Interpreting factor models

SERHIY KOZAK, STEFAN NAGEL, and SHRIHARI SANTOSH[*]

May 2, 2017

ABSTRACT

We argue that tests of reduced-form factor models and horse races between "characteristics" and "covariances" cannot discriminate between alternative models of investor beliefs. Since asset returns have substantial commonality, absence of near-arbitrage opportunities implies that the SDF can be represented as a function of a few dominant sources of return variation. As long as some arbitrageurs are present, this conclusion applies even in an economy in which all cross-sectional variation in expected returns is caused by sentiment. Sentiment investor demand results in substantial mispricing only if arbitrageurs are exposed to factor risk when taking the other side of these trades.

Reduced-form factor models are ubiquitous in empirical asset pricing. In these models, the stochastic discount factor (SDF) is represented as a function of a small number of portfolio returns. In equity market research, models such as the three-factor SDF of Fama and French (1993) and various extensions are popular with academics and practitioners alike. These models are reduced-form because they are not derived from assumptions about investor beliefs, preferences, and technology that prescribe which factors should appear in the SDF. Which interpretation should one give such a reduced-form factor model if it works well empirically?

That there exists a factor representation of the SDF is almost a tautology.[1] The economic content of the factor-model evidence lies in the fact that covariances with the factors not only explain the cross-section of expected returns, but that the factors also account for a substantial share of co-movement of stock returns. As a consequence, an investor who wants to benefit from the expected return spread between, say, value and growth stocks or recent winner and loser stocks, must invariably take on substantial factor risk exposure.

Researchers often interpret the evidence that expected return spreads are associated with exposures to volatile common factors as a distinct feature of "rational" models of asset pricing as opposed to "behavioral" models. For example, Cochrane (2011) writes:

> *Behavioral ideas—narrow framing, salience of recent experience, and so forth—*
> *are good at generating anomalous prices and mean returns in individual assets*
> *or small groups. They do not [...] naturally generate covariance. For example,*

---

[1] If the law of one price holds, one can always construct a single-factor or multi-factor representation of the SDF in which the factors are linear combination of asset payoffs (Hansen and Jagannathan 1991). Thus, the mere fact that a low-dimensional factor model "works" has no economic content beyond the law of one price.

"extrapolation" generates the slight autocorrelation in returns that lies behind momentum. But why should all the momentum stocks then rise and fall together the next month, just as if they are exposed to a pervasive, systematic risk?

In a similar vein, Daniel and Titman (1997) and Brennan, Chordia, and Subrahmanyam (1998) suggest that one can test for the relevance of "behavioral" effects on asset prices by looking for a component of expected return variation associated with stock characteristics (such as value/growth, momentum, etc.) that is orthogonal to factor covariances. This view that behavioral effects on asset prices are distinct from and orthogonal to common factor covariances is pervasive in the literature.[2]

Contrary to this standard interpretation, we argue that there is no such clear distinction between factor pricing and "behavioral" asset pricing. If sentiment—which we use as catch-all term for distorted beliefs, liquidity demands, or other distortions—affects asset prices, the resulting expected return spreads between assets should be explained by common factor covariances in similar ways as in standard rational expectations asset pricing models. The reason is that the existence of a relatively small number of arbitrageurs should be sufficient

---

[2]For example, Brennan, Chordia, and Subrahmanyam (1998) describe the reduced-form factor model studies of Fama and French as follows: "... Fama and French (FF) (1992a, b, 1993b, 1996) have provided evidence for the continuing validity of the rational pricing paradigm." The standard interpretation of factor pricing as distinct from models of mispricing also appears in more recent work. Just to provide one example, Hou, Karolyi, and Kho (2011) write: "Some believe that the premiums associated with these characteristics represent compensation for pervasive extra-market risk factors, in the spirit of a multifactor version of Merton's (1973) Intertemporal Capital Asset Pricing Model (ICAPM) or Ross's (1976) Arbitrage Pricing Theory (APT) (Fama and French 1993, 1996; Davis, Fama, and French 2000), whereas others attribute them to inefficiencies in the way markets incorporate information into prices (Lakonishok, Shleifer, and Vishny 1994; Daniel and Titman 1997; Daniel, Titman, and Wei 2001)."

to ensure that near-arbitrage opportunities—that is, trading strategies that earn extremely high Sharpe Ratios (SR)—do not exist. To take up Cochrane's example, if stocks with momentum did not rise and fall together next month to a considerable extent, the expected return spread between winner and loser stocks would not exist in the first place, because arbitrageurs would have picked this low-hanging fruit. Arbitrageurs neutralize components of sentiment-driven asset demand that are not aligned with common factor covariances, but they are reluctant to aggressively trade against components that would expose them to factor risk. Only in the latter case can the sentiment-driven demand have a substantial impact on expected returns. These conclusions apply not only to equity factor models that we focus on here, but also to no-arbitrage bond pricing models and currency factor models.

We start by analyzing the implications of absence of near-arbitrage opportunities for the reduced-form factor structure of the SDF. For typical sets of assets and portfolios, the covariance matrix of returns is dominated by a small number of factors. These empirical facts combined with absence of near-arbitrage opportunities imply that the SDF can be represented to a good approximation as a function of these few dominant factors.[3] This conclusion applies to models with sentiment-driven investors, too, as long as arbitrageurs eliminate the most extreme forms of mispricing.

If this reasoning is correct, then it should be possible to obtain a low-dimensional factor representation of the SDF purely based on information from the covariance matrix of

---

[3]This notion of absence of near-arbitrage is closely related to the interpretation of the Arbitrage Pricing Theory (APT) in Ross (1976): when discussing the empirical implementation of the APT in a finite-asset economy, Ross (p. 354) suggests bounding the maximum squared SR of any arbitrage portfolio at twice the squared SR of the market portfolio. However, our interpretation of APT-type models differs from some of the literature. For example, Fama and French (1996) (p. 75) regard the APT as a "rational" pricing model. We disagree with this narrow interpretation. The APT is just a reduced-form factor model.

returns. We show that a factor model with a small number of principal-component (PC) factors does about as well as popular reduced-form factor models do in explaining the cross-section of expected returns on anomaly portfolios. Thus, there doesn't seem to be anything special about the construction of the reduced-form factors proposed in the literature. Purely statistical factors do just as well. For typical test asset portfolios, their return covariance structure essentially dictates that the first few PC factors must explain the cross-section of expected returns.[4] Otherwise near-arbitrage opportunities would exist.

Tests of characteristics vs. covariances, like those pioneered in Daniel and Titman (1997), look for variation in expected returns that is orthogonal to factor covariances. Ex-post and in-sample such orthogonal variation always exists, perhaps even with statistical significance according to conventional criteria. It is questionable, though, whether such near-arbitrage opportunities are really a robust and persistent feature of the cross-section of stock returns. To check this, we perform a pseudo out-of-sample exercise. Splitting the sample period into subsamples, we extract the PCs from the covariance matrix of returns in one subperiod and then use the portfolio weights implied by the first subsample PCs to construct factors out-of-sample in the second subsample. While factors beyond the first few PCs contribute substantially to the maximum SR in-sample, PCs beyond the first few no longer add to the SR out-of-sample. In-sample deviations from low-dimensional factor pricing do not appear to reliably persist out of sample.

It would be wrong, however, to jump from the evidence that expected returns line up with common factor covariances to the conclusion that the idea of sentiment-driven asset

---

[4]The number of factors depends heavily on the underlying space of test assets. For instance, for Fama-French 5×5 size-B/M sorted portfolios, there are three dominant factors. For payoff spaces with weaker factor structure, the number of dominant factors is higher.

prices can be rejected. To show this, we build a model of a multi-asset market in which fully rational risk averse investors (arbitrageurs) trade with investors whose asset demands are based on distorted beliefs about the true distribution of returns (sentiment investors). We make two plausible assumptions. First, the covariance matrix of asset cash flows features a few dominant factors that drive most of the stocks' covariances. Second, sentiment investors cannot take extreme positions that would require substantial leverage or extensive use of short-selling. In this model, all cross-sectional variation in expected returns is caused by distorted beliefs and yet a low-dimensional factor model explains the cross-section of expected returns. To the extent that sentiment investor demand is orthogonal to covariances with the dominant factors, arbitrageurs elastically accommodate this demand and take the other side with minimal price concessions. Only sentiment investor demand that is aligned with covariances with dominant factors affects prices because it is risky for arbitrageurs to take the other side. As a result, the SDF in this economy can be represented to a good approximation as a function of the first few PCs, even though all deviations of expected returns from the CAPM are caused by sentiment. Therefore, the fact that a low-dimensional factor model holds is consistent with "behavioral" explanations just as much as it is consistent with "rational" explanations.

This model makes clear that empirical horse races between covariances with reduced-form factors and stock characteristics that are meant to proxy for mispricing or sentiment investor demand (as, e.g, in Daniel and Titman 1997; Brennan, Chordia, and Subrahmanyam 1998; Davis, Fama, and French 2000; and Daniel, Titman, and Wei 2001) set the bar too high for "behavioral" models: even in a world in which belief distortions affect asset prices, expected returns should line up with common factor covariances. Tests of factor models with ad-hoc macroeconomic factors (as, e.g., in Chen, Roll, and Ross 1986; Cochrane 1996; Li, Vassalou,

6

and Xing 2006; Liu and Zhang 2008) are not more informative either. As shown in Nawalkha 1997 (see, also, Shanken 1992; Reisman (1992); and Lewellen, Nagel, and Shanken 2010), if K dominant factors drive return variation and the SDF can be represented as a linear combination of these K factors, then the SDF can be represented, equivalently, by a linear combination of any K macroeconomic variables with possibly very weak correlation with the K factors.

Relatedly, theoretical models that derive relationships between firm characteristics and expected returns, taking as given an arbitrary SDF, do not shed light on the rationality of investor beliefs. Models such as Berk, Green, and Naik (1999), Johnson (2002), Liu, Whited, and Zhang (2009) or Liu and Zhang (2014), apply equally in our sentiment-investor economy as they apply to an economy in which the representative investor has rational expectations. These models show how firm investment decisions are aligned with expected returns in equilibrium, according to *firms'* first-order conditions. But these models do not speak to the question under which types of beliefs—rational or otherwise—*investors* align their marginal utilities with asset returns through their first-order conditions.

The observational equivalence between "behavioral" and "rational" asset pricing with regards to factor pricing also applies, albeit to a lesser degree, to partial equilibrium intertemporal capital asset pricing models (ICAPM) in the tradition of Merton (1973). In the ICAPM, the SDF is derived from the first-order condition of an investor who holds the market portfolio and faces exogenous time-varying investment opportunities. This leaves open the question how to endogenously generate the time-variation in investment opportunities in a way that is internally consistent with the investor's choice to hold the market portfolio. We show that time-varying investor sentiment is one possibility. If sentiment investor asset demands in excess of market portfolio weights have a single-factor structure and are mean-

reverting around zero, then the arbitrageurs' first-order condition implies an ICAPM that resembles the one in Campbell (1993) and Campbell and Vuolteenaho (2004a) in which arbitrageurs demand risk compensation both for cash-flow beta ("bad beta") and discount-rate beta ("good beta") exposure.

How can we differentiate between "rational" and "behavioral" explanations for returns? We argue that the only way to answer this question is to develop and test structural models with explicit assumptions about beliefs and preferences. Such models deliver testable predictions about the factors that should be in the SDF and the probability distributions under which a model-implied SDF prices assets. Of course, a test of a specific model of preferences and beliefs cannot yield generic conclusions about the validity of "rational" or "behavioral" approaches to asset pricing, but an empirical examination of reduced-form factor models does not help to circumvent this fundamental problem.

On the theoretical side, our work is related to Daniel, Hirshleifer, and Subrahmanyam (2001). Their model, too, includes sentiment-driven investors trading against arbitrageurs. In contrast to our model, however, the sentiment investors' position size is not constrained. As a consequence, for idiosyncratic belief distortions both the sentiment traders (mistakenly) and arbitrageurs (correctly) perceive a near-arbitrage opportunity and take huge offsetting bets against each other. With such unbounded position sizes, even idiosyncratic belief distortions can have substantial effects on prices and dominant factor covariances do not fully explain the cross-section of expected returns. We deviate from their setup because it seems plausible that sentiment investor position sizes and leverage are bounded.

On the empirical side, our paper is related to Stambaugh and Yuan (2015). They construct "mispricing factors" to explain a large number of anomalies. Our model of sentiment-driven asset prices explains why such "mispricing factors" work in explaining the cross-section

8

of expected returns. Empirically, our factor construction based on principal components is different, as the construction uses only the covariance matrix of returns and not the stock characteristics or expected returns. Kogan and Tian (2015) conduct a factor-mining exercise based on factors constructed by sorting on characteristics. They find that such factors are not robust in explaining the cross-section of expected returns out-of-sample. While we find a similar non-robustness for higher-order PC factors, we do find that the first few PC factors are robustly related to the cross-section of expected returns out-of-sample.

The rest of the paper is organized as follows. In Section I we describe the portfolio returns data that we use in this study. In Section II we lay out the implications of absence of near-arbitrage opportunities and we report the empirical results on factor pricing with principal component factors. Section III presents the model in which fully rational risk averse arbitrageurs trade with sentiment investors. Section IV develops a model with time-varying investor sentiment, which results in an ICAPM-type hedging demand.

I.   Portfolio Returns

To analyze the role of factor models empirically, we use two sets of portfolio returns. First, we use a set of 15 anomaly long-short strategies defined as in Novy-Marx and Velikov (2016) and the underlying 30 portfolios from the long and short sides of these strategies. This set of returns captures many of the most prominent features of the cross-section of stock returns discovered over the past few decades. We recreate the anomaly portfolios and use daily returns on these strategies for all estimation. Second, for comparison, we also use the 5×5 Size (SZ) and Book-to-Market (BM) sorted portfolios of Fama and French (1993).[5]

Table I provides some descriptive statistics for the anomaly long-short portfolios. Mean returns on long-short strategies range from 0.69% to 16% per year. Annualized squared SRs, shown in the second column, range from 0.01 to 0.83. Since these long-short strategies have low correlation with the market factor, these squared SRs are roughly equal to the incremental squared SR that the strategy would contribute if added to the market portfolio.

