Interpreting factor models *

Serhiy Kozak †
University of Michigan

Stefan Nagel‡
University of Michigan, NBER and CEPR

Shrihari Santosh§
University of Maryland

November 2015

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.

*We are grateful for comments from Kent Daniel, David Hirshleifer, Stijn van Nieuwerburgh, Ken Singleton, Annette Vissing-Jorgensen, participants at the American Finance Association Meetings, Copenhagen FRIC conference, NBER Summer Institute, and seminars at the University of Maryland, Michigan, MIT, and Stanford.

†Stephen M. Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109, sekozak@umich.edu

‡Stephen M. Ross School of Business and Department of Economics, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109, e-mail: stenagel@umich.edu

§Robert H. Smith School of Business, University of Maryland, e-mail: scsantosh@rhsmith.umd.edu
1 Introduction

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 tautology.\footnote{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.} 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:

\textit{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, “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 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

\(^2\)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).”
approximation as a function of these few dominant factors. 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 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. 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 be reliably persist out of sample.

It would be wrong, however, to jump from the evidence that expected returns line up with

---

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.
common factor covariances to the conclusion that the idea of sentiment-driven asset 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 (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, and Xing 2006; Liu and Zhang 2008) are not more informative either. As shown in Reisman (1992) (see, also, Shanken 1992; Nawalkha 1997; 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 (2004) in which arbitrageurs demand risk compensation only for cash-flow beta (“bad beta”) exposure, but not for discount-rate beta (“good beta”) exposure due to loadings on the transitory sentiment-demand factor.

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 covariance 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 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 2 we describe the portfolio returns data that we use in this study. In Section 3 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 4 demonstrates the model in which fully rational risk averse arbitrageurs trade with sentiment investors. Section 5 develops a model with time-varying investor sentiment, which results in an ICAPM-type hedging demand.
2 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 from Novy-Marx and Velikov (2014) 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. Second, for comparison, we also use the $5 \times 5$ Size (SZ) and Book-to-Market (BM) sorted portfolios of Fama and French (1993).

Table 1 provides some descriptive statistics for the anomaly long-short portfolios. Mean returns on long-short strategies range from 0.20% to 1.43% per month. Annualized squared SRs, shown in the second column, range from 0.02 to 1.09. 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 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 1 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 0.01 for the Beta Arbitrage strategy to 0.65 for the Size strategy. Once the first five PCs are included in the regression, the explanatory power is more uniform, with $R^2$ ranging from 0.11 for the Accruals strategy to 0.96

---

4We thank Robert Novy-Marx and Ken French for making the portfolio returns available on their websites. From those available on Novy-Marx’s website, we use those strategies that are available starting in 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 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 1: Anomalies: Returns and Principal Component Factors

The sample period is August 1963 to December 2013. The anomaly long-short strategy returns are from Novy-Marx and Velikov (2014). Average returns are reported in percent per month. 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.

<table>
<thead>
<tr>
<th>Factor</th>
<th>PC1 Return</th>
<th>Squared SR</th>
<th>PC1-2</th>
<th>PC1-3</th>
<th>PC1-4</th>
<th>PC1-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>0.33</td>
<td>0.06</td>
<td>0.12</td>
<td>0.65</td>
<td>0.73</td>
<td>0.77</td>
</tr>
<tr>
<td>Gross Profitability</td>
<td>0.41</td>
<td>0.17</td>
<td>0.00</td>
<td>0.05</td>
<td>0.11</td>
<td>0.13</td>
</tr>
<tr>
<td>Value</td>
<td>0.48</td>
<td>0.15</td>
<td>0.00</td>
<td>0.38</td>
<td>0.40</td>
<td>0.74</td>
</tr>
<tr>
<td>ValProf</td>
<td>0.83</td>
<td>0.54</td>
<td>0.01</td>
<td>0.33</td>
<td>0.33</td>
<td>0.44</td>
</tr>
<tr>
<td>Accruals</td>
<td>0.27</td>
<td>0.09</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Net Issuance (rebal.-A)</td>
<td>0.77</td>
<td>0.82</td>
<td>0.14</td>
<td>0.17</td>
<td>0.23</td>
<td>0.37</td>
</tr>
<tr>
<td>Asset Growth</td>
<td>0.37</td>
<td>0.13</td>
<td>0.05</td>
<td>0.27</td>
<td>0.27</td>
<td>0.51</td>
</tr>
<tr>
<td>Investment</td>
<td>0.57</td>
<td>0.40</td>
<td>0.02</td>
<td>0.24</td>
<td>0.24</td>
<td>0.32</td>
</tr>
<tr>
<td>Piotroski’s F-score</td>
<td>0.20</td>
<td>0.02</td>
<td>0.15</td>
<td>0.27</td>
<td>0.36</td>
<td>0.37</td>
</tr>
<tr>
<td>ValMomProf</td>
<td>1.43</td>
<td>1.09</td>
<td>0.04</td>
<td>0.46</td>
<td>0.75</td>
<td>0.81</td>
</tr>
<tr>
<td>ValMom</td>
<td>0.93</td>
<td>0.46</td>
<td>0.05</td>
<td>0.45</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Idiosyncratic Volatility</td>
<td>0.63</td>
<td>0.09</td>
<td>0.50</td>
<td>0.60</td>
<td>0.76</td>
<td>0.94</td>
</tr>
<tr>
<td>Momentum</td>
<td>1.32</td>
<td>0.45</td>
<td>0.06</td>
<td>0.12</td>
<td>0.84</td>
<td>0.95</td>
</tr>
<tr>
<td>Long Run Reversals</td>
<td>0.48</td>
<td>0.11</td>
<td>0.02</td>
<td>0.57</td>
<td>0.67</td>
<td>0.73</td>
</tr>
<tr>
<td>Beta Arbitrage</td>
<td>0.50</td>
<td>0.14</td>
<td>0.00</td>
<td>0.01</td>
<td>0.07</td>
<td>0.55</td>
</tr>
</tbody>
</table>

For the Momentum strategy, with most strategies having $R^2$ above 0.6. 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 BM 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 $5 \times 5$ 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
Figure 1: Eigenvector weights corresponding to the second and third principal components of Fama-French 25 SZ/BM portfolio returns.

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, i.e., it is similar to SMB. PC2, shown on the right, has positive weights on high B/M stocks and negative weights on low B/M stocks, i.e., 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.⁵

⁵A 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 explain the cross-section of expected returns on these portfolios about as well as the Fama-French factors do.
3 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 $t = 0, 1, 2, \ldots$. 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 $\Gamma$.

