)	(**3.49**)	(−1.23)	(0.19)	(−0.98)	(0.68)	
Panel D: Investment							
−12.80	−10.24	13.12	−24.17	−5.22	−17.40	12.88	0.62
(−**2.87**)	(<u>−2.16</u>)	(<u>2.50</u>)	(−**3.87**)	(−1.08)	(−**3.46**)	(<u>2.14</u>)	
Panel E: Intangibles							
−12.74	−5.06	13.37	−28.40	13.17	−16.35	−0.29	0.80
(−**2.80**)	(−0.54)	(<u>2.08</u>)	(−**2.72**)	(1.72)	(−1.28)	(−0.03)	

Table 7: APT model: GLS factor risk premia estimates

This table reports the factor risk price estimates for the APT seven-factor model (and corresponding nested models). The common factors are estimated from APCA applied to 278 equity portfolios. The empirical method is the two-step regression approach where the second step consists of an GLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_j denotes the risk price estimate (in %) for the jth common factor (F_j). Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{GLS} denotes the cross-sectional GLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	R^2_{GLS}
1	−12.72							−0.06
	(−**2.85**)							
2	−12.73	−14.30						1.00
	(−**2.85**)	(−**3.21**)						
3	−12.73	−14.30	7.53					1.00
	(−**2.86**)	(−**3.21**)	(1.69)					
4	−12.73	−14.31	7.53	−16.98				1.00
	(−**2.86**)	(−**3.21**)	(1.69)	(−**3.81**)				
5	−12.73	−14.31	7.53	−16.98	0.68			1.00
	(−**2.86**)	(−**3.21**)	(1.69)	(−**3.81**)	(0.15)			
6	−12.73	−14.31	7.53	−16.98	0.68	−12.88		1.00
	(−**2.86**)	(−**3.21**)	(1.69)	(−**3.81**)	(0.15)	(−**2.89**)		
7	−12.74	−14.31	7.53	−16.98	0.68	−12.88	13.79	1.00
	(−**2.86**)	(−**3.21**)	(1.69)	(−**3.81**)	(0.15)	(−**2.89**)	(**3.10**)	

Table 8: APT model: GLS factor risk premia estimates across categories

This table reports the factor risk price estimates for the APT seven-factor model. The common factors are estimated from APCA applied to 278 equity portfolios. The empirical method is the two-step regression approach where the second step consists of an GLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets are combinations of 28 different portfolios sorts that correspond to the categories defined in Table 1. For example, the value-growth category contains the BM, DUR, CFP, EP, and REV deciles. See Table 1 for a description of the different portfolio sorts and categories. λ_j denotes the risk price estimate (in %) for the jth common factor (F_j). Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{GLS} denotes the cross-sectional GLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	R^2_{GLS}
Panel A: Value-growth							
−12.88	−12.57	3.39	−25.53	6.83	−8.77	16.27	0.91
(−**2.88**)	(−**2.70**)	(0.67)	(−**3.38**)	(1.08)	(−1.63)	(**2.83**)	
Panel B: Momentum							
−12.65	−6.17	11.08	−9.78	4.53	−12.83	7.49	0.77
(−**2.83**)	(−1.21)	(<u>2.24</u>)	(<u>−2.04</u>)	(0.83)	(−1.64)	(1.08)	
Panel C: Profitability							
−12.91	−12.81	16.43	−8.46	5.13	−14.20	2.21	0.68
(−**2.89**)	(<u>−2.32</u>)	(**3.16**)	(−1.28)	(0.65)	(−1.83)	(0.33)	
Panel D: Investment							
−12.57	−12.65	9.97	−20.16	−3.69	−14.53	7.29	0.70
(−**2.82**)	(−**2.76**)	(<u>2.07</u>)	(−**3.79**)	(−0.80)	(−**3.09**)	(1.42)	
Panel E: Intangibles							
−12.33	−1.64	13.90	−21.27	14.73	−15.15	−5.60	0.66
(−**2.72**)	(−0.18)	(<u>2.27</u>)	(<u>−2.21</u>)	(<u>1.99</u>)	(−1.26)	(−0.51)	

Table 9: Descriptive statistics for equity factors

This table reports descriptive statistics for the equity factors from alternative factor models. RM, SMB/SMB^*, HML/HML^*, UMD, and LIQ denote the market, size, value, momentum, and liquidity factors, respectively. ME, IA, and ROE represent the Hou-Xue-Zhang size, investment, and profitability factors, respectively. RMW and CMA denote the Fama-French profitability and investment factors. The sample is 1972:01–2013:12. ϕ designates the first-order autocorrelation coefficient.