The factor structure of returns plays an important role in our subsequent analysis. To prepare the stage, we analyze the commonality in these anomaly strategy returns. We perform an eigenvalue decomposition of the covariance matrix of the 30 underlying portfolio returns and extract the principal components (PCs), ordered from the one with the highest

---

[5]We thank Ken French for making these returns available on his website. Monthly anomaly returns are available on Robert Novy-Marx's website. For the anomaly strategies in Novy-Marx and Velikov (2016), we use those strategies which can be constructed since 1963, are not classified as high turnover strategies, and are not largely redundant. Based on this latter exclusion criterion we eliminate the monthly-imbalanced net issuance (and use only the annually imbalanced one). We also as exclude the gross margins and asset turnover strategies which are subsumed, in terms of their ability to generate variation in expected returns, by the gross profitability strategy, as shown in Novy Marx (2013).

Table I
**Anomalies: Returns and Principal Component Factors**

The sample period is July 1966 to December 2015. The anomaly long-short strategy daily returns are as defined in Novy-Marx and Velikov (2016). Average returns and squared Sharpe ratios are reported in annualized terms. Mean returns and squared Sharpe ratios are calculated for 15 long-short anomaly strategies. Principal component factors are extracted from returns on the 30 portfolios underlying the long and short sides of these strategies.

	Mean Return	Squared SR	PC1	PC1-2	PC1-3	PC1-4	PC1-5
			PC factor-model R^2				
Size	1.29	0.01	0.08	0.11	0.60	0.64	0.69
Gross Profitability	4.56	0.17	0.03	0.05	0.13	0.16	0.50
Value	5.71	0.19	0.02	0.02	0.48	0.67	0.67
ValProf	8.45	0.63	0.08	0.10	0.34	0.38	0.46
Accruals	4.63	0.20	0.00	0.00	0.00	0.01	0.01
Net Issuance (rebal.-A)	9.06	0.83	0.15	0.26	0.27	0.38	0.40
Asset Growth	4.31	0.14	0.07	0.09	0.22	0.44	0.46
Investment	5.41	0.24	0.06	0.07	0.13	0.18	0.20
Piotroski's F-score	0.69	0.02	0.02	0.07	0.15	0.15	0.16
ValMomProf	9.59	0.43	0.01	0.44	0.63	0.70	0.80
ValMom	5.71	0.15	0.03	0.35	0.73	0.73	0.73
Idiosyncratic Volatility	9.25	0.20	0.34	0.55	0.69	0.92	0.94
Momentum	16.31	0.61	0.01	0.72	0.72	0.91	0.92
Long Run Reversals	4.22	0.09	0.01	0.01	0.40	0.52	0.58
Beta Arbitrage	7.12	0.24	0.14	0.33	0.33	0.46	0.75

eigenvalue (which explains most of the co-movement of returns) to the one with the lowest. We then run a time-series regression of each long-short strategy return on the first, the first and the second, ... , up to a regression on the PCs one to five. The last five columns in Table I report the R^2 from these regressions. Since we are looking at long-short portfolio returns here that are roughly market-neutral, the first PC naturally does not explain much of the time-series variation of returns. With the first and second PC combined, the explanatory power in terms of R^2 ranges from 0 for the Accruals strategy to 0.72 for the Momentum strategy. Once the first five PCs are included in the regression, the explanatory power is more uniform (except for Accruals), with R^2 ranging from 0.16 for the F-score strategy to 0.94 for the Idiosyncratic Volatility strategy, with most strategies having R^2 above 0.5. Thus, a substantial portion of the time-series variation in returns of these anomaly portfolios can be traced to a few common factors.

For the second set of returns from the size-B/M portfolios, it is well known from Fama and French (1993) that three factors – the excess return on the value-weighted market index (MKT), a small minus large stock factor (SMB), and a high minus low B/M factor (HML) – explain more than 90% of the time-series variation of returns. While Fama and French construct SMB and HML in a rather special way from a smaller set of six size-B/M portfolios, one obtains essentially similar factors from the first three PCs of the 5×5 size-B/M portfolio returns.

The first PC is, to a good approximation, a level factor that puts equal weight on all 25 portfolios. The first two of the remaining PCs after removing the level factor are, essentially, the SMB and HML factors. Figure 1 plots the eigenvectors. PC1, shown on the left, has positive weights on small stocks and negative weights on large stocks, that is, it is similar to SMB. PC2, shown on the right, has positive weights on high B/M stocks and negative
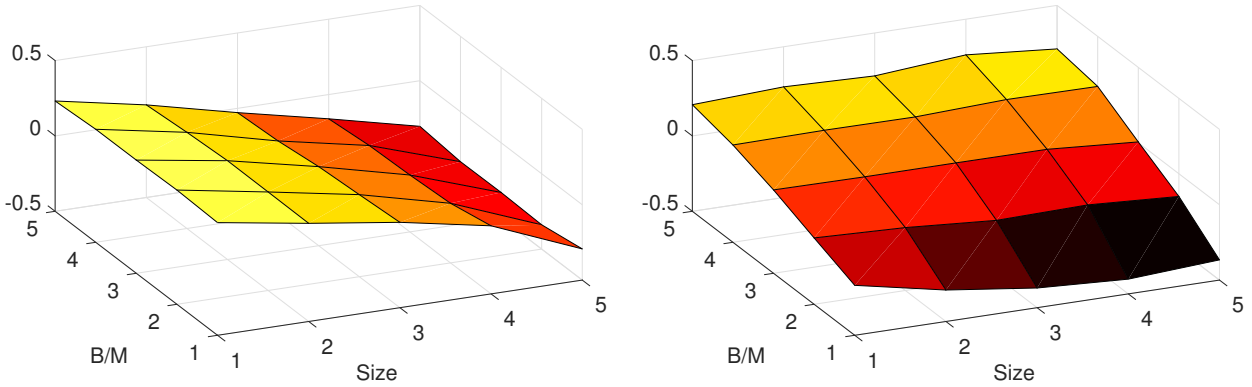
Figure 1. Eigenvectors. The figures plot eigenvectors weights corresponding to the second (left) and third (right) principal components of Fama-French 25 SZ/BM portfolio returns.

weights on low B/M stocks, that is, it is similar to HML. This shows that the Fama-French factors are not special in any way; they simply succinctly summarize cross-sectional variation in the size-B/M portfolio returns, similar to the first three PCs.[6]

---

[6]A related observation appears in Lewellen, Nagel, and Shanken (2010). Lewellen et al. note that three factors formed as linear combinations of the 25 SZ/BM portfolio returns with random weights often explain the cross-section of expected returns on these portfolios about as well as the Fama-French factors do.

II. Factor pricing and absence of near-arbitrage

We start by showing that if we have assets with a few dominating factors that drive much of the covariances of returns (i.e., small number of factors with large eigenvalues), then those factors must explain asset returns. Otherwise near-arbitrage opportunities would arise, which would be implausible even if one entertains the possibility that prices could be influenced substantially by the subjective beliefs of sentiment investors.

Consider an economy with discrete time $t = 0, 1, 2, \ldots\ldots$ There are N assets in the economy indexed by $i = 1, \ldots, N$ with a vector of returns in excess of the risk-free rate, R. Let $\mu \equiv E[R]$ and denote the covariance matrix of excess returns with Γ.

Assume that the Law of One Price (LOP) holds. The LOP is equivalent to the existence of an SDF M such that $\mathrm{E}[MR] = 0$. Note that $\mathrm{E}[\cdot]$ represents objective expectations of the econometrician, but there is no presumption here that $\mathrm{E}[\cdot]$ also represents subjective expectations of investors. Thus, the LOP does not embody an assumption about beliefs, and hence about the rationality of investors (apart from ruling out beliefs that violate the LOP).

Now consider the minimum-variance SDF in the span of excess returns, constructed as in Hansen and Jagannathan (1991) as

$$M = 1 - \mu'\Gamma^{-1}(R - \mu). \tag{1}$$

Since we work with excess returns, the SDF can be scaled by an arbitrary constant, and we normalize it to have $\mathrm{E}[M] = 1$. The variance of the SDF,

$$\mathrm{Var}\,(M) = \mu'\Gamma^{-1}\mu, \tag{2}$$

equals the maximum squared Sharpe ratio (SR) achievable from the N assets.

Now define absence of near-arbitrage as the absence of extremely high-SR opportunities (under objective probabilities) as in Cochrane and Saá-Requejo (2000). Ross (1976) also proposed a bound on the squared SR for an empirical implementation of his Arbitrage Pricing Theory in a finite-asset economy. He suggested ruling out squared SR greater than $2\times$ the squared SR of the market portfolio. Such a bound on the maximum squared SR is equivalent, via (2), to an upper bound on the variance of the SDF M that resides in the span of excess returns.

Our perspective on this issue is different than in some of the extant literature. For example, MacKinlay (1995) suggests that the SR should be (asymptotically) bounded under "risk-based" theories of the cross-section of stock returns, but stay unbounded under alternative hypotheses that include "market irrationality." A similar logic underlies the characteristics vs. covariances tests in Daniel and Titman (1997) and Brennan, Chordia, and Subrahmanyam (1998). However, ruling out extremely high-SR opportunities implies only weak restrictions on investor beliefs and preferences, with plenty of room for "irrationality" to affect asset prices. Even in a world in which many investors' beliefs deviate from rational expectations, near-arbitrage opportunities should not exist as long as some investors ("arbitrageurs") with sufficient risk-bearing capacity have beliefs that are close to objective beliefs. We can then think of the pricing equation $\mathrm{E}[MR] = 0$ as the first-order condition of the arbitrageurs' optimization problem and hence of the SDF as representing the marginal utility of the arbitrageur. The model in Section III shows that extremely high volatility of M can occur only if the wealth of arbitrageurs in the economy is small and the sentiment investors they are trading against take huge concentrated bets on certain types of risk. Given sufficient arbitrageur wealth and reasonable constraints on sentiment investors' trading, the

variance of M is bounded from above.

We now show that the absence of near-arbitrage opportunities implies that one can represent the SDF as a function of the dominant factors driving return variation. Consider the eigen-decomposition of the excess returns covariance matrix

$$\Gamma = Q\Lambda Q' \qquad \text{with} \qquad Q = (q_1, ..., q_N) \tag{3}$$

and λ_i as the diagonal elements of Λ. Assume that the first principal component (PC) is a level factor, that is, $q_1 = \frac{1}{\sqrt{N}}\iota$, where ι is a conformable vector of ones. This implies $q_k'\iota = 0$ for $k > 1$, that is, the remaining PCs are long-short portfolios. In the Appendix, Section A we show that

$$
\begin{aligned}
\text{Var}\,(M) &= (\mu'q_1)^2\lambda_1^{-1} + \mu'Q_z\Lambda_z^{-1}Q_z'\mu \\
&= \frac{\mu_m^2}{\sigma_m^2} + N\overline{\text{Var}}(\mu_i)\sum_{k=2}^{N}\frac{\overline{Corr}(\mu_i, q_{ki})^2}{\lambda_k},
\end{aligned}
\tag{4}
$$

where the z subscripts stand for matrices with the first PC removed and $\mu_m = \frac{1}{\sqrt{N}}q_1'\mu$, $\sigma_m^2 = \frac{\lambda_1}{N}$, while $\overline{\text{Var}}(.)$ and $\overline{Corr}(.)$ denote cross-sectional variance and correlation. According to this expression for SDF variance, a strong factor structure in test asset returns combined with significant cross-sectional variation in average returns implies that expected returns must line up with the first few (high-eigenvalue) PCs; otherwise $\text{Var}(M)$ would be very high and near-arbitrage opportunities would exist. To see this, note that the sum of the squared correlations of μ_i and q_{ki} is always equal to one. But the magnitude of the sum weighted by the inverse λ_k depends on which of the PCs the vector μ lines up with. If it lines up with high λ_k PCs then the sum is much lower than if it lines up with low λ_k PCs. For typical test

16

assets, eigenvalues decay rapidly beyond the first few PCs. In this case, a high correlation of μ_i with a low-eigenvalue q_{ki} would lead to an enormous maximum Sharpe Ratio.

This argument relies only on the absence of near-arbitrage opportunities; therefore, the result obtains for both rational and "behavioral" classes of models. The logic applies to any set of test assets, as long as this set has a strong factor structure. We use the anomaly portfolios in our analysis because these portfolios are well-studied in the recent literature and do exhibit these properties. If, on the other hand, a set of test assets exhibits only weak factor structure, Eq. (4) shows that these assets can't have much cross-sectional dispersion in expected returns without violating reasonable bounds on the maximum variance of the SDF (or the Sharpe ratio).

We have maintained so far that expected returns must line up with the first few principal components, otherwise high-SR opportunities would arise. We now provide empirical support for this assertion. We quantify these relationships using our two sets of test assets by asking, counterfactually, what the maximum SR of the test assets would be if expected returns did not line up, as they do in the data, with the first few (high-eigenvalue) PCs, but were instead also correlated with the higher-order PCs. To do this, we go back to Eq. (4). We assume that μ_i is correlated with the first K PCs, while the correlation with the remaining PCs is exactly zero. For simplicity of exposition, we further assume that all non-zero correlations are equal.[7] We set $\overline{\text{Var}}\,(\mu_i)$ and Λ equal to their sample values.

Figure 2 presents the results. Panel (a) shows the counterfactual squared SR for the 30 anomaly portfolios. If expected returns of these portfolios lined up equally with the first two PCs (excl. level factor) but not the higher-order ones, the squared SR would be around 2. The squared SR of the first three PCs is plotted as the dashed line in the figure for

---

[7]Since the sum of all squared correlations must add up to one, each squared correlation is $1/K$.