Assume that the Law of One Price (LOP) holds. The LOP is equivalent to the existence of an SDF $M$ such that $E[MR] = 0$. Note that $E[\cdot]$ represents objective expectations of the econometrician, but there is no presumption here that $E[\cdot]$ 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

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

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

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

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 $E[MR] = 0$ as the first-order condition of the arbitrageurs’ optimization problem and hence of the SDF as representing the marginal utility of the arbitrageur.

For example, for an arbitrageur with exponential utility (as we show below in Section 4) the first-order condition implies $M = 1 - a[R^A - E(R^A)]$, where $R^A$ represents the return on the arbitrageur’s wealth portfolio and $a$ is the arbitrageur’s risk aversion. As long as the arbitrageur can hold a relatively diversified and not too highly levered portfolio, $R^A$ will not have extremely high volatility, which keeps the variance of $M$ bounded from above. 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. Our model in Section 4 makes these arguments more precise, but for now it suffices to say that an upper bound on the Sharpe Ratio is perfectly consistent with asset prices that are largely sentiment-driven.

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

$$\Gamma = Q\Lambda Q' \quad \text{with} \quad Q = (q_1, \ldots, q_N)$$

(3)
and $\lambda_i$ as the diagonal elements of $\Lambda$. Assume that the first principal component (PC) is a level factor, i.e., $q_1 = \frac{1}{\sqrt{N}}\iota$, where $\iota$ is a conformable vector of ones. This implies $q_k'\iota = 0$ for $k > 1$, i.e., the remaining PCs are long-short portfolios. In the Appendix, Section A we show that

$$
\text{Var}(M) = (\mu'q_1)^2\lambda_1^{-1} + \mu'Q\Lambda^{-1}Q'\mu
= \frac{\mu_m^2}{\sigma_m^2} + N\text{Var}(\mu)\sum_{k=2}^{N} \frac{\text{Corr}(\mu_i, q_{ki})^2}{\lambda_k},
$$

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 $\text{Var}(.)$ and $\text{Corr}(.)$ denote cross-sectional variance and correlation. This expression for SDF variance shows that expected returns must line up with the first few (high-eigenvalue) PCs, otherwise $\text{Var}(M)$ would be huge. To see this, note that the sum of the squared correlations of $\mu_i$ and $q_{ki}$ is always equal to one. But the magnitude of the sum weighted by the inverse $\lambda_k$ depends on which of the PCs the vector $\mu$ lines up with. If it lines up with high $\lambda_k$ PCs then the sum is much lower than if it lines up with low $\lambda_k$ PCs. For typical test assets, eigenvalues decay rapidly beyond the first few PCs. In this case, a high correlation of $\mu_i$ with a low-eigenvalue $q_{ki}$ would lead to an enormous maximum Sharpe Ratio. We now turn to an empirical analysis that demonstrates this point.

### 3.1 Principal components as reduced-form factors: Evidence from anomaly portfolios

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 $5 \times 5$ size-B/M portfolios are similar to the three Fama-French factors. We now investigate the pricing performance of principal component factor models.

Table 2 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
Table 2: Explaining Anomalies with Principal Component Factors

The sample period is August 1963 to December 2013. The anomaly long-short strategy returns are from Novy-Marx and Velikov (2014). Average returns and factor-model alphas are reported in percent 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.

<table>
<thead>
<tr>
<th>Anomaly Factor</th>
<th>Mean Return</th>
<th>PC1-2</th>
<th>PC1-3</th>
<th>PC1-4</th>
<th>PC1-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>0.33</td>
<td>0.17</td>
<td>-0.46</td>
<td>-0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Gross Profitability</td>
<td>0.41</td>
<td>0.42</td>
<td>0.55</td>
<td>0.31</td>
<td>0.38</td>
</tr>
<tr>
<td>Value</td>
<td>0.48</td>
<td>0.49</td>
<td>0.01</td>
<td>0.17</td>
<td>-0.17</td>
</tr>
<tr>
<td>ValProf</td>
<td>0.83</td>
<td>0.86</td>
<td>0.46</td>
<td>0.54</td>
<td>0.36</td>
</tr>
<tr>
<td>Accruals</td>
<td>0.27</td>
<td>0.34</td>
<td>0.32</td>
<td>0.30</td>
<td>0.27</td>
</tr>
<tr>
<td>Net Issuance (rebal.-A)</td>
<td>0.77</td>
<td>0.88</td>
<td>0.80</td>
<td>0.59</td>
<td>0.43</td>
</tr>
<tr>
<td>Asset Growth</td>
<td>0.37</td>
<td>0.45</td>
<td>0.14</td>
<td>0.18</td>
<td>-0.06</td>
</tr>
<tr>
<td>Investment</td>
<td>0.57</td>
<td>0.61</td>
<td>0.35</td>
<td>0.36</td>
<td>0.25</td>
</tr>
<tr>
<td>Piotroski’s F-score</td>
<td>0.20</td>
<td>0.36</td>
<td>0.65</td>
<td>0.25</td>
<td>0.18</td>
</tr>
<tr>
<td>ValMomProf</td>
<td>1.43</td>
<td>1.51</td>
<td>0.95</td>
<td>0.22</td>
<td>0.38</td>
</tr>
<tr>
<td>ValMom</td>
<td>0.93</td>
<td>1.03</td>
<td>0.50</td>
<td>-0.29</td>
<td>-0.32</td>
</tr>
<tr>
<td>Idiosyncratic Volatility</td>
<td>0.63</td>
<td>1.10</td>
<td>1.51</td>
<td>0.67</td>
<td>0.25</td>
</tr>
<tr>
<td>Momentum</td>
<td>1.32</td>
<td>1.47</td>
<td>1.16</td>
<td>-0.49</td>
<td>-0.17</td>
</tr>
<tr>
<td>Long Run Reversals</td>
<td>0.48</td>
<td>0.42</td>
<td>-0.24</td>
<td>0.21</td>
<td>0.05</td>
</tr>
<tr>
<td>Beta Arbitrage</td>
<td>0.50</td>
<td>0.52</td>
<td>0.58</td>
<td>0.25</td>
<td>-0.19</td>
</tr>
</tbody>
</table>

| PC factors’ max. squared SR | 3.86 | 0.10 | 0.49 | 1.47 | 1.70 | 1.72 |

\( \chi^2 \)-pval. for zero pricing errors

For comparison:

| 25 SZ/BM | 2.44 | 0.23 | 0.37 | 0.65 | 0.76 | 0.77 |
| \( \chi^2 \)-pval. for zero pricing errors | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| MKT, SMB, and HML | 0.59 | - | - | - | - | - |

14
returns. The upper panel in Table 2 reports the pricing errors, i.e., the intercepts or alphas, from these regressions. The raw mean excess return (in percent per month) 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 1.51 percent per month. Adding PC2 and PC3 to the factor model drastically shrinks the pricing errors. With five factors, the maximum (absolute) alpha is 0.43.