	Mean (%)	Stdev. (%)	Min. (%)	Max. (%)	ϕ
RM	0.53	4.61	−23.24	16.10	0.08
SMB	0.20	3.13	−16.39	22.02	0.01
HML	0.39	3.01	−12.68	13.83	0.15
UMD	0.71	4.46	−34.72	18.39	0.07
LIQ	0.43	3.57	−10.14	21.01	0.09
ME	0.31	3.14	−14.45	22.41	0.03
IA	0.44	1.87	−7.13	9.41	0.06
ROE	0.57	2.62	−13.85	10.39	0.10
SMB^*	0.23	3.07	−15.26	19.05	0.03
HML^*	0.40	3.00	−12.61	13.88	0.15
RMW	0.29	2.25	−17.60	12.24	0.18
CMA	0.37	1.96	−6.76	8.93	0.14

Table 10: Correlations of equity factors

This table reports the pairwise correlations among the equity factors from alternative factor models. RM, SMB/SMB^*, HML/HML^*, UMD, and LIQ denote the market, size, value, momentum, and liquidity factors, respectively. ME, IA, and ROE represent the Hou-Xue-Zhang size, investment, and profitability factors, respectively. RMW and CMA denote the Fama-French profitability and investment factors. The sample is 1972:01–2013:12.

	RM	SMB	HML	UMD	LIQ	ME	IA	ROE	SMB^*	HML^*	RMW	CMA
RM	1.00	0.28	−0.32	−0.14	−0.04	0.25	−0.36	−0.18	0.25	−0.32	−0.23	−0.39
SMB		1.00	−0.23	−0.01	−0.02	0.95	−0.23	−0.39	0.99	−0.23	−0.44	−0.12
HML			1.00	−0.15	0.04	−0.07	0.69	−0.09	−0.11	1.00	0.15	0.70
UMD				1.00	−0.04	0.00	0.04	0.50	−0.03	−0.15	0.09	0.02
LIQ					1.00	−0.03	0.03	−0.07	−0.02	0.04	0.02	0.04
ME						1.00	−0.12	−0.31	0.98	−0.07	−0.38	−0.01
IA							1.00	0.06	−0.15	0.69	0.10	0.90
ROE								1.00	−0.38	−0.09	0.67	−0.09
SMB^*									1.00	−0.11	−0.39	−0.05
HML^*										1.00	0.15	0.70
RMW											1.00	−0.03
CMA												1.00

Table 11: Correlations between PCA and equity factors

This table reports pairwise correlations between the estimated PCA common factors (F_j) and equity factors. RM, SMB/SMB^*, HML/HML^*, UMD, and LIQ denote the market, size, value, momentum, and liquidity factors, respectively. ME, IA, and ROE represent the Hou-Xue-Zhang size, investment, and profitability factors, respectively. RMW and CMA denote the Fama-French profitability and investment factors. The sample is 1972:01–2013:12.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7
RM	−0.99	0.04	0.01	0.01	−0.02	−0.01	0.02
SMB	−0.28	0.22	−0.59	−0.19	0.40	−0.22	−0.08
HML	0.30	−0.79	−0.18	0.01	−0.01	−0.03	0.22
UMD	0.15	0.13	0.40	−0.80	0.09	−0.05	0.06
LIQ	0.02	−0.06	−0.10	−0.03	−0.03	0.14	0.12
ME	−0.25	0.06	−0.56	−0.22	0.49	−0.22	−0.07
IA	0.35	−0.65	0.03	−0.07	−0.11	−0.37	−0.01
ROE	0.17	−0.02	0.76	−0.20	0.26	0.09	−0.11
SMB^*	−0.26	0.10	−0.61	−0.19	0.45	−0.20	−0.09
HML^*	0.30	−0.79	−0.18	0.01	−0.01	−0.03	0.22
RMW	0.21	−0.25	0.59	0.19	0.33	0.22	0.00
CMA	0.37	−0.62	−0.11	−0.10	−0.13	−0.42	−0.01