(a) 30 anomaly portfolios (in excess of level factor)   (b) 5×5 Size-B/M portfolios (in excess of level factor)
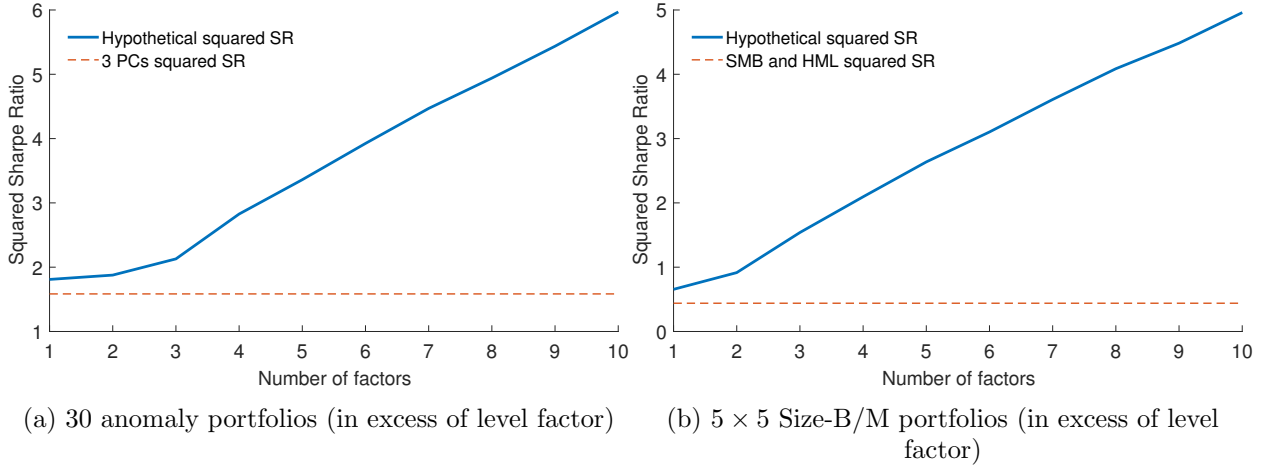
Figure 2. Hypothetical Sharpe ratios. The figures show hypothetical Sharpe Ratios if expected returns line up with first K (high-eigenvalue; excl. PC1) principal components. Panel (a) uses 30 anomaly portfolios as test assets. Panel (b) focuses on 25 Size-B/M portfolios.

comparison. If expected returns lined up instead equally with the first 10 PCs, the squared SR would be almost 6. Panel (b) shows a similar analysis for the 5×5 size-B/M portfolios. Here, too, the counterfactual squared SR increases rapidly with K. If expected returns lined up equally with (only) the first two PCs (excl. level factor), the squared SR would be higher than the sum of the squared SRs of SMB and HML (SMB and HML do not price the cross-section of the 5×5 size-B/M portfolios perfectly). However, if expected returns were correlated equally with the first 10 PCs, the squared SR would reach around 5.

We can look at the data from another perspective. Fixing $\mathrm{Var}\,(M)$ and $\overline{\mathrm{Var}}\,(\mu_i)$, we can bound the minimum cross-sectional R^2 of a factor model based on K PCs, $\sum_2^K \overline{Corr}\,(\mu_i, q_{ki})^2$. Figure 3 presents the results. The plots show that reasonable bounds on the maximum squared Sharpe ratio imply that a low dimensional factor model constructed from high eigenvalue PCs will provide a good approximation to expected returns. Looking at Panel (a) (30 anomaly portfolios), a bound of $SR^2 < 2$ implies a three factor model will deliver greater

(a) 30 anomaly portfolios (in excess of level factor) (b) 5×5 Size-B/M portfolios (in excess of level factor)
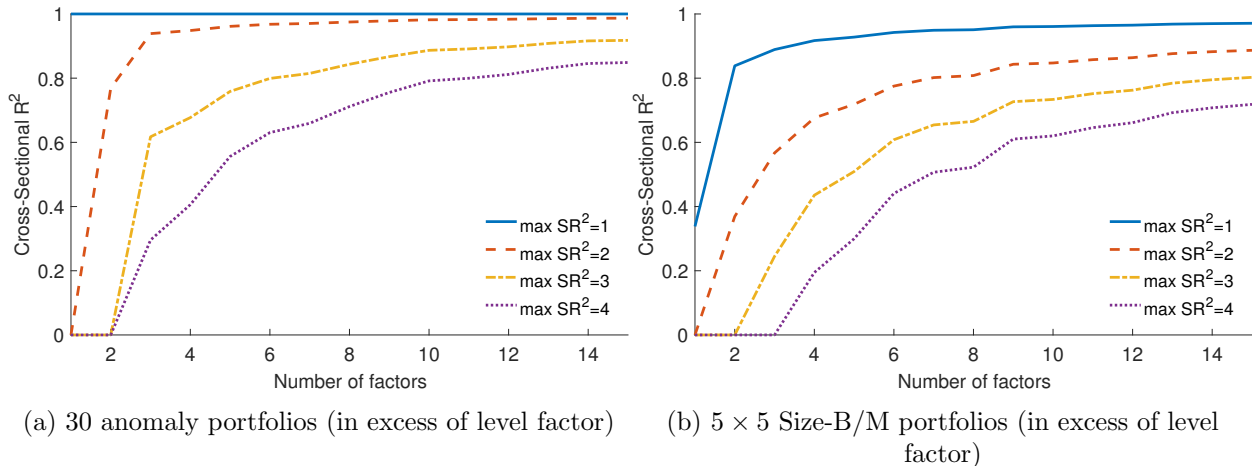
Figure 3. Minimum cross-sectional R^2. The figures plot minimum cross-sectional R^2 from a K factor model (high-eigenvalue; excl. PC1) for various bounds on $\mathrm{Var}\,(M)$. Panel (a) uses 30 anomaly portfolios as test assets. Panel (b) focuses on 25 Size-B/M portfolios.

than 90% cross-sectional R^2. The strength of this argument, of course, heavily depends on the underlying space of test assets. For payoff spaces with weaker factor structure or lower variation in expected returns, number of factors in the SDF (the value of K) could be higher for any given bound on $\mathrm{Var}\,(M)$. This is seen in panel (b), which presents the results for Fama French 25 BM/SZ portfolios. Since this set of assets has significantly lower $\overline{\mathrm{Var}}\,(\mu_i)$ compared to the anomaly portfolios, but similar factor structure (eigenvalues of covariance matrix), the bounds are looser.

A.   *Principal components as reduced-form factors: pricing performance*

Based on the no-near-arbitrage logic developed above, it should not require a judicious construction of factor portfolios to find a reduced-form SDF representation. Brute statistical force should do. We already showed earlier in Figure 1 that the first three principal components of the 5×5 size-B/M portfolios are similar to the three Fama-French factors. We now investigate the *pricing* performance of principal component factor models. Our focus,

therefore, is on how well covariances with factors explain expected returns rather than how well factors explain variances. This is an important distinction. It is perfectly possible for a factor to be important in explaining return variance but to play no role in pricing.

Table II shows that the first few PCs do a good job of capturing cross-sectional variation in expected returns of the anomaly portfolios. We run time-series regressions of the 15 long-short anomaly excess returns on the principal component factors extracted from 30 underlying portfolio returns. The upper panel in Table II reports the pricing errors, that is, the intercepts or alphas, from these regressions. The raw mean excess return (in % per year) is shown in the first column, alphas for specifications with an increasing number of PC factors in the second to sixth column. With just the first PC (PC1; roughly the market) as a single factor, the SDF does not fit well. Alphas reach magnitudes up to 17% per year. Adding PC2 and PC3 to the factor model drastically shrinks the pricing errors. With five factors, the maximum (absolute) alpha is 6.

The bottom panel reports the (ex post) maximum squared SR of the anomaly portfolios (4.23) and the maximum squared SR of the PC factors. With five factors, the highest-SR combination of the factors achieves a squared SR of 1.77. This is still considerably below the maximum squared SR of the anomaly portfolios and the p-values from a χ^2-test of the zero-pricing error null hypothesis rejects at a high level of confidence. However, it is important to realize that this pricing performance of the PC1-5 factor model is actually better than the performance of the Fama-French factor model in pricing the 5×5 size-B/M portfolios—which is typically regarded as a successful factor model. As the table shows, the maximum squared SR of the 5×5 size-B/M portfolios is 3.82. But the squared SR of MKT, SMB, and HML is only 0.81. As the table shows, PC1-3, a combination of the first three PCs of the size-B/M portfolios (incl. level factor), has a squared SR of 0.88 and gets slightly closer

Table II
Explaining Anomalies with Principal Component Factors

The sample period is July 1966 to December 2015. The anomaly long-short strategy daily returns are as defined in Novy-Marx and Velikov (2016). Average returns and factor-model alphas are reported in % per month. Squared Sharpe Ratios are reported in annualized terms. Mean returns and alphas are calculated for 15 long-short anomaly strategies. Maximum squared Sharpe ratios and principal component factors are extracted from returns on the 30 portfolios underlying the long and short sides of these strategies.

		PC1	PC1-2	PC1-3	PC1-4	PC1-5
	Mean Return	PC factor-model alphas				
Size	1.29	2.47	4.89	2.07	3.29	4.08
Gross Profitability	4.56	3.96	2.33	3.37	4.33	2.48
Value	5.71	6.28	6.63	3.80	0.97	1.16
ValProf	8.45	9.45	8.02	6.34	5.33	4.43
Accruals	4.63	4.57	3.94	3.98	4.25	4.39
Net Issuance (rebal.-A)	9.06	10.30	6.55	6.87	5.31	4.87
Asset Growth	4.31	5.31	3.42	2.07	-0.64	-0.27
Investment	5.41	6.31	5.25	4.38	3.16	3.57
Piotroski's F-score	0.69	0.94	-0.32	0.17	0.21	0.12
ValMomProf	9.59	10.17	-0.39	-2.44	-0.59	-1.95
ValMom	5.71	6.51	-2.95	-5.86	-5.66	-5.83
Idiosyncratic Volatility	9.25	13.13	2.52	5.00	0.08	-0.74
Momentum	16.31	17.12	-2.41	-2.34	2.12	2.57
Long Run Reversals	4.22	4.76	4.69	1.92	-0.53	-1.48
Beta Arbitrage	7.12	8.87	1.85	1.57	-1.00	1.23
	Max. sq. SR	PC factors' max. squared SR				
All anomalies	4.23	0.11	1.34	1.45	1.69	1.77
χ^2-pval. for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
For comparison:						
25 SZ/BM	3.82	0.27	0.27	0.88	1.19	1.23
χ^2-pval. for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MKT, SMB, and HML	0.81	-	-	-	-	-

to the mean-variance frontier than the Fama-French factors. While the PC factor models and the Fama-French factor model are statistically rejected at a high level of confidence, the fact that the Fama-French model is typically viewed as successful in explaining the size-B/M portfolio returns suggests that one should also view the PC1-3 factor model as successful. In terms of the distance to the mean-variance frontier, the PC1-5 factor model for the anomalies in the upper panel is even better at explaining the cross-section of anomaly returns than the Fama-French model in explaining the size-B/M portfolio returns.

Overall, this analysis shows that one can construct reduced-form factor models simply from the principal components of the return covariance matrix. There is nothing special, for example, about the construction of the Fama-French factors. Intended or not, the Fama-French factors are similar to the first three PCs of the size-B/M portfolios and they perform similarly well in explaining the cross-section of average returns of those portfolios.

While reduced-form factor model tests cannot help to differentiate between competing models of investor beliefs, summarizing the cross-section of expected returns in a parsimonious fashion with a factor-SDF can still be useful to quantify the extent of the challenge faced by asset pricing models. If the dominant factors earn substantial Sharpe Ratios, then a rational expectations model would require large technological shocks (as in Bansal and Yaron 2004) or high effective risk aversion (e.g., due to habits as in Campbell and Cochrane 1999). A "behavioral" model such as the one in Section 4 would require substantial belief distortions and a significant wealth share of sentiment investors.

B. Characteristics vs. covariances: In-sample and out-of-sample

Daniel and Titman (1997) and Brennan, Chordia, and Subrahmanyam (1998) propose tests that look for expected return variation that is correlated with firm *characteristics* (e.g.,

B/M), but not with reduced-form factor model *covariances*. Framed in reference to our analysis above, this would mean looking for cross-sectional variation in expected returns that is orthogonal to the first few PCs—which implies that it must be variation that lines up with some of the higher-order PCs. The underlying presumption behind these tests is that "irrational" pricing effects should manifest themselves as mispricings that are orthogonal to covariances with the first few PCs.

From the evidence in Table II that the ex-post squared SR obtainable from the first few PCs falls short, by a substantial margin, of the ex-post squared SR of the test assets, one might be tempted to conclude that (i) there is actually convincing evidence for mispricing orthogonal to factor covariances, and (ii) that therefore the approach of looking for mispricing unrelated to factor covariances is a useful way to test behavioral asset pricing models. After all, at least ex-post, average returns appear to line up with components of characteristics that are orthogonal to factor covariances.

We think that this conclusion would not be warranted. First, there is certainly substantial sampling error in the ex-post squared SR. Of course, the χ^2-test in Table II takes the sampling error into account and still rejects the low-dimensional factor models.[8] However, there are additional reasons to suspect that high ex-post SR are not robust indicators of persistent near-arbitrage opportunities. Short-lived near-arbitrage opportunities might exist for a while before being recognized and eliminated by arbitrageurs. Data-snooping biases (Lo and MacKinlay 1990) further overstate the in-sample SR.

To shed light on this robustness issue, we perform a pseudo-out-of-sample (OOS) analysis. We split our sample period in two halves, treating the first half as our in-sample period, and

---

[8]Bootstrap simulation which accounts for the significant non-normality of daily returns leads to similar rejection of the factor models, albeit with higher p-values.
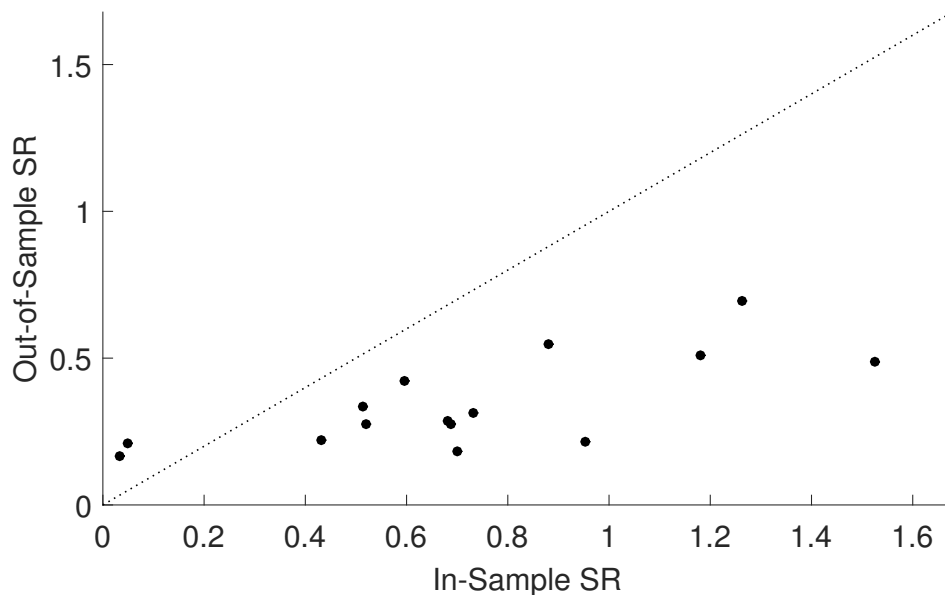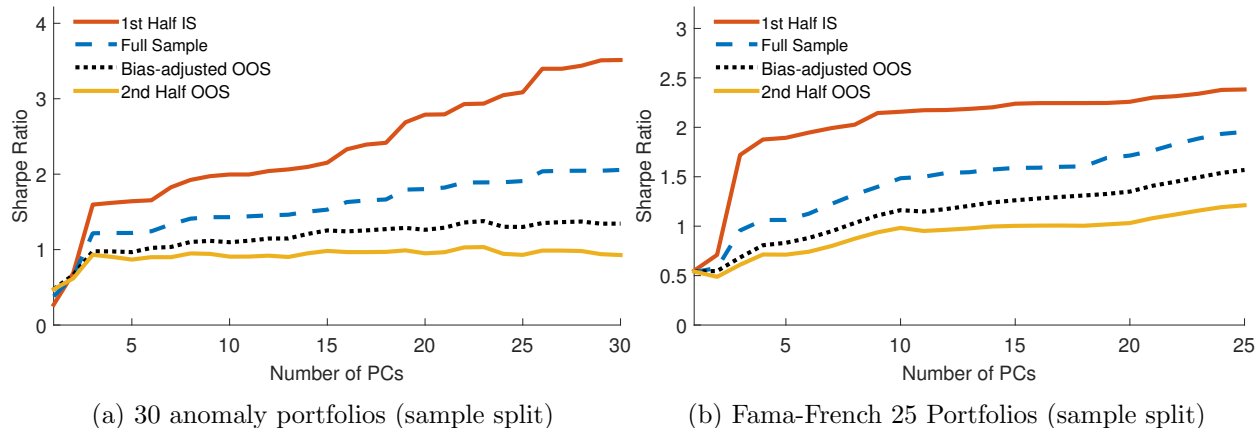
Figure 4. In-sample and out-of-sample Sharpe ratios. The figure shows in-sample and out-of-sample Sharpe ratios of 15 anomaly long-short strategies. The sample period is split into two halves. In-sample SR are those in the first subperiod. Out-of-sample SR are those in the second sub-period. Sharpe ratios are annualized.