The bottom panel reports the (ex post) maximum squared SR of the anomaly portfolios (3.86) and the maximum squared SR of the PC factors. With five factors, the highest-SR combination of the factors achieves a squared SR of 1.72. This is still considerably below the maximum squared SR of the anomaly portfolios and the $p$-values from a $\chi^2$-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 $5 \times 5$ size-B/M portfolios—which is typically regarded as a success. As the Table shows, the maximum squared SR of the $5 \times 5$ size-B/M portfolios is 2.44. But the squared SR of MKT, SMB, and HML is only 0.59. 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 0.65 and gets slightly closer 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 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.
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 do so 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 equation (4). We assume that $\mu_i$ is correlated with $K$ 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. Since the sum of all squared correlations must add up to one, each squared correlation is then $1/K$. From (4) it is clear that the lowest possible SDF volatility arises if the $K$ PCs with non-zero correlation with $\mu_i$ are the first $K$ with the highest eigenvalues. Thus, we have

$$\text{Var}(M) \geq \frac{\mu_m^2}{\sigma_m^2} + \frac{N}{K} \text{Var}(\mu_i) \sum_{k=2}^{K} \frac{1}{\lambda_k}. \quad (5)$$

We now use the principal components extracted from the empirical covariance matrix of our test assets to calculate the bound (5) for different values of $K$.

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 1.2. The squared SR of the Fama-French factors is plotted as the dashed line in the figure for comparison. If expected returns lined up instead equally with the first 10 PCs, the squared SR would reach around 6.

Panel (b) shows a similar analysis for the $5 \times 5$ size-B/M portfolios. Here, too, the counterfactual squared SR increase with $K$. If expected returns lined up equally with the first two PCs (excl. level factor), the squared SR would be approximately equal to the sum of the squared SRs of SMB and HML. However, if expected returns were correlated equally with the first 10 PCs, the squared SR would reach around 4.
Figure 2: Hypothetical Sharpe Ratios if expected returns line up with first $K$ (high-eigenvalue; excl. PC1) principal components.
3.2 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 2 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 mispricings 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 \( \chi^2 \)-test in Table 2 takes the sampling error into account and still rejects the low-dimensional factor models. However, there are additional reasons to suspect that high ex-post SR are not robust indicators of persistent near-arbitrage opportunities. Data-snooping biases can overstate the in-sample SR. Short-lived near-arbitrage opportunities might exist for a while, without being a robust, persistent feature of the cross-section of expected returns.

To shed light on this robustness issue, we perform pseudo-out-of-sample analyses. We split our sample period in two halves, and we treat the first half as our in-sample period, and the second half as our out-of-sample period. We start with a univariate perspective with the 15 anomaly long-short portfolios. Figure 3 plots the in-sample squared SR in the first subperiod on the x-axis and the ratio of out-of-sample to in-sample squared SR on the y-axis. The figure shows that there
In Figure 3, panel (a), we consider all 30 portfolios underlying the 15 long-short strategies jointly. Focusing first on the in-sample period in the first half of the sample, we look at the maximum

---

6In 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.
Figure 4: In-sample and out-of-sample maximum squared Sharpe Ratios (annualized) of first $K$ principal components (incl. level factor). In panels (a) and (b) the sample period is split into two halves. We extract PCs in the first sub-period and calculate SR-maximizing combination of first $K$ PCs in first subperiod. We then apply the portfolio weights implied by this combination in the out-of-sample period (second sub-period). In panels (c) and (d) we randomly sample (without replacement) half of the returns to extract PCs and calculate 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; average squared SRs are shown.
squared SR that can be obtained from a combination of the first $K$ principal components (incl. level factor). The blue solid line in the figure plots the result. With $K = 3$, the maximum squared SR is around 1.2, but raising $K$ further raises the squared SR above 4 for $K = 15$. 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. The red dashed line in the figure shows the result. Not surprisingly, overall SR are lower out of sample. Most importantly, it makes virtually no difference whether one picks $K = 5$ or $K = 15$—the out-of-sample squared SR is about the same and stays mostly around 1. Hence, while the higher-order PCs add substantially to the squared SR in sample, they provide no incremental improvement of the SR in the out-of-sample 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. In panel (b) we repeat the same analysis for the $5 \times 5$ size-B/M portfolios and their PC factors. The results are similar.

In Figure 4, panels (c) and (d), we perform a bootstrap estimation. First, we randomly sample (without replacement) half of the returns to extract PCs and calculate the SR-maximizing combination of the 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; average squared SRs are shown. Panel (c) shows the results for anomaly portfolios. In panel (d) we repeat the same analysis for the $5 \times 5$ size-B/M portfolios and their PC factors. Similar to our findings that used a sample split, we show that the higher-order PCs provide very little incremental improvement of the SR in the out-of-sample period.

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 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.
4 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 $t = 0, 1, 2, \ldots$. There are $N$ stocks in the economy indexed by $i = 1, \ldots, N$. The supply of each stock is normalized to $1/N$ shares. A risk-free bond is available in perfectly elastic supply at an interest rate of $R_F = 0$. 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 $\Gamma$ features a few dominant factors that drive most of the stocks’ covariances. Since prices are constant in this IID case, the covariance matrix of returns equals the covariance matrix of dividends, $\Gamma$. Consider its eigenvalue decomposition $\Gamma = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$. Assume that the first PC is a level factor, with identical constant value for each element of the corresponding eigenvector $q_1 = \sqrt{1/N}$. Then, the variance of returns on the market portfolio is

$$\sigma_m^2 = \text{Var}(R_{m,t+1}) = N^{-2} q'_1 \mathbf{\Lambda}_1 q_1 = N^{-1} \lambda_1.$$ 

All other principal components, by construction, are long-short portfolios, i.e., $q'_k = 0$ 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, i.e.,

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

s.t. $W_{t+1} = (W_t - C_t) + y'R_{t+1}$,

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

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

(6)

Sentiment investors, the second group, are present in measure \( \theta \). 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 \( \delta \). Their risky asset demand vector is

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

(7)

where we assume that \( \delta' \theta = 0 \). 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 \( \delta \), 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 \( \delta' \theta = 0 \).