Table 12: Regressions of equity factors onto PCA factors

This table reports results for multiple regressions of equity factors onto the estimated PCA common factors (F_j) and equity factors. RM, SMB/SMB^*, HML/HML^*, UMD, and LIQ denote the market, size, value, momentum, and liquidity factors, respectively. ME, IA, and ROE represent the Hou-Xue-Zhang size, investment, and profitability factors, respectively. RMW and CMA denote the Fama-French profitability and investment factors. The first line associated with each factor reports the slope estimates, while the second line displays heteroskedasticity-robust t-ratios (in parentheses). R^2 denotes the coefficient of determination. The sample is 1972:01–2013:12. Bolded and underlined t-ratios denote statistical significance at the 1% and 5% levels, respectively.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	R^2
RM	−0.046	0.002	0.000	0.000	−0.001	−0.000	0.001	0.99
	(**−166.03**)	(**5.32**)	(1.12)	(1.05)	(**−3.54**)	(−0.88)	(**3.03**)	
SMB	−0.009	0.007	−0.018	−0.006	0.012	−0.007	−0.003	0.72
	(**−10.46**)	(**6.33**)	(**−19.96**)	(**−5.55**)	(**15.44**)	(**−7.64**)	(**−2.91**)	
HML	0.009	−0.024	−0.005	0.000	−0.000	−0.001	0.007	0.81
	(**11.75**)	(**−27.33**)	(**−6.54**)	(0.45)	(−0.67)	(−1.05)	(**9.99**)	
UMD	0.007	0.006	0.018	−0.036	0.004	−0.002	0.003	0.86
	(**6.76**)	(**4.51**)	(**15.05**)	(**−34.88**)	(**3.47**)	(−2.43)	(**2.62**)	
LIQ	0.001	−0.002	−0.004	−0.001	−0.001	0.005	0.004	0.05
	(0.32)	(−0.95)	(−1.69)	(−0.70)	(−0.54)	(**2.96**)	(2.27)	
ME	−0.008	0.002	−0.018	−0.007	0.015	−0.007	−0.002	0.72
	(**−9.74**)	(1.46)	(**−18.09**)	(**−5.59**)	(**18.37**)	(**−6.77**)	(−2.41)	
IA	0.007	−0.012	0.000	−0.001	−0.002	−0.007	−0.000	0.70
	(**11.88**)	(**−21.61**)	(0.59)	(−1.41)	(**−3.65**)	(**−12.64**)	(−0.22)	
ROE	0.005	−0.001	0.020	−0.005	0.007	0.002	−0.003	0.74
	(**6.22**)	(−0.62)	(**26.36**)	(**−5.97**)	(**8.01**)	(**3.49**)	(**−3.83**)	
SMB^*	−0.008	0.003	−0.019	−0.006	0.014	−0.006	−0.003	0.73
	(**−9.89**)	(**2.95**)	(**−22.50**)	(**−5.95**)	(**17.83**)	(**−7.23**)	(**−3.20**)	
HML^*	0.009	−0.024	−0.005	0.000	−0.000	−0.001	0.007	0.81
	(**11.78**)	(**−27.47**)	(**−6.54**)	(0.43)	(−0.66)	(−1.04)	(**10.02**)	
RMW	0.005	−0.006	0.013	0.004	0.007	0.005	0.000	0.65
	(**7.03**)	(**−5.26**)	(**13.64**)	(**4.58**)	(**8.48**)	(**6.38**)	(0.15)	
CMA	0.007	−0.012	−0.002	−0.002	−0.003	−0.008	−0.000	0.75
	(**14.66**)	(**−17.29**)	(**−3.17**)	(**−2.98**)	(**−4.38**)	(**−15.95**)	(−0.30)	