the second half as our OOS period. We start with a univariate perspective with the 15 anomaly long-short portfolios. Figure 4 plots the in-sample SR in the first subperiod on the x-axis and the OOS SR on the y-axis. The figure shows that there is generally a substantial deterioration. OOS SR are, on average, less than half as large as the in-sample values and almost all of them are lower in the OOS period. Furthermore, the strategies that hold up best are those that have relatively low in-sample SR. This is one first indication that high in-sample SR do not readily lead to high OOS SR.

Our pseudo-OOS evidence is consistent with recent work by McLean and Pontiff (2016) that examines the *true* OOS performance of a large number of cross-sectional return predictors that appeared in the academic literature in recent decades. They find a substantial

(a) 30 anomaly portfolios (sample split)   (b) Fama-French 25 Portfolios (sample split)

Figure 5. IS and OOS maximum Sharpe Ratios. The figures plot in-sample and out-of-sample maximum Sharpe Ratios (annualized) of first K principal components (incl. level factor) of 30 anomaly long and short portfolio returns (Panel a) and 20 size-B/M portfolios (Panel b). We split the sample period into two halves then extract PCs in the first sub-period and calculate the SR-maximizing combination of the first K PCs using sample means and covariances. We then apply the portfolio weights implied by this combination in the out-of-sample period (second sub-period). Full sample Sharpe ratios (blue dashed) are shown for comparison. The black dotted line represents the mean-bias-adjusted OOS Sharpe ratio with bias calculated from a bootstrap simulation.

decay in returns from the researchers' in-sample period to the OOS period after the publication of the academic study. Most relevant for our purposes is their finding that the predictors with higher in-sample t-statistics are the ones that experience the biggest decay.[9]

In Figure 5, panel (a), we consider all 30 portfolios underlying the 15 long-short strategies jointly. Focusing first on the full sample, we look at the maximum SR that can be obtained from a combination of the first K principal components (incl. level factor). The dashed blue line shows a significant increase in Sharpe ratio beyond the first few PCs, graphically demonstrating the rejection of low-dimensional factor models (formally shown in Table II). This pattern may, however, be spurious. Data-snooping biases and transient near-arbitrage

---

[9]In private correspondence, Jeff Pontiff provided us with estimation results showing that a stronger decay is also present for predictors with high in-sample SR. We thank Jeff for sending us those results.

opportunities can inflate the full sample Sharpe ratios. Such biases should manifest as very high Sharpe ratios in the early part of the data followed by significantly lower Sharpe ratios later. We observe exactly this pattern when looking at the red (first half) and yellow (second half OOS) lines. In-sample (first half) with $K = 3$, the maximum SR is around 1.6, but raising K further raises the SR to nearly 4 for $K = 30$. However, out of sample, the picture looks different. For each K, we now take the asset weights that yield the maximum SR from the first K PCs in the first subperiod, and we apply these weights to returns from the second subperiod.[10] The solid yellow line in the figure shows the resulting OOS Sharpe ratio. Not surprisingly, overall SR are lower out of sample. Most importantly, it makes virtually no difference whether one picks $K = 3$ or $K = 30$ – the OOS SR is about the same and stays mostly around 1. Hence, while the higher-order PCs add substantially to the SR in sample, they provide no incremental improvement of the SR in the OOS period. Whatever these higher-order PCs were picking up in the in-sample period is not a robust feature of the cross-section of expected return that persists out of sample.

One potential concern with these analyses is that mean-variance optimal portfolios constructed using in-sample estimates of means and covariances naturally perform poorly out of sample. If so, the diminished SR contribution of low-eigenvalue PCs in the OOS analysis could be an artifact of an inefficient method of constructing portfolios. It is important to realize, however, that our PC-based method is already "robust" by design and thus avoids usual pitfalls of naïve mean-variance optimization. A major problem with mean-variance optimization is that eigenvalues of the sample covariance matrix are too "extreme" in the sense that large sample eigenvalues overestimate and small ones underestimate the true ei-

---

[10]Results are similar if we recalculate optimal weights each date using data available up to that point (both expanding and rolling windows).

genvalues. Using these incorrect eigenvalues results in portfolios with extreme weights, due to the inversion of the tiny eigenvalues. Ignoring the small eigenvalue PC portfolios is, hence, a form of regularization well suited to the OOS portfolio problem.[11]

To the extent there are any remaining concerns, we address them in three ways. First, our covariance estimates throughout the paper are calculated from daily returns. This largely eliminates any downward bias in OOS Sharpe ratios due to sampling error in covariances, which could be substantial if we used monthly data instead. Second, we solve for optimal portfolio weights using the realized second half covariance matrix and obtain similar OOS performance (not shown). Finally, we use a bootstrap simulation to quantify the remaining bias resulting from uncertainty in estimated means.[12] The black dotted line shows that the mean-bias-adjusted OOS Sharpe ratio is only moderately higher than the unadjusted value. Importantly, the conclusion that Sharpe ratios don't significantly increase beyond the first few PCs is unchanged. In panel (b) we repeat the same analysis for the 5×5 size-B/M portfolios and their PC factors. The results are similar.

In summary, the empirical evidence suggests that reduced-form factor models with a few principal component factors provide a good approximation of the SDF, as one would

---

[11]The closest counterpart to our procedure in the linear regression setting are the Principal Component Regression (PCR – a procedure that replaces explanatory variables in a regression with their first few principal components) and the ridge regression (a continuous version of PCR – a procedure that continuously shrinks small PCs of explanatory variables relatively more than high PCs). See Friedman et al. (2001) for more details.

[12]We randomly sample (without replacement) half of the returns to extract PCs and calculate the SR-maximizing combination of first K PCs in the subsample. We then apply the portfolio weights implied by this combination in the out-of-sample period (remainder of the data). The procedure is repeated 1,000 times. Bias is calculated as the mean of the bootstrap distribution minus the full sample value.

expect if near-arbitrage opportunities do not exist. However, as we discuss in the rest of the paper, this fact tells us little about the "rationality" of investors and the degree to which "behavioral" effects influence asset prices.

III.   Factor pricing in economies with sentiment investors

We now show that mere absence of near-arbitrage opportunities has limited economic content. We model a multi-asset market in which fully rational risk averse investors (arbitrageurs) trade with investors whose asset demands are driven by distorted beliefs (sentiment investors).

Consider an IID economy with discrete time $t = 0, 1, 2, ...$ There are N stocks in the economy indexed by $i = 1, ..., N$. The supply of each stock is normalized to $1/N$ shares. A risk-free bond is available in perfectly elastic supply at a gross interest rate of $R_F > 1$. Stock i earns time-t dividends D_{it} per share. Collect the individual-stock dividends in the column vector D_t. We assume that $D_t \sim \mathcal{N}(0, \Gamma)$.

We assume that the covariance matrix of asset cash flows Γ features a few dominant factors that drive most of the stocks' covariances. This assumption is consistent with empirical evidence in Ball, Sadka, and Sadka (2009) who show that there is strong factor structure in fundamentals. Since prices are constant in our IID case, the covariance matrix of returns equals the covariance matrix of dividends, Γ. Therefore, even with belief distortions, returns inherit a strong factor structure. Consider further the eigenvalue decomposition of the covariance matrix $\Gamma = Q \Lambda Q'$. Assume that the first PC is a level factor, with identical constant value for each element of the corresponding eigenvector $q_1 = \iota N^{-1/2}$. Then, the variance of returns on the market portfolio is

$$\sigma_m^2 = \mathrm{Var}(R_{m,t+1}) = N^{-2} \iota' q_1 q_1' \iota \lambda_1 = N^{-1} \lambda_1. \tag{5}$$

All other principal components, by construction, are long-short portfolios, that is, $\iota' q_k = 0$ for $k > 1$.

There are two groups of investors in this economy. The first group comprises competitive rational arbitrageurs in measure $1 - \theta$. The representative arbitrageur has CARA utility with absolute risk aversion a. In this IID economy, the optimal strategy for the arbitrageur is to maximize next period wealth, that is,

$$\max_y E\left[-\exp(-aW_{t+1})\right] \tag{6}$$

$$\text{s.t.} \quad W_{t+1} = (W_t - C_t)R_F + y'R_{t+1}, \tag{7}$$

where $R_{t+1} \equiv P_{t+1} + D_{t+1} - P_t R_F$ is a vector of dollar excess returns. From arbitrageurs' first-order condition and their budget constraint, we obtain their asset demand

$$y_t = \frac{1}{a}\Gamma^{-1}E[R_{t+1}] \tag{8}$$

where expectations are taken under the objective measure.

The second group comprises sentiment investors who have biased expectations; in making their decisions these investors use state probabilities which differ from the objective probabilities used by arbitrageurs. Sentiment investors are present in measure θ. Like arbitrageurs, they have CARA utility with absolute risk aversion a and they face a similar budget constraint, but they have an additional sentiment-driven component to their demand δ. Their risky asset demand vector is

$$x_t = \frac{1}{a}\Gamma^{-1}E[R_{t+1}] + \delta. \tag{9}$$

where we assume that $\delta'\iota = 0$. The first term is the rational component of the demand,

equivalent to the arbitrageur's demand. The second term is the sentiment investors' excess demand δ, which is driven by investors' behavioral biases or misperceptions of the true distribution of returns. This misperception is only cross-sectional; there is no misperception of the market portfolio return distribution since $\delta'\iota = 0$.

If δ were completely unrestricted, then prices could be arbitrarily strongly distorted even if arbitrageurs are present. Unbounded δ would imply that sentiment investors can take unbounded portfolio positions, including high levels of leverage and unbounded short sales. This is not plausible. Extensive short selling and high leverage is presumably more likely for arbitrageurs than for less sophisticated sentiment-driven investors. For this reason, we constrain the sentiment investors' "extra" demand due to the belief distortion to

$$\delta'\delta \leq 1. \tag{10}$$

This constraint is a key difference between our model and the models like Daniel, Hirshleifer, and Subrahmanyam (2001). In their model, no such constraint is imposed. As a consequence, when sentiment investors (wrongly) perceive a near-arbitrage opportunity, they are willing to take an extremely levered bet on this perceived opportunity. Arbitrageurs in turn are equally willing to take a bet in the opposite direction to exploit the actual near-arbitrage opportunity generated by the sentiment investor demand. Since sentiment investors are equally aggressive in pursuing their perceived opportunity as arbitrageurs are in pursuing theirs, mispricing can be big even for "idiosyncratic" mispricings. Imposing the constraint (10) prevents sentiment investors from taking such extreme positions, which is arguably realistic. By limiting the cross-sectional sum of squared deviations from rational weights in this way, the maximum deviation that we allow in an individual stock is, approximately, one

that results in a portfolio weight of ± 1 in one stock and $1/N \pm 1/N$ in all others.[13] Thus, the constraint still allows sentiment investors to have rather substantial portfolio tilts, but it prevents the most extreme ones.

Market clearing,

$$\theta\delta + \frac{1}{a}\Gamma^{-1}E[R_{t+1}] = \frac{1}{N}\iota, \tag{11}$$

implies

$$E[R_{t+1}] - \mu_m\iota = -a\theta\Gamma\delta, \tag{12}$$

where $\mu_m \equiv (1/N)\iota'E[R_{t+1}]$ and we used the fact that, due to the presence of the level factor, ι is an eigenvector of Γ and so $\Gamma^{-1}\iota = \frac{1}{\lambda_1}\iota = \frac{1}{N\sigma_m^2}\iota$. Moreover, we used $\mu_m = a\sigma_m^2$. Then, after substituting into arbitrageurs optimal demand, we get

$$y = \frac{1}{N}\iota - \theta\delta. \tag{13}$$

Consequently, we obtain the SDF,

$$M_{t+1} = 1 - a\left(R - E\left[R\right]\right)'y$$

$$= 1 - a[R_{m,t+1} - \mu_m] + a(R_{t+1} - E[R_{t+1}])'\theta\delta, \tag{14}$$

---

[13]In equilibrium, a representative rational investor with objective expectations would hold the market portfolio with weights $1/N$. Deviating to a weight of 1 in one stock and to zero in all the other $N-1$ stocks therefore implies a sum of squared deviations of $(1 - 1/N)^2 + (N-1)/N^2 = 1 - 1/N \approx 1$ and exactly zero mean deviation.

and the SDF variance,

$$\text{Var}(M) = a^2 \sigma_m^2 + a^2 \theta^2 \delta' \Gamma \delta. \tag{15}$$

The effect of δ on the factor structure and the volatility of the SDF depends on how δ lines up with the PCs. To characterize the correlation of δ with the PCs, we express δ as a linear combination of PCs,

$$\delta = Q\beta, \tag{16}$$

with $\beta_1 = 0$. Note that $\delta'\delta = \beta'Q'Q\beta = \beta'\beta$ so the constraint (10) can be expressed in terms of β:

$$\beta'\beta \le 1. \tag{17}$$

A. Dimensionality of the SDF

All deviations from the CAPM in the cross-section of expected returns in our model are caused by sentiment. If the share of sentiment investors was zero, the CAPM would hold. However, as we now show, for sentiment investors' belief distortions to generate a cross-section of expected stock returns with Sharpe ratios comparable to what is found in empirical data, the SDF must have a low-dimensional factor representation.