If \( \delta \) was completely unrestricted, then prices could be arbitrarily strongly distorted even if arbitrageurs are present. Unbounded \( \delta \) 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

\[ \delta' \delta \leq 1. \]  

(8)

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 (8) 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 ± 1/N in
all others.\(^7\) Thus, the constraint still allows sentiment investors to have rather substantial portfolio
tilts, but it prevents the most extreme ones.

Market clearing,
\[
\theta \delta + \frac{1}{a} \Gamma^{-1} E[R_{t+1}] = \frac{1}{N} \mu^t,
\]
implies
\[
E[R_{t+1}] - \mu^m t = -a \theta^t \mu^m,
\]
where \(\mu^m \equiv (1/N) t' E[R_{t+1}]\) and we used the fact that, due to the presence of the level factor, \(t\)
is an eigenvector of \(\Gamma\) and so \(\Gamma^{-1} t = \frac{1}{\lambda} t = \frac{1}{N \sigma^2 m} t\). Moreover, we used \(\mu^m = a \sigma^2 m\). Then, after
substituting into arbitrageurs optimal demand, we get
\[
y = \frac{1}{N} t - \theta \delta.
\]
Consequently, we obtain the SDF,
\[
M_{t+1} = 1 - a (R - E[R])' y
= 1 - a [R_{m,t+1} - \mu^m] + a (R_{t+1} - E[R_{t+1}])' \theta \delta,
\]
\(^7\)In equilibrium, the 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,
\[ \text{Var}(M) = a^2 \sigma_m^2 + a^2 \theta^2 \delta \Gamma \delta. \] (13)

The effect of \( \delta \) on the factor structure and the volatility of the SDF depends on how \( \delta \) lines up with the PCs. To characterize the correlation of \( \delta \) with the PCs, we express \( \delta \) as a linear combination of PCs,
\[ \delta = Q \beta, \] (14)
with \( \beta_1 = 0 \). Note that \( \delta' \delta = \beta' Q' Q \beta = \beta' \beta \) so the constraint (8) can be expressed in terms of \( \beta \):
\[ \beta' \beta \leq 1. \] (15)

### 4.1 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 (14) and (13) to obtain excess SDF variance, expressed, for comparison, as a fraction of the SDF variance accounted for by the market factor,
\[ V(\beta) = \frac{\text{Var}(M) - a^2 \sigma_m^2}{a^2 \sigma_m^2} \]
\[ = \frac{\theta^2}{\sigma_m^2} \delta' \Gamma \delta \]
\[ = \kappa^2 \sum_{k=2}^{N} \beta_k^2 \lambda_k \] (16)

where \( \kappa \equiv \frac{\theta}{\sigma_m} \). From equation (16) we see that SDF excess variance is linear in the eigenvalues of the covariance matrix, with weights \( \beta_k^2 \). For the sentiment-driven demand component \( \delta \) to have a large impact on SDF variance and hence the maximum Sharpe Ratio, the \( \beta_k \) corresponding to high
eigenvalues must have a big absolute value. This means that $\delta$ must line up primarily with the high-eigenvalue (volatile) principal components of asset returns. The constraint (15) implies that if $\beta$ 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 5 illustrates this with data based on the covariance matrix of actual portfolios used as $\Gamma$ 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 1. Returns are in excess of the level factor. We set $\beta$ 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. (16) provides the excess variance of the SDF in each case.

Figure 5 plots the result with $K$ on the horizontal axis. For both sets of portfolios, a substantial SDF excess variance can be achieved only if $\delta$ 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 $\delta$ also lines up with low eigenvalue PCs, then the limited amount of variation in $\delta$ permitted by the constraint (8) is neutralized to a large extent 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.

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 (8)) or
Figure 5: 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 $5 \times 5$ size-B/M portfolios; the red dashed curve is based on 30 anomaly long and short portfolios.

if arbitrage capital was largely absent. None of these two alternatives seems plausible.

4.2 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 3.2, 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 $\delta$ as being associated with certain stock characteristics. For example, elements of $\delta$ 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 (10) and hence cross-sectional variation in expected returns is

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

$$= a^2 \theta^2 \beta' \Lambda^2 \beta. \tag{17}$$

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

$$a^2 \theta^2 \sum_{k=2}^{K} \beta_k^2 \lambda_k^2. \tag{18}$$

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 $\beta$, we can compute the proportion of cross-sectional variation in expected returns explained by the first $K$ principal components, i.e., the ratio of (18) to (17), and the ratio of (the upper bound of) cross-sectional variance in expected returns, (17), to squared expected excess market returns. Depending on the choice of the elements of the $\beta$ 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 $\beta$ subject to the constraint (15). In Figure 6 we plot the right envelope, that is, the maximal cross-sectional expected return variation for a given level of share explained by the first $K$ PCs.\(^8\)

As Figure 6 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 $5 \times 5$ size-B/M portfolios and slightly below 0.60 for the anomaly portfolios (depicted with

\(^8\)Appendix section B provides more details on the construction of Figure 6.
Figure 6: Characteristics vs. covariances: Cross-sectional variation in expected returns explained by first two principal components for $5 \times 5$ 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.

dashed vertical lines on the plot). To achieve these levels of cross-sectional variation in expected returns, virtually all expected return variation has 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.

4.3 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. 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. 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 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 patterns; Johnson (2002) builds a related model based on a reduced-form SDF in a paper with the title “Rational Momentum Effects.”
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 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 $\Pi_{t+1}$ 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

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

The SDF that appears in this objective function is not restricted any further. Hence, the SDF could be, for example, the SDF (12) 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

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

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

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

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

If we let $\Pi_{t+1}$ 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),

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

31
where \( M_{t+1} \) is the SDF (12) that reflects the sentiment investor demand.

Thus, a firm with high \( E[\Pi_{t+1}] \) (relative to other firms) must either have high investment or a strongly negative \( \text{Cov}(M_{t+1}, D_{t+1}) \) (which implies a high expected return). Similarly, 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 profitability 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 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 $\varepsilon_i$ with the SDF. Projects with very negative $\gamma_i = \text{Cov}(\varepsilon_i, M)$ have a high expected return, i.e., a high cost of capital, and are rejected. Ones with less negative $\gamma_i$ are taken on by the firm. Again, it is important to keep in mind that $\gamma_i$ is a covariance with a generic SDF. Other than the existence of such an SDF, nothing has been assumed that would imply that $\gamma_i$ 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 $\gamma_i$ of the still-alive projects, which we denote $\gamma_t$. Since expected excess returns are equal to the negative of the covariance with $M$, it follows that

$$E[R_{t+1}] = f(n_t, \gamma_t) \tag{24}$$

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

\[ E[R_{t+1}] = a_0 + a_1(B_t/MV_t) + a_2(1/M_t), \tag{25} \]

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 \( \gamma_i \) (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 \( \gamma_i \).