Table 13: Empirical models: OLS factor risk premia estimates

This table reports the factor risk price estimates for the empirical multifactor models. The empirical method is the two-step regression approach where the second step consists of an OLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_M, $\lambda_{SMB}/\lambda_{SMB^*}$, $\lambda_{HML}/\lambda_{HML^*}$, λ_{UMD}, and λ_{LIQ} denote the risk price estimates (in %) for the market, size, value, momentum, and liquidity factors, respectively. λ_{ME}, λ_{IA}, and λ_{ROE} represent the risk prices associated with the Hou-Xue-Zhang size, investment, and profitability factors, respectively. λ_{RMW} and λ_{CMA} denote the risk price estimates for the Fama-French profitability and investment factors. Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{OLS} denotes the cross-sectional OLS R^2 while R^2_C is the cross-sectional constrained R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

	λ_M	$\lambda_{SMB}/\lambda_{SMB^*}$	$\lambda_{HML}/\lambda_{HML^*}$	λ_{UMD}	λ_{LIQ}	λ_{ME}	λ_{IA}	λ_{ROE}	λ_{RMW}	λ_{CMA}	R^2_{OLS}	R^2_C
1	0.59	−0.22	0.35								0.02	−0.16
	(**2.84**)	(−1.28)	(<u>2.33</u>)									
2	0.60	−0.05	0.39	0.70							0.49	0.44
	(**2.88**)	(−0.31)	(<u>2.52</u>)	(**3.32**)								
3	0.60	−0.20	0.37		−0.38						0.03	−0.22
	(**2.88**)	(−1.18)	(<u>2.39</u>)		(−1.04)							
4	0.57					0.25	0.32	0.35			0.42	0.29
	(**2.78**)					(1.52)	(**2.98**)	(**2.60**)				
5	0.55	0.08	0.19						0.16	0.32	0.24	0.11
	(**2.66**)	(0.48)	(1.25)						(1.29)	(**3.18**)		
6	0.56	0.01							0.16	0.24	0.21	0.05
	(**2.72**)	(0.07)							(1.34)	(<u>2.22</u>)		

Table 14: Empirical models: GLS factor risk premia estimates

This table reports the factor risk price estimates for the empirical multifactor models. The empirical method is the two-step regression approach where the second step consists of an GLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_M, $\lambda_{SMB}/\lambda_{SMB^*}$, $\lambda_{HML}/\lambda_{HML^*}$, λ_{UMD}, and λ_{LIQ} denote the risk price estimates (in %) for the market, size, value, momentum, and liquidity factors, respectively. λ_{ME}, λ_{IA}, and λ_{ROE} represent the risk prices associated with the Hou-Xue-Zhang size, investment, and profitability factors, respectively. λ_{RMW} and λ_{CMA} denote the risk price estimates for the Fama-French profitability and investment factors. Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{GLS} denotes the cross-sectional GLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

	λ_M	$\lambda_{SMB}/\lambda_{SMB^*}$	$\lambda_{HML}/\lambda_{HML^*}$	λ_{UMD}	λ_{LIQ}	λ_{ME}	λ_{IA}	λ_{ROE}	λ_{RMW}	λ_{CMA}	R^2_{GLS}
FF3	0.57	0.14	0.38								0.17
	(**2.77**)	(1.01)	(**2.77**)								
C4	0.57	0.15	0.39	0.62							0.39
	(**2.77**)	(1.03)	(**2.80**)	(**3.09**)							
PS4	0.57	0.14	0.38		0.19						0.18
	(**2.76**)	(1.00)	(**2.75**)		(1.01)						
HXZ4	0.56					0.25	0.24	0.26			0.16
	(**2.74**)					(1.75)	(**2.78**)	(<u>2.16</u>)			
FF5	0.56	0.18	0.37						0.11	0.28	0.22
	(**2.74**)	(1.25)	(**2.66**)						(1.04)	(**3.13**)	
FF4	0.56	0.18							0.11	0.28	0.17
	(**2.74**)	(1.26)							(1.04)	(**3.14**)	