We combine (16) and (15) to obtain excess SDF variance, expressed, for comparison, as a fraction of the SDF variance accounted for by the market factor,

$$\begin{aligned}
V(\beta) &\equiv \frac{\text{Var}(M) - a^2\sigma_m^2}{a^2\sigma_m^2} \\
&= \frac{\theta^2}{\sigma_m^2} \delta'\Gamma\delta \\
&= \kappa^2 \sum_{k=2}^{N} \beta_k^2 \lambda_k
\end{aligned} \tag{18}$$

33

where $\kappa \equiv \frac{\theta}{\sigma_m}$. From Eq. (18) we see that SDF excess variance is linear in the eigenvalues of the covariance matrix, with weights β_k^2. For the sentiment-driven demand component δ to have a large impact on SDF variance and hence the maximum Sharpe Ratio, the β_k corresponding to high eigenvalues must have a big absolute value. This means that δ must line up primarily with the high-eigenvalue (volatile) principal components of asset returns. The constraint (17) implies that if β did line up with some of the low-eigenvalue PCs instead, the loadings on high-eigenvalue PCs would be substantially reduced and hence the variance of the SDF would be low. As a consequence, either the SDF can be approximated well by a low-dimensional factor model with the first few PCs as factors, or the SDF can't be volatile and hence Sharpe Ratios only very small.

We now assess this claim quantitatively. Figure 6 illustrates this with data based on the covariance matrix of actual portfolios used as Γ and with $\theta = 0.5$. We consider two sets of portfolios: (i) 25 SZ/BM portfolios and (ii) 30 anomaly portfolios underlying the long and short positions in the 15 anomalies in Table I. Returns are in excess of the level factor. We set β to have equal weight on the first K PCs, and zero on the rest. Thus, low K implies that the SDF has a low-dimensional factor representation in terms of the PCs, high K implies that it is a high-dimensional representation in which the high-eigenvalue PCs are not sufficient to represent the SDF. Eq. (18) provides the excess variance of the SDF in each case.

Figure 6 plots the result with K on the horizontal axis. For both sets of portfolios, a substantial SDF excess variance can be achieved only if δ lines up with the first few (high-eigenvalue) PCs and hence the SDF is driven by a small number of principal component factors. If K is high so that δ also lines up with low eigenvalue PCs, then the limited amount of variation in δ permitted by the constraint (10) is neutralized to a large extent
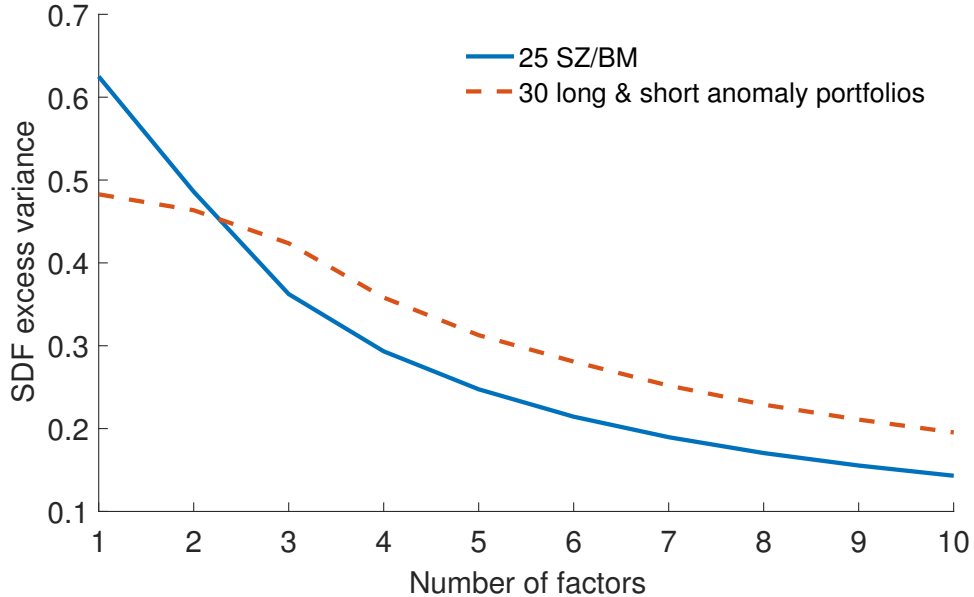
Figure 6. SDF excess variance. The plot shows SDF excess variance, $V(\beta)$, achieved when sentiment investor demands $\delta = Q\beta$ line up equally with first K principal components (ex level factor). The blue solid curve corresponds to 5×5 size-B/M portfolios; the red dashed curve is based on 30 anomaly long and short portfolios.

by arbitrageurs. This is because arbitrageurs find it attractive to trade against sentiment demand if doing so does not require taking on risk exposure to high-eigenvalue PCs.

Similarly to what we argued in Section II, the value of K depends on the payoff space being considered. If the underlying test assets had a weaker factor structure than in the examples we analyzed, sentiment could line up with a higher number of PCs; the "sentiment-based" SDF would contain higher number of factors and could still exhibit high excess variance.

Crucially, our analysis does not directly require there to be a strong factor structure in (biased) beliefs.[14] Instead, the model implies that arbitrageurs activities will ensure that

---

[14]Equation (16) is simply a tautology: since the matrix of eigen-vectors of dividends (returns) Q is of full rank, it forms a basis in $\mathbb{R}^N$ and hence we can "project" δ onto Q as β. Our analysis merely shows that

only belief components that are aligned with loadings on major common factors can have substantial price effects. Other belief components may exist, but they should not affect prices much in equilibrium.

In summary, if the SDF can be represented by a low-dimensional factor model with the first few PCs as factors, this does not necessarily imply that pricing is "rational." Even in an economy in which all deviations from the CAPM are caused by sentiment, one would still expect the SDF to have such a low-dimensional factor representation because only sentiment-driven demand that lines up with the main sources of return co-movement should have much price impact when arbitrageurs are present in the market. Our analysis shows that one could avoid this conclusion only if sentiment investors could take huge leverage and short positions (which would violate our constraint (10)) or if arbitrage capital was largely absent. Neither of these two alternatives seems plausible.

B. Characteristics vs. covariances

Our model sheds further light on the meaning of characteristics vs. covariances tests as in Daniel and Titman (1997), Brennan, Chordia, and Subrahmanyam (1998), and Davis, Fama, and French (2000). As noted in Section II.B, the underlying presumption behind these tests is that "irrational" pricing effects should manifest as mispricing that is orthogonal to covariances with the first few PCs (which implies that mispricing must instead be correlated with low-eigenvalue PCs).

To apply our model to this question, we can think of the belief distortion δ as being

_____

unless the first few elements of such projection (β) are "large", sentiment-driven demands have negligible effect on equilibrium prices. If the first few elements do turn out to be large, the remaining loadings must be "small" due to the constraint in Equation (17).

associated with certain stock characteristics. For example, elements of δ could be high for growth stocks with low B/M due to overextrapolation of recent growth rates or for stocks with low prior 12-month returns due to underreaction to news. We examine whether it is possible that a substantial part of cross-sectional variation in expected returns can be orthogonal to covariances with the first few PCs.

Equilibrium expected returns in our model are given by (12) and hence cross-sectional variance in expected returns is

$$\frac{1}{N}(E[R_{t+1}] - \mu_m \iota)'(E[R_{t+1}] - \mu_m \iota) = a^2 \theta^2 \delta' \Gamma' \Gamma \delta$$
$$= a^2 \theta^2 \beta' \Lambda^2 \beta. \tag{19}$$

The cross-sectional variance in expected returns explained by the first K PCs is

$$a^2 \theta^2 \sum_{k=2}^{K} \beta_k^2 \lambda_k^2. \tag{20}$$

We set $\theta = 0.5$ and take the covariance matrix from empirically observed portfolio returns using two sets of portfolios: the 25 SZ/BM portfolios (with $K = 2$), and the 30 anomaly portfolios (with $K = 3$), both in excess of the level factor. For any choice of β, we can compute the proportion of cross-sectional variation in expected returns explained by the first K principal components, that is, the ratio of (20) to (19), and the ratio of (the upper bound of) cross-sectional variance in expected returns, (19), to squared expected excess market returns. Depending on the choice of the elements of the β vector, various combinations of cross-sectional expected return variance and the share explained by the first K principal components are possible. We search over these by varying the elements of β subject to the constraint (17). In Figure 7 we plot the right envelope, that is, the maximal cross-sectional

37
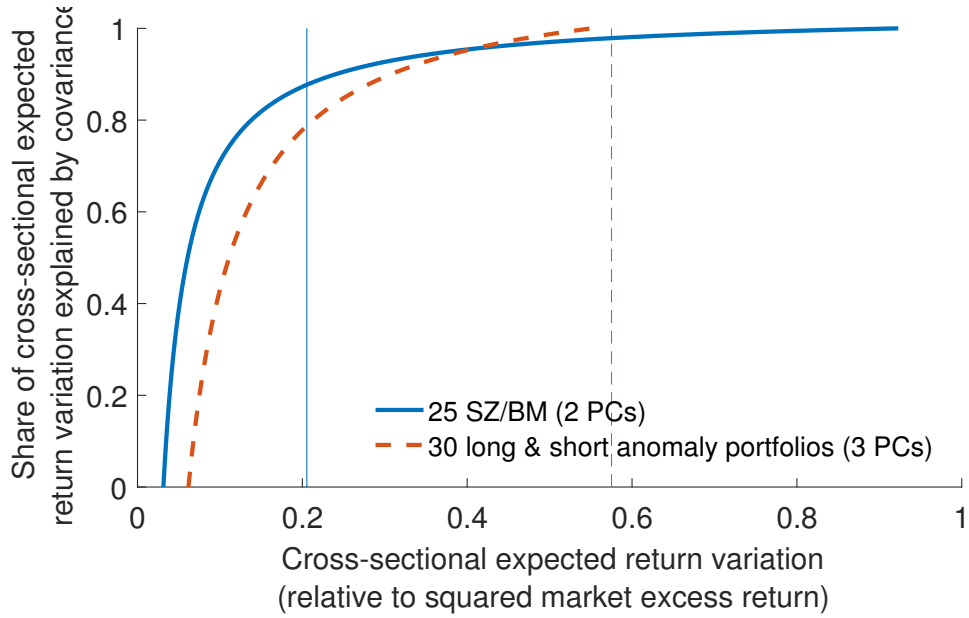
Figure 7. Characteristics vs. covariances. Cross-sectional variation in expected returns explained by the first two principal components for 5×5 size-B/M portfolios and 3 principal components for anomaly long and short portfolio. Portfolio returns are represented in excess of the level factor. Vertical lines depict in-sample estimates of the ratio of cross-sectional variation in expected returns and the squared market excess return for two sets of portfolios.

expected return variation for a given level of share explained by the first K PCs.[15]

As Figure 7 shows, it is not possible to generate much cross-sectional variation in expected returns without having the first two principal components of size-B/M portfolios (in excess of the level factor) and 3 principal components of the 30 anomaly portfolios explain almost all the cross-sectional variation in expected returns of their respective portfolios. For comparison, the ratio of cross-sectional variation in expected returns and the squared market excess return is around 0.20 for the 5×5 size-B/M portfolios and slightly below 0.60 for the anomaly portfolios (depicted with dashed vertical lines on the plot). To achieve these levels of cross-sectional variation in expected returns, virtually all expected return variation has

---

[15]Appendix section B provides more details on the construction of Figure 7.

to be aligned with loadings on the first few principal components.

Thus, despite the fact that all deviations from the CAPM in this model are due to belief distortions, a horse race between characteristics and covariances as in Daniel and Titman (1997) cannot discriminate between a rational and a sentiment-driven theory of the cross-section of expected returns. Covariances and expected returns are almost perfectly correlated in this model—if they weren't, near-arbitrage opportunities would arise, which would not be consistent with the presence of some rational investors in the model.

C. Analyst Forecast Bias as Proxy for Investor Misperceptions

The preceding model and analysis show that loadings on a few large PCs should "explain" the cross-section of expected returns even in "behavioral" models, not just in "rational" models. To devise tests that are more informative about investor beliefs, researchers must exploit additional predictions of the model that relate returns to other data such as macroeconomic variables, information on portfolio holdings, or data on investor beliefs.

Here we briefly explore whether data on beliefs is broadly consistent with the model in Section III. Following La Porta (1996) and Engelberg, McLean, and Pontiff (2015) we use analyst forecasts as proxy for sentiment investor expectations. Under this assumption we can shed light on the connection between belief distortions and factor loadings that's implied by our model.[16]

The prediction of the model in Section III is that belief distortions should "line up" with the large principal component factors in returns; otherwise they would not have much impact on equilibrium asset prices. We now provide some suggestive evidence in favor of this mechanism.

---

[16]We thank an anonymous referee for suggesting this analysis of analyst forecasts.

Following DellaVigna and Pollet (2007), we define the scaled earnings surprise (forecast error) for firm i in quarter q as

$$s_{i,q} = \frac{EPS_{i,q} - \widehat{EPS}_{i,q}}{P_{i,q}}, \tag{21}$$

where $\widehat{EPS}_{i,q}$ is the consensus (mean) forecast from I/B/E/S, $EPS_{i,q}$ is the realized earnings per share, and $P_{i,q}$ is the (split-adjusted) share price five trading days prior to the earnings announcement.[17] If analyst forecasts satisfied rational expectations, the mean surprise should be zero. Defining bias as the difference between the analyst forecast and the rational expectations forecast, the law of iterated expectations implies $bias = -\mathrm{E}\,[s_{i,q}]$. To explore bias in the cross-section, we form calendar-time "portfolios" of earnings surprises for each anomaly.

As a concrete example, consider portfolio sorted on size (equity market capitalization). For each calendar quarter and size decile, we calculate the value-weighted average surprise. We base portfolio assignments on information known prior to the current quarter and use beginning of quarter market capitalization when computing value-weights. Since $s_{i,q}$ is a per-share measure scaled by price, the value-weighted average has an interpretation of a buy-and-hold portfolio "surprise". Finally, we compute *bias* as the negative of the time-series average of the quarterly portfolio surprise. Positive *bias* indicates analyst "optimism" whereas negative bias reflects "pessimism".