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 (25), 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 4). Thus, in terms of investor beliefs and preferences, the BGN model is as much a “behavioral” model as it is a “rational” model.
5 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 4 to the dynamic case when sentiment demand is time-varying. Like in the previous setup, there are $N$ stocks, $i = 1, \ldots, N$, 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 arbitrageurs and sentiment investors consistent with the equilibrium demand in the static model (see (11) and the market clearing condition), but now subject to an IID stochastic shock. For sentiment investors we have

$$x_t = \frac{1}{N} t + (1 - \theta) \delta \xi_t, \quad (26)$$

and for arbitrageurs,

$$y_t = \frac{1}{N} t - \theta \delta \xi_t, \quad (27)$$

where $\xi_{t+1} \sim \mathcal{N}(0, \omega^2)$ is a time-varying component (scalar) of their demand. We assume $\delta$ has a level component and a component orthogonal to the level component. The setup effectively assumes a single factor in sentiment investors demand.

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

$$J_t(W_t, \xi_t) = \max_{(C_s, y_s), s \geq t} \mathbb{E}_t \left[ -\sum_{s=t}^{\infty} \beta^s \exp(-\alpha C_s) \right], \quad (28)$$
where the maximization is subject to

\[
W_{t+1} = (W_t - C_t) R_F + y_t' R_{t+1},
\]

(29)

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

We define the market portfolio as \( R_{M,t+1} = \frac{1}{N} \xi' R_{t+1} \). We guess that prices and the log value function are linear in \( \xi_t \),

\[
P_t = a_0 + a_1 \xi_t \quad (30)
\]

\[
J_t(W_t, \xi_t) = -\beta_t \exp(-\gamma W_t - b_0 - b_1 \xi_t). \quad (31)
\]

In Appendix C we solve for the constants \( a_i \) and \( b_i \) and establish that equilibrium expected returns are given by

\[
E(R_{t+1}) = \gamma \text{Cov}(R_{t+1}, R_{M,t+1} - E_t R_{M,t+1}) + \frac{\gamma}{R_F} \text{Cov}(R_{t+1}, E_{t+1} R_{M,t+2})
\]

Thus, we get an ICAPM similar to Campbell (1993, Eq.23). The degree of presence of sentiment traders does not show up directly, but it is indirectly in \( \text{Cov}(R_{t+1}, E_{t+1}[R_{M,t+2}]) \), because as \( \theta \) goes to zero, this covariance shrinks to zero. Alternatively, note that \( \text{Cov}(D_{t+1}, R_{M,t+1} - E_t[R_{M,t+1}]) = \gamma \Gamma_t \) and so we can write

\[
E[R_{t+1}] = \gamma \text{Cov}(D_{t+1}, R_{M,t+1} - E_t[R_{M,t+1}])
\]

(32)

This is a “bad beta, good beta” specification as in Campbell and Vuolteenaho (2004), but here with a zero risk premium for the “good” beta, i.e., the discount rate beta. The “good” beta disappears because the hedging demand due to time variation in expected returns goes in the opposite direction to the discount rate component of the market return, and exactly cancels out when returns are i.i.d. (so that low returns today lead to an immediate one-to-one increase in expected returns for the next period). Arbitrageurs therefore do not demand a risk premium for discount-rate beta.
exposure, because expected return variation only has transitory effects on their wealth. Only the
cash-flow beta ("bad") beta is compensated with a risk premium.

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.
6 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


Appendix

A Absence of near-arbitrage

This section presents the derivation of the SDF Variance. Define

\[
\omega_m = \frac{1}{\sqrt{N}} q_1
\]
\[
\mu_m = \omega'_m \mu
\]
\[
\sigma^2_m = \omega'_m \Gamma \omega_m
\]

The last definition implies that \( \sigma^2_m = \frac{\lambda_1}{N} \).

Then

\[
\text{Var}(M) = \frac{\mu^2_m}{\sigma^2_m} + (\mu - \mu_m)' Q_z \Lambda_z^{-1} Q'_z (\mu - \mu_m)
\]

Let

\[
\omega_k = \frac{1}{\sqrt{N}} q_k
\]
\[
\mu_k = \omega'_k \mu
\]
\[
\sigma^2_k = \omega'_k \Gamma \omega_k
\]

The last definition implies that \( \sigma^2_k = \frac{\lambda_k}{N} \). \( \sigma^2_k \) 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

\[
\text{Var}(M) = \frac{\mu^2_m}{\sigma^2_m} + \sum_{k=2}^{N} N^2 \frac{\text{Cov}(\mu_i, q_{ki})^2}{\lambda_k}
\]
\[
= \frac{\mu^2_m}{\sigma^2_m} + \text{Var}(\mu_i) \sum_{k=2}^{N} \frac{\text{Corr}(\mu_i, q_{ki})^2}{\sigma^2_k}
\]

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 \( \sigma^2_k \) depends on which of the PCs \( \mu \) lines up with. If it lines up with high \( \sigma^2_k \) PCs then the sum is much lower than if it lines up with low \( \sigma^2_k \) PCs. Thus, if expected returns line up with low-eigenvalue PCs, then we get much higher SR.

B Characteristics vs. covariances

This subsection provides additional detail on the construction of the Figure 6.

Figure 6 plots the right envelope of the set generated of all \( \beta \) that satisfy restriction (15).
To construct this right envelope we put all weight of $\beta$ onto two eigenvectors: the eigenvector associated with the highest eigenvalue (1-st eigenvector) and the $(K+1)$-th eigenvector, i.e. the eigenvector associated with the highest principal component from the remainder of $N-K$ principal components not used in equation (18)). We then vary weights on these two components in a way that satisfies (15). For each set of weights, we compute the ratio of (18) to (17), and the ratio of (the upper bound of) cross-sectional variance in expected returns, (17), to squared expected excess market returns.

C Dynamic Model

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

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

(42)

The model can be easily specialized to the case considered in Section 5 (also Case 2 below) by setting $\mu = \phi = 0$.