Table 15: Anomalies and empirical factors

This table reports R^2 estimates from multiple regressions of return spreads onto factors associated with alternative multifactor models. The "high-minus-low" spreads in returns are associated with 28 market anomalies. See Table 1 for a description of the different portfolio sorts. The models are the Carhart four-factor model (C4), Hou-Xue-Zhang four-factor model (HXZ4), Fama-French five-factor model (FF5), and a restricted version of FF5 (FF4). The sample is 1972:01–2013:12.

	C4	HXZ4	FF5	FF4
BM	0.69	0.47	0.72	0.47
DUR	0.51	0.21	0.53	0.21
CFP	0.59	0.24	0.60	0.26
EP	0.56	0.20	0.59	0.23
REV	0.45	0.52	0.50	0.46
MOM	0.86	0.28	0.07	0.03
SUE	0.14	0.19	0.06	0.04
ABR	0.13	0.06	0.04	0.02
IM	0.62	0.18	0.04	0.02
ABR*	0.23	0.07	0.06	0.05
ROE	0.43	0.78	0.63	0.62
GPA	0.09	0.18	0.36	0.27
NEI	0.33	0.43	0.34	0.27
RS	0.28	0.33	0.27	0.19
IA	0.34	0.52	0.52	0.51
ACI	0.11	0.08	0.08	0.06
NSI	0.28	0.33	0.45	0.44
CEI	0.57	0.53	0.62	0.57
PIA	0.13	0.27	0.31	0.31
IG	0.16	0.29	0.25	0.25
IVC	0.09	0.18	0.22	0.22
IVG	0.19	0.33	0.30	0.30
NOA	0.08	0.00	0.11	0.02
OA	0.03	0.10	0.10	0.10
POA	0.27	0.33	0.35	0.34
PTA	0.28	0.34	0.35	0.34
OCA	0.17	0.28	0.22	0.20
OL	0.02	0.15	0.24	0.24

Table 16: Augmented empirical models: OLS factor risk premia estimates

This table reports the factor risk price estimates for the augmented empirical multifactor models with seven factors each. The empirical method is the two-step regression approach where the second step consists of an OLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_M, $\lambda_{SMB}/\lambda_{SMB^*}$, $\lambda_{HML}/\lambda_{HML^*}$, and λ_{UMD} denote the risk price estimates (in %) for the market, size, value, and momentum factors, respectively. λ_{ME}, λ_{IA}, and λ_{ROE} represent the risk prices associated with the Hou-Xue-Zhang size, investment, and profitability factors, respectively. λ_{RMW} and λ_{CMA} denote the risk price estimates for the Fama-French profitability and investment factors. λ_{GPA}, λ_{OL}, λ_{OA}, λ_{ABR}, λ_{NOA}, λ_{ACI}, and λ_{RS} denote the risk prices for factors (long-short portfolios) associated with the GPA, OL, OA, ABR, NOA, ACI, and RS anomalies, respectively. Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{OLS} denotes the cross-sectional OLS R^2 while R^2_C is the cross-sectional constrained R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

Panel A: C4								
λ_M	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_{GPA}	λ_{OL}	λ_{OA}	R^2_{OLS}	R^2_C
0.59	−0.01	0.46	0.64	0.26	0.54	0.44	0.61	0.57
(**2.84**)	(−0.09)	(**3.00**)	(**3.04**)	(1.49)	(**2.76**)	(**2.58**)		
Panel B: HXZ4								
λ_M	λ_{ME}	λ_{IA}	λ_{ROE}	λ_{ABR}	λ_{OL}	λ_{NOA}	R^2_{OLS}	R^2_C
0.59	0.18	0.39	0.27	0.70	0.38	0.34	0.58	0.43
(**2.84**)	(1.10)	(**3.80**)	(<u>1.98</u>)	(**3.64**)	(1.85)	(1.94)		
Panel C: FF5								
λ_M	λ_{SMB^*}	λ_{HML^*}	λ_{RMW}	λ_{CMA}	λ_{ACI}	λ_{ABR}	R^2_{OLS}	R^2_C
0.58	0.06	0.33	0.10	0.26	0.51	0.86	0.60	0.47
(**2.80**)	(0.37)	(<u>2.18</u>)	(0.83)	(<u>2.55</u>)	(**2.75**)	(**3.90**)		
Panel D: FF4								
λ_M	λ_{SMB^*}	λ_{RMW}	λ_{CMA}	λ_{NOA}	λ_{RS}	λ_{ABR}	R^2_{OLS}	R^2_C
0.56	0.14	0.12	0.35	0.40	−0.00	0.85	0.59	0.50
(**2.72**)	(0.87)	(1.01)	(**3.42**)	(<u>2.41</u>)	(−0.01)	(**3.79**)		