If biased beliefs are responsible for CAPM return anomalies, we should find that portfolios with large α have, on average, pessimistic analyst forecasts and vice versa. Table III presents

---

[17]We use the I/B/E/S unadjusted Surprise History file which includes announcements from 1993-2014. Results are similar if we use the unadjusted Summary Statistics file.

the results (we normalize the ordering to go from low to high CAPM α as one reads from left to right). The H-L column conceptually maps to the distortion, δ, in (9). The pattern in analyst forecast bias aligns with the model prediction; bias is systematically more negative for high α portfolios. This analysis is related to the findings in Engelberg, McLean, and Pontiff (2015) that anomaly returns are an order of magnitude higher on earnings days and that anomaly signals predict analyst forecast errors in a way that suggests anomalous returns result from biased expectations.

Our model has an additional implication. If the bound on total bias in (10) is "tight", then behavioral distortions must line up with the large eigenvectors of returns in order to have any impact on equilibrium prices and expected returns. Letting $\tilde{\alpha}$ be the CAPM error of the base assets, Eq. (12) can be transformed as

$$\alpha \propto -\Lambda\beta, \tag{22}$$

where $\alpha \equiv Q'\tilde{\alpha}$ is the CAPM error of PC portfolios. Since Λ is diagonal, Eq. (22) can be read equation-by-equation as $\alpha_i \propto -\lambda_i\beta_i$. This implies α "lines up" in the cross-section with $-\Lambda\beta$. A tight bound implies the first few elements of β are "large" (in magnitude) and the remaining β_j are "small".

Table IV presents estimates of β (behavioral bias, δ, rotated into PC space). For ease of interpretation, all PC portfolios are normalized to have positive expected return (α). The first column shows that most of the large magnitude (significant) β_i are negative, consistent with positive α's. The third column presents $\frac{\beta_i^2}{\beta'\beta}$, the fraction of "total bias" accounted for by each PC. The first PC represents 53%, as predicted if the bound on total bias is tight. Importantly, the model predicts that equilibrium CAPM α arise not necessarily for PCs

41

Table III
Anomalies: Analyst Forecast Bias

We define earnings surprise relative to analyst forecasts as in DellaVigna and Pollet (2007). We value-weight the stock-level surprises by beginning of quarter equity market capitalization to construct calendar quarter portfolio-level surprises. The table reports the negative of the time-series average of the quarterly portfolio-level surprise series. Columns P1 and P10 correspond to short and long ends of long-short strategies (column H-L), respectively. We include intermediate portfolios in columns P4 and P7 (portfolios 4 and 7, respectively). For Piotroski's F-score stocks are sorted into two portfolios due to discreteness of the underlying characteristic variable. t-statistics are Newey-West with 4 lags (1 year). The sample period is 1993Q1 to 2014Q4.

	Low α (P1)	P4	P7	High α (P10)	H-L (δ)	t (δ)
Size	-5.0	-6.7	-5.9	-4.7	0.3	0.3
Gross Profitability	-6.4	-7.9	-5.2	-3.8	2.6	3.0
Value	-3.8	-5.9	-9.5	-9.0	-5.2	4.4
ValProf	-5.1	-4.8	-6.5	-7.7	-2.6	3.0
Accruals	-4.1	-5.5	-5.2	-6.6	-2.5	3.1
Net Issuance (rebal.-A)	-5.6	-6.2	-6.1	-6.0	-0.4	0.4
Asset Growth	-4.0	-4.7	-6.2	-8.1	-4.2	4.5
Investment	-4.3	-4.8	-5.4	-9.9	-5.6	4.6
ValMomProf	-3.8	-5.3	-5.0	-9.6	-5.8	6.2
ValMom	-2.8	-5.2	-6.8	-12.1	-9.3	8.4
Idiosyncratic Volatility	-6.1	-6.2	-6.2	-3.7	2.4	1.5
Momentum	-2.9	-4.8	-5.8	-9.0	-6.1	3.2
Long Run Reversals	-4.6	-4.8	-7.0	-14.2	-9.6	4.3
Beta Arbitrage	-5.9	-6.6	-6.3	-3.0	2.9	3.7
Piotroski's F-score	-5.8	-	-	-5.0	0.8	2.0
Mean	-4.6	-5.7	-6.2	-7.7	-3.1	7.0
Median	-4.5	-5.4	-6.1	-7.9	-3.4	-

Table IV
Forecast Bias in PC Portfolios

The table uses PC portfolios based on 15 long-short anomaly strategies. β is $Q'\delta$ where Q is the eigenvector matrix computed from return covariances, $\Gamma = Q'\Lambda Q$ and δ is the analyst forecast bias from Table III. Mean returns computed using full sample of daily data. t-statistics are Newey-West with 4 lags (1 year). The sample period is 1993Q1 to 2014Q4.

PC	β	t(β)	% of $\beta'\beta$	$-\lambda\beta$	E(R)	% of $\beta'\Lambda^2\beta$
1	-12.6	8.5	52.9	11.4	10.8	90.1
2	2.7	1.3	2.4	-1.6	5.1	1.7
3	-5.6	4.5	10.3	2.3	0.2	3.5
4	1.3	1.3	0.6	-0.4	2.3	0.1
5	-7.9	6.7	20.8	2.3	2.7	3.7
6	-2.5	3.5	2.1	0.6	1.7	0.2
7	2.0	1.6	1.3	-0.4	0.8	0.1
8	4.6	3.5	7.0	-0.8	1.3	0.4
9	-1.6	2.2	0.8	0.2	0.5	0.0
10	-1.0	0.9	0.3	0.1	1.0	0.0
11	0.8	0.7	0.2	-0.1	1.0	0.0
12	-0.1	0.1	0.0	0.0	0.4	0.0
13	1.3	1.8	0.5	-0.1	0.7	0.0
14	0.2	0.3	0.0	-0.0	2.8	0.0
15	1.4	2.4	0.7	-0.1	0.2	0.0

with large β_i, but rather for those with large $\lambda_i\beta_i$ (where λ_i is the variance of returns). The fourth and fifth columns give $-\lambda_i\beta_i$ and $\mathrm{E}[R_i]$, which have 0.8 cross-sectional correlation, consistent with the proportionality prediction of Eq. (22).

Finally, the last column shows the proportion of cross-sectional variance in expected returns, $\beta'\Lambda^2\beta$ (see section III.B), accounted for by each PC. According to the model, this is the partial-R^2 in a regression of individual asset CAPM-$\tilde{\alpha}$ on PC factor loadings. The estimates imply that loadings on the first PC should "explain" 90% of the anomaly returns, suggesting essentially a single factor model of the cross-section (plus a level factor). The strong lining up of belief distortions, δ, with the large eigenvectors of return covariances, Q, suggests the model's mechanism is more than a just theoretical possibility and is plausibly important in generating observed patterns in expected returns.

D. Investment-based expected stock returns

So far our focus has been on the interpretation of empirical reduced-form factor models. There is a related literature that uses reduced-form specifications of the SDF in models of firm decisions with the goal of deriving predictions about the cross-section of stock returns. Our critique that reduced-form factor models have little to say about the beliefs and preferences of investors applies to these models, too.

The models in this literature feature firms that make optimal investment decisions. They generate the prediction that stock characteristics such as the book-to-market ratio, firm size, investment, and profitability should be correlated with expected returns. We discuss two classes of such models. In the first one, firms continuously adjust investment, subject to adjustment costs. One recent example is Lin and Zhang (2013). In the second class, firms are presented with randomly arriving investment opportunities that differ in systematic risk.

44

The firm can either take or reject an arriving project. A prominent example of a model of this kind is Berk, Green, and Naik (1999) (BGN).

Our focus is on the question of whether these models have anything to say about the reason why investors price some stocks to have higher expected returns than others. These theories are often presented as *rational* theories of the cross-section of expected returns that are contrasted with *behavioral* theories in which investors are not fully rational.[18] However, a common feature of these models is that firms optimize taking as given a generic SDF that is not restricted any further. Existence of such a generic SDF requires nothing more than the absence of arbitrage opportunities. Thus, these models make essentially no assumption about investor preferences and beliefs. As a consequence, these models cannot deliver any conclusions about investor preferences or beliefs. As our analysis above shows, it is perfectly possible to have an economy in which all cross-sectional variation in expected returns is caused by sentiment, and yet an SDF not only exists, but it also has a low-dimensional structure in which the first few principal components drive SDF variation, similar to many popular reduced-form factor models. For this reason, models that focus on firm optimization, taking a generic SDF as given, cannot answer the question about investor rationality.

To illustrate, consider a model of firm investment similar to the one in Lin and Zhang (2013). Firms operate in an IID economy, and they take the SDF as given when making real

---

[18]To provide a few examples, BGN, p. 1553, motivate their analysis by pointing to these competing explanations and commenting that "these competing explanations are difficult to evaluate without models that explicitly tie the characteristics of interest to risks and risk premia."; Daniel, Hirshleifer, and Subrahmanyam (2001) cite BGN as a "rational model of value/growth effects"; Grinblatt and Moskowitz (2004) include BGN among "rational risk-based explanations" of past-returns related cross-sectional predictability patters; Johnson (2002) builds a related model based on a reduced-form SDF in a paper with the title "Rational Momentum Effects."

investment decisions. At each point in time, a firm has a one-period investment opportunity. For an investment I_t the firm will make profit Π_{t+1} per unit invested. The firm faces quadratic adjustment costs and the investment fully depreciates after one period. The full depreciation assumption is not necessary for what we want to show, but it simplifies the exposition. To reduce clutter, we also drop the i subscripts for each firm.

Every period, the firm has the objective

$$\max_{I_t} -I_t - \frac{c}{2}I_t^2 + E[M_{t+1}\Pi_{t+1}I_t]. \tag{23}$$

The SDF that appears in this objective function is not restricted any further. Hence, the SDF could be, for example, the SDF (14) from our earlier example economy in which all cross-sectional variation in expected returns is due to sentiment. Taking this SDF as given, we get the firm's first-order condition

$$I_t = -\frac{1}{c} + E[M_{t+1}\Pi_{t+1}] \tag{24}$$

$$= -\frac{1}{c} + E[M_{t+1}] + E[\Pi_{t+1}] + \mathrm{Cov}(M_{t+1}, \Pi_{t+1}). \tag{25}$$

Since the economy features IID shocks, I_t is constant over time, that is, we can write $I_t = I$. The firm's cash flow net of (recurring) investment each period, is

$$D_{t+1} = I\Pi_{t+1} - \frac{c}{2}I^2 - I. \tag{26}$$

If we let Π_{t+1} be normally distributed, this fits into our earlier framework as the cash-flow generating process (with a slight modification to allow for a positive average cash flow and

heterogeneous expected profitability across firms),

$$I = -\frac{1}{c} + E[M_{t+1}] + E[\Pi_{t+1}] + \frac{1}{I}\text{Cov}(M_{t+1}, D_{t+1}), \qquad (27)$$

where M_{t+1} is the SDF (14) that reflects the sentiment investor demand.

Thus, a firm with high $E[\Pi_{t+1}]$ (relative to other firms) must either have high invest-ment or a strongly negative $\text{Cov}(M_{t+1}, D_{t+1})$ (which implies a high expected return). Simi-larly, a firm with high I must either have high profitability or a not very strongly negative $\text{Cov}(M_{t+1}, D_{t+1})$ (which implies a low expected return). Thus, together I and $E[\Pi_{t+1}]$ should explain cross-sectional variation in $\text{Cov}(M_{t+1}, D_{t+1})$ and hence in expected returns.

These relationships arise because firms align their investment decisions with the SDF and the expected return–which is their cost of capital—that they face in the market. From the viewpoint of the firm in this type of model, it is irrelevant whether cross-sectional variation in expected returns is caused by sentiment or not. The implications for firm investment and for the relation between expected returns, investment and profitability are observationally equivalent. Thus, the empirical evidence in Fama and French (2006), Hou, Xue, and Zhang (2014), Novy Marx (2013) that investment and profitability are related, cross-sectionally, to expected stock returns is to be expected in a model in which firms optimize. Moreover, as long as the firm optimizes, the Euler equation $E[M_{t+1}R_{t+1}] = 1$ also holds for the firm's investment return, as in Liu, Whited, and Zhang (2009), again irrespective of whether *investors* are rational or have distorted beliefs.

Testing whether empirical relationships between expected returns, investment, and pro-fitability exist in the data is a test of a model of firm decision-making, but not a test of a model of how investors price assets. Evidence on these empirical relationships does not

47

help resolve the question of how to specify investor beliefs and preferences. Only models that make assumptions about these beliefs and preferences—which result in restrictions on the SDF—can deliver testable predictions that could potentially help discriminate between competing models of how investors price assets.

For example, if one couples a model of firm investment with a standard rational-expectations consumption Euler equation on the investor side (e.g., as in Gomes, Kogan, and Zhang (2003)), then the model makes testable predictions about the identity of the risk-factor in the SDF: covariances with consumption growth should explain the cross-section of expected returns. In this example, modeling of firm investment can provide insights on the relationship between firm characteristics and choices and the systematic consumption risk of the firm, but the firm-investment side of the model does not provide any predictions about the nature of the risks investors care about and what the prices of those risks are.

Turning to the second class of models, we focus on the version of Berk, Green, and Naik (1999) (BGN) with constant interest rates, which is sufficient to produce the key predictions of their model. BGN assert the existence of a generic SDF M that is not restricted any further apart from an auxiliary assumption that M is log-normal. Hence, this SDF could represent, for example, an SDF that arises in an economy in which sentiment causes all cross-sectional variation in expected returns, as in our earlier example economy. All of their conclusions about the relationships between expected returns, firms' book-to-market ratios, and firm size would arise in this model irrespective of the specification of investor beliefs and preferences (rational, behavioral, or otherwise).