Bellman equation is given by

$$J_t(W_t, \xi_t) = \max_{C_t, y_t} \{ -\beta^t \exp(-\alpha C_t) + E_t[J_{t+1}(W_{t+1}, \xi_{t+1})]\}$$  

(43)

Guess

$$P_t = a_0 + a_1 \xi_t$$  

(44)

$$J_t(W_t, \xi_t) = -\beta^t \exp(-\gamma W_t - b_0 - b_1 \xi_t - b_2 \xi_t^2)$$  

(45)

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

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

(46)

and hence

$$E_t[R_{t+1}] = -r_F a_0 - R_F a_1 \xi_t + a_1 \hat{\xi}_t$$  

(47)

$$E_t[W_{t+1}] = (W_t - C_t)R_F - y_t'(r_F a_0 + R_F a_1 \xi_t - a_1 \hat{\xi}_t)$$  

(48)

$$\text{Var}_t(W_{t+1}) = y_t'(\Gamma + a_1 a_1' \omega^2) y_t$$  

(49)

$$\text{Cov}_t(W_{t+1}, \xi_{t+1}) = y_t' a_1 \omega^2$$  

(50)

$$E_t(\xi_{t+1}^2) = \omega^2 + \hat{\xi}_t^2$$  

(51)

$$\text{Cov}_t(\xi_{t+1}, \xi_{t+1}^2) = 2 \omega^2 \hat{\xi}_t$$  

(52)

$$\text{Cov}_t(W_{t+1}, \xi_{t+1}^2) = 2 \omega^2 \hat{\xi}_t y_t' a_1$$  

(53)

$$\text{Var}_t(\xi_{t+1}^2) = 2 \omega^4 + 4 \hat{\xi}_t^2 \omega^2$$  

(54)
where we used the following derivations:

\[
E_t \left( \xi_{t+1}^2 \right) = \text{Var}_t[\xi_{t+1}] + [E_t(\xi_{t+1})]^2 = \omega^2 + \hat{\xi}_t^2
\]

\[
\text{Cov}_t(\xi_{t+1}, \xi_{t+1}^2) = E_t \left[ (\xi_{t+1} - \hat{\xi}_t) \left( \xi_{t+1}^2 - \hat{\xi}_t^2 - \omega^2 \right) \right]
\]

\[
= E_t \left[ (\xi_{t+1} - \hat{\xi}_t)^2 \right] + E_t \left[ (\xi_{t+1} - \hat{\xi}_t) \left( 2\xi_{t+1}\hat{\xi}_t - 2\hat{\xi}_t^2 - \omega^2 \right) \right]
\]

\[
= 2\omega^2\hat{\xi}_t
\]

\[
\text{Var}_t(\xi_{t+1}^2) = \text{Var}_t \left[ (\xi_{t+1} - \hat{\xi}_t)^2 + 2\hat{\xi}_t (\xi_{t+1} - \hat{\xi}_t) \right]
\]

\[
= \omega^4 \text{Var}_t \left[ \left( \frac{\xi_{t+1} - \hat{\xi}_t}{\omega} \right)^2 \right] + 4\hat{\xi}_t^2 \omega^2
\]

\[
= 2\omega^4 + 4\hat{\xi}_t^2 \omega^2
\]

where the second line follows because the third moment of a mean zero normally-distributed random variable is zero, and the last line follows as the variance of \(\chi^2(1)\) distribution.

Then

\[
E_t \left[ J_{t+1}(W_{t+1}, \xi_{t+1}) \right] = -\beta^{t+1} \exp \left( -\gamma E_t[W_{t+1}] - b_0 - b_1\hat{\xi}_t - b_2 \left( \omega^2 + \hat{\xi}_t^2 \right) \right)
\]

\[
+ \frac{1}{2} \gamma^2 \text{Var}_t(W_{t+1}) + \frac{1}{2} b_1^2 \omega^2 + \frac{1}{2} b_2^2 \text{Var}_t(\xi_{t+1}^2)
\]

\[
+ \gamma b_1 \text{Cov}_t(W_{t+1}, \xi_{t+1}) + \gamma b_2 \text{Cov}_t(W_{t+1}, \xi_{t+1}^2)
\]

\[
+ b_1 b_2 \text{Cov}_t(\xi_{t+1}, \xi_{t+1}^2) \tag{55}
\]

First-order condition for \(y_t\),

\[
0 = \gamma (r_F a_0 + R_F a_1 \xi_t - a_1\hat{\xi}_t) + \gamma^2 (\Gamma + a_1 a_1' \omega^2) y_t + \gamma b_1 a_1 \omega^2 + 2\gamma b_2 a_1 \omega^2 \hat{\xi}_t \tag{56}
\]

which we can solve for

\[
y_t = -\frac{1}{\gamma} \left( \Gamma + a_1 a_1' \omega^2 \right)^{-1} \left( r_F a_0 + R_F a_1 \xi_t - a_1\hat{\xi}_t + b_1 a_1 \omega^2 + 2 b_2 a_1 \omega^2 \hat{\xi}_t \right) \tag{57}
\]

Plugging this solution into the market clearing condition, and rearranging, we get

\[
-\gamma (\Gamma + a_1 a_1' \omega^2) \left( \frac{1}{N} \tau - \theta \delta \xi_t \right) - b_1 a_1 \omega^2 - 2 b_2 a_1 \omega^2 \hat{\xi}_t = r_F a_0 + R_F a_1 \xi_t - a_1\hat{\xi}_t \tag{58}
\]

Since the MC has to hold for any value of \(\xi_t\), we can apply the method of undetermined coefficients.
and get

\[ a_0 = -\frac{\gamma}{r_F}(\Gamma + a_1a'_1\omega^2) - \frac{b_1}{r_F}a_1\omega^2 - 2\frac{b_2}{r_F}a_1\omega^2\mu + \frac{1}{r_F}a_1\mu \]  
\[ a_1 = \frac{\gamma}{R_F - \phi + 2b_2\phi\omega^2}(\Gamma + a_1a'_1\omega^2)\theta\delta \]  

From the latter equation, we obtain

\[ a_1 \left( 1 - a'_1\delta\omega^2\theta \frac{\gamma}{r_F - \phi + 2b_2\phi\omega^2} \right) = \frac{\gamma}{R_F - \phi + 2b_2\phi\omega^2} \Gamma \delta\theta \]  

Pre-multiplying with \( \delta \) we get a quadratic equation in \( a'_1 \delta \),

\[ c_1(a'_1\delta)^2 - (a'_1\delta) + c_2 = 0 \]  

which we can solve for the positive solution (that we will not need below, though).