Table 17: Augmented empirical models: GLS factor risk premia estimates

This table reports the factor risk price estimates for the augmented empirical multifactor models with seven factors each. The empirical method is the two-step regression approach where the second step consists of an GLS cross-sectional regression of average portfolio excess returns on factor betas. The testing assets represent 278 portfolios associated with 28 portfolio sorts. See Table 1 for a description of the different portfolio sorts. λ_M, $\lambda_{SMB}/\lambda_{SMB^*}$, $\lambda_{HML}/\lambda_{HML^*}$, and λ_{UMD} denote the risk price estimates (in %) for the market, size, value, and momentum factors, respectively. λ_{ME}, λ_{IA}, and λ_{ROE} represent the risk prices associated with the Hou-Xue-Zhang size, investment, and profitability factors, respectively. λ_{RMW} and λ_{CMA} denote the risk price estimates for the Fama-French profitability and investment factors. λ_{GPA}, λ_{OL}, λ_{OA}, λ_{ABR}, λ_{NOA}, λ_{ACI}, and λ_{RS} denote the risk prices for factors (long-short portfolios) associated with the GPA, OL, OA, ABR, NOA, ACI, and RS anomalies, respectively. Below the risk price estimates are displayed t-statistics based on Shanken's standard errors (in parentheses). The column labeled R^2_{GLS} denotes the cross-sectional GLS R^2. The sample is 1972:01–2012:12. Underlined and bold t-ratios denote statistical significance at the 5% and 1% levels, respectively.

Panel A: C4

λ_M	λ_{SMB}	λ_{HML}	λ_{UMD}	λ_{GPA}	λ_{OL}	λ_{OA}	R^2_{GLS}
0.74	−0.67	3.61	0.45	0.34	0.37	0.27	1.00
(**3.62**)	(**−4.82**)	(**26.88**)	($\underline{2.19}$)	($\underline{2.29}$)	($\underline{2.17}$)	(1.94)	

Panel B: HXZ4

λ_M	λ_{ME}	λ_{IA}	λ_{ROE}	λ_{ABR}	λ_{OL}	λ_{NOA}	R^2_{GLS}
0.65	0.03	0.44	0.25	0.73	0.39	0.39	1.00
(**3.18**)	(0.18)	(**5.13**)	($\underline{2.09}$)	(**5.15**)	($\underline{2.27}$)	(**2.80**)	

Panel C: FF5

λ_M	λ_{SMB^*}	λ_{HML^*}	λ_{RMW}	λ_{CMA}	λ_{ACI}	λ_{ABR}	R^2_{GLS}
0.56	0.35	0.39	0.11	0.30	0.26	0.73	1.00
(**2.71**)	($\underline{2.46}$)	(**2.81**)	(1.08)	(**3.31**)	(1.84)	(**5.15**)	

Panel D: FF4

λ_M	λ_{SMB^*}	λ_{RMW}	λ_{CMA}	λ_{NOA}	λ_{RS}	λ_{ABR}	R^2_{GLS}
0.67	0.17	0.01	0.11	0.39	0.30	0.73	1.00
(**3.28**)	(1.20)	(0.08)	(1.28)	(**2.81**)	($\underline{1.97}$)	(**5.15**)	