Firms in their model are presented with randomly arriving and dying investment projects that all have the same expected profitability and scale, but differ randomly in the covariance of their cash-flows shocks ε_i with the SDF. Projects with very negative $\gamma_i = \text{Cov}(\varepsilon_i, M)$

have a high expected return, that is, a high cost of capital, and are rejected. Ones with less negative γ_i are taken on by the firm. Again, it is important to keep in mind that γ_i is a covariance with a generic SDF. Other than the existence of such an SDF, nothing has been assumed that would imply that γ_i has to represent "rationally priced" risk. Each firm also has an (identical) stock of growth options from the future arrival of new investment projects. Since expected profitability is assumed to be constant in this model and since we work with the constant-interest rate version, the value of these growth options is simply the value of a risk-free bond. At a given point in time, the firm's return covariance with M is then determined by the number of projects, n_t, the firm has taken on in the past that are still alive (relative to constant stock of riskless growth options) and by the aggregated γ_i of the still-alive projects, which we denote γ_t. Since expected excess returns are equal to the negative of the covariance with M, it follows that

$$E[R_{t+1}] = f(n_t, \gamma_t) \qquad (28)$$

for some function $f(.)$. As BGN show (see their equation 45), this leads to a linear relationship between expected returns, the book-to-market ratio and market value,

$$E[R_{t+1}] = a_0 + a_1(B_t/MV_t) + a_2(1/M_t), \qquad (29)$$

where B_t/MV_t depends positively on n_t (as having more ongoing projects reduce the weight on the riskless growth options) and positively on γ_i (as higher expected return lowers market value), while $1/MV_t$ depends negatively on n_t (as more projects taken on raise market value) and positively on γ_i.

Nowhere in this derivation is there any assumption that would restrict investor preferences

and beliefs any further than asserting the existence of an SDF. Thus, if BGN's model of firm decision-making is correct, the conclusions that expected returns are linear in B/MV and $1/MV$, as in (29), would apply in any world in which an SDF exists, even if all cross-sectional variation in expected returns is caused by sentiment (as in our model in Section III). Thus, in terms of investor beliefs and preferences, the BGN model is as much a "behavioral" model as it is a "rational" model.

IV. Factor pricing in economies with sentiment investors: Dynamic case

In this section we show that the observational equivalence between "behavioral" and "rational" asset pricing with regards to factor pricing also applies, albeit to a lesser degree, to partial equilibrium intertemporal capital asset pricing models (ICAPM) in the tradition of Merton (1973). To demonstrate this, we specify and solve a dynamic model with time-varying investor sentiment.

We model the economy in a discrete time and infinite horizon framework. The setup is an extension of the IID model in Section III to the dynamic case when sentiment demand is time-varying. Like in the previous setup, there are N stocks, $i = 1, ..., N$, each in supply of $1/N$ shares, with per-period dividends $D_t \sim \mathcal{N}(0, \Gamma)$. The risk-free one-period bond is in perfectly elastic supply at a constant interest rate of r_F. Define the gross interest rate as $R_F = 1 + r_F$. Finally, we assume there exists a measure $(1 - \theta)$ of arbitrageurs. We model the asset demands of sentiment investors as exogenous and equal to the equilibrium demand in the static model (see (13) and the market clearing condition), but additionally subject to an IID stochastic shock. For sentiment investors we therefore have

$$x_t = \frac{1}{N}\iota + (1 - \theta)\delta\xi_t. \tag{30}$$

Market clearing implies the following demand for arbitrageurs,

$$y_t = \frac{1}{N}\iota - \theta\delta\xi_t, \tag{31}$$

where $\xi_{t+1} \sim \mathcal{N}(0, \omega^2)$ is a time-varying component (scalar) of their demand and δ gives the

direction of sentiment distortion. We assume that δ has a level component, $\delta'\iota \neq 0$, so that market discount rates are not constant. The setup effectively assumes a single time-series factor in sentiment investors demand.[19]

We solve for prices consistent with these equilibrium demands. Arbitrageurs maximize their life-time exponential utility

$$J_t(W_t, \xi_t) = \max_{(C_s, y_s), s \geq t} \mathrm{E}_t \left[-\sum_{s=t}^{\infty} \beta^s \exp(-\alpha C_s) \right], \tag{32}$$

where the maximization is subject to

$$W_{t+1} = (W_t - C_t)R_F + y_t'R_{t+1}, \tag{33}$$

where $R_{t+1} \equiv P_{t+1} - R_F P_t + D_{t+1}$.

We define the market portfolio as $R_{M,t+1} \equiv \frac{1}{N}\iota'R_{t+1}$. We guess that prices are linear in ξ_t and the log value function is quadratic in ξ_t,

$$P_t = a_0 + a_1 \xi_t \tag{34}$$

$$J_t(W_t, \xi_t) = -\beta^t \exp(-\gamma W_t - b_0 - b_1 \xi_t - b_2 \xi_t^2). \tag{35}$$

In Appendix C we show how to solve for the constants a_i and b_i and establish that equilibrium expected returns are given by

$$\mathrm{E}(R_t) = \gamma \mathrm{Cov}(R_t, R_{M,t}) + \tilde{\delta}_{DR}\mathrm{Cov}(R_t, \mathrm{E}_t[R_{M,t+1}]), \tag{36}$$

---

[19]Campbell and Kyle (1993) present a related single-asset model in which ξ follows an Ornstein-Uhlenbeck process.

where $\gamma = \alpha \frac{r_F}{R_F} > 0$ and $\tilde{\delta}_{DR} > 0$ are two positive prices of risk (see Appendix C; Case 2).

Thus, we get an ICAPM similar to Campbell (1993, Eq.25) or a "bad beta, good beta" specification as in Campbell and Vuolteenaho (2004b). The degree of presence of sentiment traders is indirectly reflected in $\text{Cov}(R_t, \text{E}_t[R_{M,t+1}])$. This covariance shrinks to zero as θ goes to zero.

In summary, the analysis shows that time-varying investor sentiment can give rise to an ICAPM-like SDF. As in our static model in the previous section, this model is "behavioral" and "risk-based" at the same time. Deviations from the static CAPM are caused by sentiment, but from the viewpoint of the arbitrageurs, time-varying sentiment generates hedging demands, because it makes the arbitrageurs' investment opportunities time-varying. When evaluating how aggressively to accommodate sentiment investor demand in a particular stock, arbitrageurs consider the covariance of the stock's return with the sentiment-driven investment opportunity state variable. As a result, expected returns reflect this state-variable risk.

V.   Conclusions

Reduced-form factor models are useful to provide a parsimonious summary of the cross-section of asset returns. Yet, their success or failure in explaining the cross-section of asset returns does not help to answer the question whether asset pricing is "rational." As we have shown, even if all cross-sectional variation in expected returns is driven by belief distortions on the part of some investors, a low-dimensional SDF with the first few principal components of returns as factors should still explain asset prices. This only requires that near-arbitrage opportunities are absent. For the same reason, tests that look for stock characteristics capture expected return variation in the cross-section that is orthogonal to common factor covariances are unlikely to be of much help in answering that question either. Therefore, tests of reduced-form factor models cannot shed light on questions regarding the "rationality" of investors.

In fact, the framing of the question concerning investor "rationality" is unhelpfully imprecise in the first place. The arbitrageurs in our model are rational. From their viewpoint, expected returns are consistent with the risk premia that they require as compensation for tilting their portfolio weights away from the market portfolio. But it is the sentiment investor demand that arbitrageurs accommodate which causes these risk premia. Thus, there is no dichotomy between "risk-based" and "behavioral" asset pricing in this model.

The only path to a better understanding of investor beliefs is to develop and test structural asset pricing models with specific assumptions about investor beliefs and preferences that deliver predictions about the factors that should be in the SDF and the probability distribution under which this SDF prices assets. While we discussed these issues in the context of equity markets research, similar conclusions apply to reduced-form no-arbitrage models in bond and currency market research.

The recognition that factor covariances should explain cross-sectional variation in expected returns even in a model of sentiment-driven asset prices should also be useful for the development of models that meet the Cochrane (2011) challenge presented in the introduction of our paper. The answer to his question could be that some components of sentiment-driven asset demands are aligned with covariances with important common factors, some are orthogonal to these factor covariances. Trading by arbitrageurs eliminates the effects of the orthogonal asset demand components, but those that are correlated with common factor exposures survive because arbitrageurs are not willing to accommodate these demands without compensation for the factor risk exposure.

REFERENCES

Ball, Ray, Gil Sadka, and Ronnie Sadka, 2009, Aggregate earnings and asset prices, *Journal of Accounting Research* 47, 1097–1133.

Bansal, R., and A. Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options and security returns, *Journal of Finance* 54, 1553–1607.

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

Campbell, J., 1993, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–512.

Campbell, John, and John Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., and Albert S. Kyle, 1993, Smart money, noise trading and stock price behavior, *Review of Economic Studies* 60, 1–34.

Campbell, John Y, and Luis M Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433–495.

Campbell, John Y., and Tuomo Vuolteenaho, 2004a, Bad beta, good beta, *American Economic Review* 94, 1249–1275.

Campbell, John Y., and Tuomo Vuolteenaho, 2004b, Inflation illusion and stock prices,

American Economic Review Papers and Proceedings 94, 19–23.

Chen, Nai-Fu, Richard Roll, and Stephen A Ross, 1986, Economic forces and the stock market, *Journal of Business*, 383–403.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

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

Cochrane, John H., and Jesus Saá-Requejo, 2000, Beyond arbitrage: Good-deal asset price bounds in incomplete markets, *Journal of Political Economy* 108, 79–119.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, Kent, Sheridan Titman, and KC Wei, 2001, Explaining the cross-section of stock returns in japan: Factors or characteristics?, *Journal of Finance* 56, 743–766.

Daniel, Kent D, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *The Journal of Finance* 56, 921–965.

Davis, James, Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

DellaVigna, Stefano, and Joshua M. Pollet, 2007, Demographics and industry returns, *American Economic Review* 97, 1667–1702.

Engelberg, Joseph, R David McLean, and Jeffrey Pontiff, 2015, Anomalies and news.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 23–49.

Fama, Eugene F., and Kenneth R. French, 1996, Mulitifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–87.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Gomes, J., L. Kogan, and L. Zhang, 2003, Equilibrium cross-section of returns, *Journal of Political Economy* 111, 693–732.

Grinblatt, Mark, and Tobias J. Moskowitz, 2004, Predicting stock price movements from past returns: The role of consistency and tax-loss selling, *Journal of Financial Economics* 71, 541–579.

Hansen, Lars P., and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hou, Kewei, G. Andrew Karolyi, and Bong-Chan Kho, 2011, What factors drive global stock returns?, *Review of Financial Studies* 24, 2527–2574.

Hou, Kewei, Chen Xue, and Lu Zhang, 2014, Digesting anomalies, *Review of Financial Studies*.

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Kim, Tong Suk, and Edward Omberg, 1996, Dynamic nonmyopic portfolio behavior, *Review of financial studies* 9, 141–161.

Kogan, Leonid, and Mary Tian, 2015, Firm characteristics and empirical factor models: a model-mining experiment, Working paper, MIT.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *The Journal of Finance* 51, 1715–1742.

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

Li, Qing, Maria Vassalou, and Yuhang Xing, 2006, Sector investment growth rates and the cross section of equity returns, *Journal of Business* 79, 1637–1665.

Lin, X., and Lu Zhang, 2013, The investment manifesto, *Journal of Monetary Economics* 60, 351—66.

Liu, Laura X., Toni M. Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105–1139.

Liu, Laura Xiaolei, and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Liu, Laura Xiaolei, and Lu Zhang, 2014, A neoclassical interpretation of momentum, *Journal of Monetary Economics* 67, 109–128.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

MacKinlay, A Craig, 1995, Multifactor models do not explain deviations from the capm, *Journal of Financial Economics* 38, 3–28.

McLean, David R., and Jeffrey Pontiff, 2016, Does Academic Research Destroy Stock Return Predictability?, *Journal of Finance* 71, 5–32.

Merton, Robert C, 1973, An intertemporal capital asset pricing model, *Econometrica: Journal of the Econometric Society*, 867–887.

Nawalkha, Sanjay K, 1997, A multibeta representation theorem for linear asset pricing theories, *Journal of Financial Economics* 46, 357–381.

Novy Marx, Robert, 2013, The Other Side of Value: The Gross Profitability Premium, *Journal of Financial Economics* 108, 1–28.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Reisman, Haim, 1992, Reference variables, factor structure, and the approximate multi-beta representation, *Journal of Finance* 47, 1303–1314.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–60.

Shanken, Jay, 1992, The current state of the arbitrage pricing theory, *Journal of Finance* 47, 1569–1574.

Stambaugh, Robert F., and Yu Yuan, 2015, Mispricing factors, Working paper, Wharton School.

Appendix A: Absence of near-arbitrage

This section presents the derivation of the SDF Variance. Define

$$\omega_m = \frac{1}{\sqrt{N}} q_1 \tag{A1}$$

$$\mu_m = \omega_m' \mu \tag{A2}$$

$$\sigma_m^2 = \omega_m' \Gamma \omega_m. \tag{A3}$$

The last definition implies that $\sigma_m^2 = \frac{\lambda_1}{N}$. Then

$$\text{Var}\,(M) = \frac{\mu_m^2}{\sigma_m^2} + (\mu - \mu_m)' Q_z \Lambda_z^{-1} Q_z' (\mu - \mu_m). \tag{A4}$$

Let

$$\omega_k = \frac{1}{\sqrt{N}} q_k \tag{A5}$$

$$\mu_k = \omega_k' \mu \tag{A6}$$

$$\sigma_k^2 = \omega_k' \Gamma \omega_k. \tag{A7}$$

The last definition implies that $\sigma_k^2 = \frac{\lambda_k}{N}$. σ_k^2 is decreasing from second to higher-order PCs proportional to eigenvalue. We refer to R_k as the return on the zero-investment portfolio associated with the k-th principal component.

Then

$$\text{Var}(M) = \mu_m^2 / \sigma_m^2 + \sum_{k=2}^{N} N^2 \frac{\overline{\text{Cov}}(\mu_i, q_{ki})^2}{\lambda_k} \tag{A8}$$

$$= \mu_m^2 / \sigma_m^2 + \overline{\text{Var}}(\mu_i) \sum_{k=2}^{N} \frac{\overline{Corr}(\mu_i, q_{ki})^2}{\sigma_k^2}. \tag{A9}$$

Covariance is a cross-sectional covariance, and for the second line we used the fact that q_{ki} is mean zero and has variance N^{-1}. The sum of the squared correlations is equal to one. But the sum weighted by the inverse σ_k^2 depends on which of the PCs μ lines up with. If it lines up with high σ_k^2 PCs then the sum is much lower than if it lines up with low σ_k^2 PCs. Thus, if expected returns line up with low-eigenvalue PCs, then we get much higher SR.

Appendix B: Characteristics vs. covariances

This subsection provides additional detail on the construction of the Figure 7.