We can rewrite (57) as

\[
\left( y_t'(\Gamma + a_1a'_1\omega^2)\right) y_t + y_t' \left( r_Fa_0 + R_Fa_1\xi_t - a_1\xi_t \right) + b_1 \left( y_t'a_1\omega^2 \right) + b_2 \left( 2\omega^2\xi_t y_t'a_1 \right) = 0 \]  

Substituting in variance and covariance from (49) and (50), multiplying through by \( \gamma \), and subtracting one-half the variance term on both sides,

\[
\frac{1}{2}\gamma^2\text{Var}_t(W_{t+1}) + \gamma y_t'(r_Fa_0 + R_Fa_1\xi_t - a_1\xi_t) + \gamma b_1\text{Cov}_t(W_{t+1},\xi_{t+1}) + \gamma b_2\text{Cov}_t(W_{t+1},\xi_{t+1}^2) = \frac{1}{2}\gamma^2\text{Var}_t(W_{t+1}) \gamma^2 = \frac{1}{2}\left( \frac{1}{N}t - \theta\delta\xi_t \right)'(\Gamma + a_1a'_1\omega^2)(\frac{1}{N}t - \theta\delta\xi_t)\gamma^2 \]  

We can write the right-hand side as

\[ \lambda_0 + \lambda_1\xi_t + \lambda_2\xi_t^2 \]  

where

\[ \lambda_1 = \frac{1}{N}t' (\Gamma + a_1a'_1\omega^2) \delta\theta\gamma^2 \]  
\[ \lambda_2 = \theta^2\delta' (\Gamma + a_1a'_1\omega^2) \delta \]  

Now, going back to (55), we can write

\[
\text{Et}[J_{t+1}(W_{t+1},\xi_{t+1})] = -\beta^{t+1} \exp \left( -\gamma(W_t - C_t)R_F + \lambda_0 + \lambda_1\xi_t + \lambda_2\xi_t^2 \right) \]
where
\[
\begin{align*}
\dot{\lambda}_0 &= \lambda_0 - b_0 - b_1 \mu - b_2 (\omega^2 + \mu^2) + \frac{1}{2} b_1^2 \omega^2 + b_2^2 \omega^2 (\omega^2 + 2 \mu^2) + 2 b_1 b_2 \omega^2 \mu \\
\dot{\lambda}_1 &= \lambda_1 - b_1 \phi - 2 b_2 \mu \phi + 4 b_2^2 \mu \phi \omega^2 + 2 b_1 b_2 \phi \omega^2 \\
\dot{\lambda}_2 &= \lambda_2 - b_2 \phi^2 + 2 b_2^2 \omega^2 \phi^2.
\end{align*}
\]

Now we evaluate the first-order condition for consumption
\[
U'(C_t) = \frac{\partial E_t [J_{t+1}(W_{t+1}, \xi_{t+1})]}{\partial C_t}
\]
(68)

After taking logs,
\[
\log \alpha - \alpha C_t = \log(\beta \gamma R_F) - \gamma R_F (W_t - C_t) + \dot{\lambda}_0 + \dot{\lambda}_1 \xi_t + \dot{\lambda}_2 \xi_t^2
\]
which we can solve for
\[
C_t = \frac{\gamma R_F}{\alpha + \gamma R_F} W_t + \frac{1}{\alpha + \gamma R_F} \left[ \log \left( \frac{\alpha}{\beta \gamma R_F} \right) - \dot{\lambda}_0 - \dot{\lambda}_1 \xi_t - \dot{\lambda}_2 \xi_t^2 \right]
\]

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
\[
\exp(-\gamma W_t - b_0 - b_1 \xi_t - b_2 \xi_t^2)
= \exp(-\alpha C_t) + \beta \exp[-(W_t - C_t) R_F + \dot{\lambda}_0 + \dot{\lambda}_1 \xi_t + \dot{\lambda}_2 \xi_t^2]
\]
\[
= \exp \left( - \frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t \right) \exp \left\{ - \frac{\alpha}{\alpha + \gamma R_F} \log \left( \frac{\alpha}{\beta \gamma R_F} \right) \right\} \exp \left\{ \frac{\alpha}{\alpha + \gamma R_F} \left( \dot{\lambda}_0 + \dot{\lambda}_1 \xi_t + \dot{\lambda}_2 \xi_t^2 \right) \right\}
+ \beta \exp \left( - \frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t \right) \exp \left\{ \frac{\gamma R_F}{\alpha + \gamma R_F} \log \left( \frac{\alpha}{\beta \gamma R_F} \right) \right\} \exp \left\{ \frac{\alpha}{\alpha + \gamma R_F} \left( \dot{\lambda}_0 + \dot{\lambda}_1 \xi_t + \dot{\lambda}_2 \xi_t^2 \right) \right\}
\]
\[
= \left( \frac{\alpha}{\beta \gamma R_F} \right)^{-\frac{\alpha + \gamma R_F}{\alpha + \gamma R_F}} \left( 1 + \beta \left( \frac{\alpha}{\beta \gamma R_F} \right) \right) \exp \left\{ \frac{\alpha}{\alpha + \gamma R_F} \dot{\lambda}_0 \right\}
\times \exp \left( - \frac{\alpha \gamma R_F}{\alpha + \gamma R_F} W_t \right) \exp \left\{ \frac{\alpha}{\alpha + \gamma R_F} \left( \dot{\lambda}_1 \xi_t + \dot{\lambda}_2 \xi_t^2 \right) \right\}
\]
(69)

Comparing coefficients, we get
\[
\gamma = \frac{\alpha \gamma R_F}{\alpha + \gamma R_F}, \quad \text{i.e.,} \quad \gamma = \frac{\gamma R_F}{R_F}
\]
(70)
\[
b_1 = - \frac{\alpha}{\alpha + \gamma R_F} \dot{\lambda}_1, \quad \text{i.e.,} \quad b_1 = - \frac{\dot{\lambda}_1}{R_F}
\]
(71)
\[
b_2 = - \frac{\alpha}{\alpha + \gamma R_F} \dot{\lambda}_2, \quad \text{i.e.,} \quad b_2 = - \frac{\dot{\lambda}_2}{R_F}
\]
(72)

and one can solve along similar lines for $b_0$.  