Figure 7 plots the right envelope of the set generated of all β that satisfy restriction (17). To construct this right envelope we put all weight of β onto two eigenvectors: the eigenvector associated with the highest eigenvalue (1-st eigenvector) and the $(K+1)$-th eigenvector, that is, the eigenvector associated with the highest principal component from the remainder of $N - K$ principal components not used in equation (20)). We then vary weights on these two components in a way that satisfies (17). For each set of weights, we compute the ratio of (20) to (19), and the ratio of (the upper bound of) cross-sectional variance in expected returns, (19), to squared expected excess market returns.

Appendix C: Dynamic Model

We solve a more general case of the model by assuming that sentiment investors demand follows an AR(1): $\xi_{t+1} \sim \mathcal{N}(\hat{\xi}_t, \omega^2)$, and the mean of ξ_{t+1} is given by

$$\hat{\xi}_t \equiv \mu + \phi \xi_t. \tag{C1}$$

The model can be easily specialized to the case considered in Section IV by setting $\mu = \phi = 0$.

Bellman equation is given by

$$J_t(W_t, \xi_t) = \max_{C_t, y_t} \left\{ -\beta^t \exp(-\alpha C_t) + \mathrm{E}_t \left[J_{t+1}(W_{t+1}, \xi_{t+1}) \right] \right\}. \tag{C2}$$

Guess

$$P_t = a_0 + a_1 \xi_t \tag{C3}$$

$$J_t(W_t, \xi_t) = -\beta^t \exp(-\gamma W_t - b_0 - b_1 \xi_t - b_2 \xi_t^2). \tag{C4}$$

where a_0 and a_1 are vectors of constants and γ, b_0, b_1, and b_2 are scalars. Note that, based on this guess,

$$R_{t+1} = D_{t+1} + a_1(\xi_{t+1} - \xi_t) - r_F a_0 - r_F a_1 \xi_t. \tag{C5}$$

Then

$$\mathrm{E}_t \left[J_{t+1} \right] = - \beta^{t+1} \exp \left(-\gamma \left[(W_t - C_t) R_F - y_t' (r_F a_0 + R_F a_1 \xi_t) \right] - b_0 \right)$$

$$\times \mathrm{E}_t \left[\exp \left(-\gamma y_t' D_{t+1} \right) \right]$$

$$\times \mathrm{E}_t \left[\exp \left(- (b_1 + \gamma y_t' a_1) \xi_{t+1} - b_2 \xi_{t+1}^2 \right) \right], \tag{C6}$$

where

$$\mathrm{E}_t \left[\exp\left(-\gamma y_t' D_{t+1} \right) \right] = \exp\left(\frac{1}{2} \gamma^2 y_t' \Gamma y_t \right) \tag{C7}$$

as the expectation of log-normal. We can rewrite the last expectation in (C6) as

$$\mathrm{E}_t \left[\exp\left(-\left(b_1 + \gamma y_t' a_1 \right) \xi_{t+1} - b_2 \xi_{t+1}^2 \right) \right] = \exp\left(b_2 \lambda_t^2 \right) \times \mathrm{E}_t \left[\exp\left(-b_2 \omega^2 \zeta_{t+1}^2 \right) \right], \tag{C8}$$

where $\zeta_{t+1} = \frac{\xi_{t+1} + \lambda_t}{\omega} \sim \mathrm{N}\left(\hat{\zeta}_t, 1 \right)$, $\hat{\zeta}_t = \frac{\hat{\xi}_t + \lambda_t}{\omega}$, and $\lambda_t = \frac{b_1 + \gamma y_t' a_1}{2b_2}$. Using the moment generating function of the noncentral χ^2 distribution we compute the expectation as follows[20]:

$$\mathrm{E}_t \left[\exp\left(-b_2 \omega^2 \zeta_{t+1}^2 \right) \right] = \frac{1}{\sqrt{1 + 2b_2\omega^2}} \exp\left(\frac{-b_2 \omega^2 \hat{\zeta}_t^2}{1 + 2b_2\omega^2} \right). \tag{C9}$$

The first-order condition for y_t yields:

$$0 = \gamma(r_F a_0 + R_F a_1 \xi_t) + \gamma^2 \Gamma y_t - \frac{1}{1 + 2b_2\omega^2} \left(\frac{b_1 + \gamma y_t' a_1}{2b_2} + \hat{\xi}_t \right) \gamma a_1 + \frac{b_1 + \gamma y_t' a_1}{2b_2} \gamma a_1. \tag{C10}$$

We can compute a_0 and a_1 by plugging in the market clearing condition for y_t in equation (31) and applying the method of undetermined coefficients (since market clearing has to hold for any value of ξ_t). We omit these calculations for brevity.

Now, going back to (C6), we can write

$$\mathrm{E}_t \left[J_{t+1}(W_{t+1}, \xi_{t+1}) \right] = -\beta^{t+1} \exp\left(-\gamma(W_t - C_t) R_F + \hat{\lambda}_0 + \hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2 \right), \tag{C11}$$

where $\hat{\lambda}_0$, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ can be easily solved for by plugging in equations (C7), (C8), (C9), and (31) into (C6).

---

[20]We assume the technical condition $-b_2\omega^2 < 0.5$ is satisfied.

Now we evaluate the first-order condition for consumption

$$U'(C_t) = \frac{\partial \mathrm{E}_t\left[J_{t+1}(W_{t+1}, \xi_{t+1})\right]}{\partial C_t}. \tag{C12}$$

After taking logs,

$$\log \alpha - \alpha C_t = \log(\beta \gamma R_F) - \gamma R_F(W_t - C_t) + \hat{\lambda}_0 + \hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2, \tag{C13}$$

which we can solve for

$$C_t = \frac{\gamma R_F}{\alpha + \gamma R_F} W_t + \frac{1}{\alpha + \gamma R_F}\left[\log\left(\frac{\alpha}{\beta \gamma R_F}\right) - \hat{\lambda}_0 - \hat{\lambda}_1 \xi_t - \hat{\lambda}_2 \xi_t^2\right]. \tag{C14}$$

Substituting this solution into the Bellman equation, and using the fact that $1 - \gamma R_F/(\alpha + \gamma R_F) = \alpha/(\alpha + \gamma R_F)$ to calculate, we get

$$
\begin{aligned}
&\exp(-\gamma W_t - b_0 - b_1 \xi_t - b_2 \xi_t^2) \\
&= \exp(-\alpha C_t) + \beta \exp[-\gamma(W_t - C_t)R_F + \hat{\lambda}_0 + \hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2] \\
&= \exp\left(-\frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t\right) \exp\left\{-\frac{\alpha}{\alpha + \gamma R_F} \log\left(\frac{\alpha}{\beta \gamma R_F}\right)\right\} \exp\left\{\frac{\alpha}{\alpha + \gamma R_F}\left(\hat{\lambda}_0 + \hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2\right)\right\} \\
&\quad + \beta \exp\left(-\frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t\right) \exp\left\{\frac{\gamma R_F}{\alpha + \gamma R_F} \log\left(\frac{\alpha}{\beta \gamma R_F}\right)\right\} \exp\left\{\frac{\alpha}{\alpha + \gamma R_F}\left(\hat{\lambda}_0 + \hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2\right)\right\} \\
&= \left(\frac{\alpha}{\beta \gamma R_F}\right)^{-\frac{\alpha}{\alpha + \gamma R_F}} \left(1 + \beta\left(\frac{\alpha}{\beta \gamma R_F}\right)\right) \exp\left\{\frac{\alpha}{\alpha + \gamma R_F} \hat{\lambda}_0\right\} \\
&\quad \times \exp\left(-\frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t\right) \exp\left\{\frac{\alpha}{\alpha + \gamma R_F}\left(\hat{\lambda}_1 \xi_t + \hat{\lambda}_2 \xi_t^2\right)\right\}. \tag{C15}
\end{aligned}
$$

Comparing coefficients, we get

$$\gamma = \frac{\alpha\gamma R_F}{\alpha + \gamma R_F} = \alpha \frac{r_F}{R_F} \tag{C16}$$

$$b_1 = -\frac{\alpha}{\alpha + \gamma R_F}\hat{\lambda}_1 = -\frac{\hat{\lambda}_1}{R_F} \tag{C17}$$

$$b_2 = -\frac{\alpha}{\alpha + \gamma R_F}\hat{\lambda}_2 = -\frac{\hat{\lambda}_2}{R_F}, \tag{C18}$$

and one can solve along similar lines for b_0.

Asset pricing. Having solved for these coefficients, we can now look at asset pricing. Expected returns are given by

$$E_t(R_{t+1}) = -r_F a_0 - R_F a_1 \xi_t + a_1 \hat{\xi}_t. \tag{C19}$$

Combining equations (C10) and (C19) yields:

$$E_t(R_{t+1}) = a_1 \hat{\xi}_t + \gamma \Gamma y_t - \frac{1}{1 + 2b_2\omega^2}\left(\frac{b_1 + \gamma y_t' a_1}{2b_2} + \hat{\xi}_t\right)a_1 + \frac{b_1 + \gamma y_t' a_1}{2b_2}a_1 \tag{C20}$$

$$= \gamma \Gamma y_t + \frac{\gamma}{1 + 2b_2\omega^2}\omega^2 a_1 a_1' y_t + \frac{b_1 + 2b_2\hat{\xi}_t}{1 + 2b_2\omega^2}\omega^2 a_1. \tag{C21}$$

Note that the following identities hold:

$$\text{Cov}_t(D_{t+1}, R_{A,t+1}) = \Gamma y_t \tag{C22}$$

$$\text{Cov}_t(R_{t+1}, \xi_{t+1}) = \omega^2 a_1, \tag{C23}$$

where $R_{A,t+1} = R'_{t+1} y_t$ is the return on arbitrageur's investment portfolio. We plug these identities into equation (C21) to obtain the pricing equation

$$E_t(R_{t+1}) = \delta_{CF} \times \text{Cov}_t(D_{t+1}, R_{A,t+1}) + \delta_{\xi,t} \times \text{Cov}_t(R_{t+1}, \xi_{t+1}), \tag{C24}$$

where $\delta_{CF} = \gamma$ and $\delta_{\xi,t} = \frac{\gamma a'_1 y_t}{1+2b_2\omega^2} + \frac{b_1 + 2b_2\hat{\xi}_t}{1+2b_2\omega^2}$.

Case 1. Sentiment investors demands are i.i.d. with non-zero mean, that is, $\hat{\xi}_t = \mu, \phi = 0.$ We can easily condition down the equation above in this case:

$$E(R_t) = \delta_{CF} \times \text{Cov}(D_t, R_{A,t}) + \delta_{\xi} \times \text{Cov}(R_t, \xi_t), \tag{C25}$$

where we used the fact that $E_{t-1}[D_t] = 0$ and $E_{t-1}[\xi_t]$ are constant. All prices of risk in this case are constant, δ_{CF} is as above and $\delta_{\xi} = \frac{\gamma a'_1 \left(\frac{1}{N}\iota - \theta\mu\delta\right)}{1+2b_2\omega^2} + \frac{b_1 + 2\mu b_2}{1+2b_2\omega^2}$.

Case 2. Sentiment investors demands are i.i.d. with zero mean, that is, $\mu = \phi = 0.$ This is the case described in Section IV. Specialize (C21) and condition down:

$$E(R_t) = \gamma\Gamma\frac{1}{N}\iota + \left(\frac{\gamma\frac{1}{N}a'_1\iota + b_1}{1+2b_2\omega^2}\right) \times \omega^2 a_1 \tag{C26}$$

$$= \gamma\text{Cov}(D_t, R_{M,t}) + \left(\frac{\gamma\frac{1}{N}a'_1\iota + b_1}{1+2b_2\omega^2}\right)\text{Cov}(R_t, \xi_t). \tag{C27}$$

We used the fact that $E(\xi_t) = 0$ (Case 2) and $R_{A,t+1} = R_{M,t+1} - \theta\xi_t R_{\delta,t+1}$, where $R_{\delta,t+1} = R'_{t+1}\delta$ is the return on a portfolio driven by sentiment demand distortions.

Further note that

$$\mathrm{Cov}\left(R_t, R_{M,t}\right) = \left[\mathrm{Cov}\left(D_t, R_{M,t}\right) + \frac{1}{N}a_1'\iota \times \omega^2 a_1\right] + R_F^2 \frac{1}{N}a_1'\iota \times \omega^2 a_1, \qquad \text{(C28)}$$

where the term in the square brackets reflects the expectation of conditional covariance and the last term is the covariance of conditional expectations. Plugging this in (C27) yields:

$$\mathrm{E}\left(R_t\right) = \gamma\mathrm{Cov}\left(R_t, R_{M,t}\right) + \tilde{\delta}_\xi \mathrm{Cov}\left(R_t, \xi_t\right), \qquad \text{(C29)}$$

where

$$\tilde{\delta}_\xi = \frac{b_1 + \gamma\frac{1}{N}a_1'\iota - \gamma\left(1 + R_F^2\right)\left(1 + 2b_2\omega^2\right)\frac{1}{N}a_1'\iota}{1 + 2b_2\omega^2} \qquad \text{(C30)}$$

$$= \frac{b_1}{1 + 2b_2\omega^2} - \gamma R_F^2 \frac{1}{N}a_1'\iota - \frac{2b_2\omega^2}{1 + 2b_2\omega^2}\gamma\frac{1}{N}a_1'\iota. \qquad \text{(C31)}$$

Without loss of generality, assume that $\delta'\iota > 0$, that is, positive shocks to ξ_t correspond to positive sentiment demand for the market portfolio. One can also show that the price of the aggregate market portfolio is increasing in sentiment demand, $a_1'\iota > 0$, and the arbitrageur's value function is decreasing in the amount of aggregate sentiment (increasing in aggregate market expected returns), $b_1 < 0$, $b_2 > 0$ (see, for instance, Campbell and Kyle (1993), Campbell and Viceira (1999), Kim and Omberg (1996)). These facts together imply that $\tilde{\delta}_\xi < 0$.

Finally, note that

$$\mathrm{Cov}\left(R_t, \mathrm{E}_t R_{M,t+1}\right) = -R_F\frac{1}{N}a_1'\iota \times \mathrm{Cov}\left(R_t, \xi_t\right). \qquad \text{(C32)}$$

We can therefore rewrite (C29) as follows

$$\mathrm{E}\left(R_t\right) = \gamma \mathrm{Cov}\left(R_t, R_{M,t}\right) + \tilde{\delta}_{DR}\mathrm{Cov}\left(R_t, \mathrm{E}_t R_{M,t+1}\right), \tag{C33}$$

where $\tilde{\delta}_{DR} = -\tilde{\delta}_\xi \left(R_F \frac{1}{N} a_1' \iota\right)^{-1} > 0.$