46
**Asset pricing.** Having solved for these coefficients, we can now look at asset pricing. Rewrite equation (63) as

$$E_t(R_{t+1}) = \gamma (\Gamma + a_1 a'_1 \omega^2) y_t + \omega^2 b_1 a_1 + 2 \omega^2 \xi_t b_2 a_1 \quad (73)$$

**Case 1.** \(\xi_t\) is a non-zero constant, \(\hat{\xi}_t = \mu\). It follows that \(\phi = 0\), \(a_1 = \frac{\gamma}{R_F} (\Gamma + a_1 a'_1 \omega^2) \theta \delta\), \(b_1 = -\frac{1}{R_F} \lambda_1 = -\gamma \frac{\omega}{1} a_1' a_1\), and \(b_2 = -\frac{1}{\theta} \theta' a_1\). Furthermore, \(\text{Cov}(R_{t+1}, E_{t+1}[R_{M,t+2}]) = -a_1 a'_1 \frac{1}{\omega} R_F \omega^2\), \(\text{Cov}(R_{t+1}, E_{t+1}[R_{\delta,t+2}]) = -a_1 a'_1 \delta R_F \omega^2\) and so (73) yields:

$$E_t(R_{t+1}) = \gamma \text{Cov}_t(R_{t+1}, R_{A,t+1}) + \frac{\gamma}{R_F} \text{Cov}_t(R_{t+1}, E_{t+1} R_{M,t+2}) + \frac{2 \mu}{R_F} \text{Cov}_t(R_{t+1}, E_{t+1} R_{\delta,t+2}) \quad (74)$$

where \(\gamma = a_1 \frac{R_F}{R_F}\), \(\text{Cov}_t(R_{t+1}, R_{A,t+1}) = \text{Cov}_t(R_{t+1}, R_{M,t+1}) - \xi_t \theta \text{Cov}_t(R_{t+1}, R_{\delta,t+1})\), \(R_A\) is the return on arbitrageur’s investment portfolio, and \(R_\delta\) is the return on long-short portfolio driven by demands of sentiment investors.

We can rewrite (74) with two terms only:

$$E_t(R_{t+1}) = \gamma \text{Cov}_t(R_{t+1}, R_{A,t+1}) + \psi \text{Cov}_t(R_{t+1}, \xi_{t+1})$$

$$= \gamma \text{Cov}_t(R_{t+1}, R_{A,t+1}) + \hat{\psi} \text{Cov}_t(R_{t+1}, E_{t+1} R_{M,t+2})$$

where \(\psi = \omega^2 (b_1 + 2 \mu^2 b_2), \hat{\psi} = -\frac{\psi}{\frac{1}{\theta} \theta' a_1 R_F \omega^2}\).

Taking expectations of both sides gives

$$E(R_{t+1}) = \gamma \text{Cov}(R_{t+1}, R_{A,t+1} - E_t R_{A,t+1}) + \psi \text{Cov}(R_{t+1}, \xi_{t+1})$$

$$= \gamma \text{Cov}(R_{t+1}, R_{A,t+1} - E_t R_{A,t+1}) + \hat{\psi} \text{Cov}(R_{t+1}, E_{t+1} R_{M,t+2}) \quad (75)$$

where

$$\text{Cov}(R_{t+1}, R_{A,t+1} - E_t R_{A,t+1}) = \text{Cov}(R_{t+1}, R_{M,t+1} - E_t R_{M,t+1})$$

$$- \mu \theta \text{Cov}(R_{t+1}, R_{\delta,t+1} - E_t R_{\delta,t+1})$$

Equation (75) is convenient for empirical estimation given that we have an empirical proxy for the sentiment investor flow vector \(\xi_t\).

**Case 2.** \(\xi_t\) is zero, \(\hat{\xi}_t = 0\). It follows that \(\mu = \phi = 0\),

$$E(R_{t+1}) = \gamma \text{Cov}(R_{t+1}, R_{M,t+1} - E_t R_{M,t+1}) + \frac{\gamma}{R_F} \text{Cov}(R_{t+1}, E_{t+1} R_{M,t+2})$$

Thus, we get an ICAPM similar to Campbell (1993, equation (23)). The degree of presence of sentiment traders does not show up directly, but it is indirectly in \(\text{Cov}(R_{t+1}, E_{t+1}[R_{M,t+2}])\), because as \(\theta\) goes to zero, this covariance shrinks to zero. Alternatively, note that \(\text{Cov}(D_{t+1}, \cdots -
\( E_t[R_{M,t+1}] = \gamma \Gamma \frac{1}{N} t \) and so we can write
\[
E[R_{t+1}] = \gamma \text{Cov}(D_{t+1}, R_{M,t+1} - E_t[R_{M,t+1}])
\]
(76)

This is a bad beta, good beta specification as in Campbell and Vuolteenaho (2004), but here with a zero risk premium for the “good” beta, i.e., the discount rate beta.

**Case 3.** \( \xi_t \) is AR(1), \( \xi_t = \mu + \phi \xi_t \). Similarly to Case 1, we can derive the following equation
\[
E_t \left( R_{t+1} \right) = \gamma \text{Cov}_t \left( R_{t+1}, R_{A,t+1} \right) + \psi_t \text{Cov}_t \left( R_{t+1}, \xi_{t+1} \right)
\]
\[
= \gamma \text{Cov}_t \left( R_{t+1}, R_{A,t+1} \right) + \hat{\psi}_t \text{Cov}_t \left( R_{t+1}, E_{t+1}R_{M,t+2} \right)
\]
where \( \psi_t = \omega^2 \left( b_1 + 2\xi_t^2 b_2 \right) \), \( \hat{\psi}_t = -\frac{\psi_t}{\gamma} \). Note that in this case the price of discount rate risk is time-varying. Unconditionally we get:
\[
E \left( R_{t+1} \right) = \gamma \text{Cov} \left( R_{t+1}, R_{A,t+1} - E_t R_{A,t+1} \right) + \tilde{\psi} \text{Cov} \left( R_{t+1}, \tilde{\xi}_{t+1} - E_t \tilde{\xi}_{t+1} \right)
\]
\[
= \gamma \text{Cov} \left( R_{t+1}, R_{A,t+1} - E_t R_{A,t+1} \right) + \tilde{\psi} \text{Cov} \left( R_{t+1}, E_{t+1}R_{M,t+2} - E_t R_{M,t+2} \right)
\]
where \( \tilde{\psi} = \omega^2 \left( b_1 + 2b_2E \left[ \xi_t^2 \right] \right) \), \( \tilde{\psi} = -\frac{\tilde{\psi}}{\gamma} \).