The Economics of Factor Timing*

Valentin Haddad
UCLA & NBER

Serhiy Kozak
University of Michigan

Shrihari Santosh
University of Maryland

June 17, 2018

Abstract

An optimal factor timing portfolio is equivalent to a conditional SDF. We use economic restrictions to determine and characterize both empirically. With these restrictions, we find that long-short equity factors are strongly and robustly predictable. A number of these portfolios have small or zero average price of risk, suggesting that the economic risks investors worry about conditionally are often very different from those they worry about on average. This manifests in the very different compositions of the conditional and unconditional SDFs. For bonds and foreign exchange strategies, long-short portfolios sorted on maturity or interest rate differential are also predictable.

*Haddad: valentin.haddad@anderson.ucla.edu; Kozak: sekozak@umich.edu; Santosh: shrihari@umd.edu.

This paper was previously circulated under the title "Predicting Relative Returns." We thank Mikhail Chernov, John Cochrane, Julien Cujean, Robert Dittmar, Stefano Giglio, Bryan Kelly, Ralph Koijen, Hugues Langlois, Lars Lochstoer, Mark Loewenstein, Tyler Muir, Stefan Nagel, Nikolai Roussanov, Avanidhar Subrahmanyan, Michael Weber and seminar participants at AFA, FIRS, Maryland, Michigan for helpful comments and suggestions.

1 Introduction

Aggregate stock returns are predictable over time (e.g., Shiller 1981, Fama and French 1988), creating the scope for investors to engage in market timing. Factors beyond the aggregate market are sources of risk premia in the cross-section of assets (e.g., Fama and French 1993), creating the basis for factor investing.[1] How can we combine these two ideas and construct the optimal *factor timing* portfolio, which unifies cross-sectional and time-series predictability of returns? Determining this portfolio empirically appears difficult because of the high dimensionality of the space of dynamic strategies. A crucial observation, however, is that this problem is equivalent to a fundamental economic question: what is the *conditional* stochastic discount factor?[2] This equivalence has powerful implications, which we exploit: we use economic restrictions on stochastic discount factors (SDFs) to discipline the construction of the optimal factor timing portfolio in the data. We then use the estimated factor timing portfolio to characterize the properties of the conditional SDF.

We discover significant and robust benefits to factor timing, benefits which are elusive when using purely statistical methods. A pure factor timing portfolio with no average exposure to the factors earns a Sharpe ratio of 0.85, with most of this performance coming from timing risks other than the market.[3] Time variation in the compensation for factor risk is an important ingredient of both the conditional SDF and the optimal factor timing portfolio. The compositions of the conditional and unconditional SDFs are often substantially different, and the former is much more volatile.

Our starting point is the cross-section of fifty standard stock "anomaly" portfolios that have been put forward in previous work as capturing cross-sectional variation in expected returns. We rely on two economic restrictions. First, we assume that a factor model for stocks holds with respect to these portfolios, or a subset of them.[4] This is the basis of factor timing: if factors determine expected returns, they should be the building blocks of trading strategies. Second, we assume that the pricing kernel implied by this factor model does not generate excessively high *conditional* Sharpe ratios. Such near-arbitrage opportunities would be eliminated by arbitrageurs in equilibrium (Kozak et al., 2017). The assumption implies that the time series variation in factor risk premia is mostly driven by time-varying exposure of the SDF to the largest sources of variation in realized factor returns. If this

[1] This is also commonly referred to as "smart beta" investing.

[2] Hansen and Jagannathan (1991) show that the minimum variance SDF is affine in the return on the maximum Sharpe ratio portfolio.

[3] A pure timing portfolio is a portfolio with a zero average weight in all risky assets.

[4] This assumption forms, for instance, the essence of the methodology proposed by Brandt et al. (2009).

were not the case, small components would be highly predictable, generating implausibly large *conditional* Sharpe ratios. We measure the largest sources of variation by the first two principal components (PCs) of returns, which explain more than a third of the variation in realized returns. We study their predictability in addition to that of the market portfolio, using the net book-to-market ratio of these portfolios as a simple measure to predict their returns.

These two PC portfolios of anomalies are very strongly predictable. Their own book-to-market ratios predicts their future annual returns with an out-of-sample R^2 of more than 20%, about twice as large as that of predicting the market returns. The risk premium variation for the two components is only weakly related to that of the market; more than 75% of their variation is orthogonal to changes in market risk premium. This predictability of the dominant PC portfolios captures common variation in the risk premia and allows us to predict returns of individual anomaly strategies with a median out-of-sample R^2around 9%. The observation that factor returns are robustly predictable lends support to the enterprise of factor timing. It also sheds new light on the economic importance of some of these anomaly strategies. Interestingly, even strategies with relatively weak unconditional performance, such as size or leverage, exhibit substantial time-series predictability. Similarly, our first principal component portfolio commands a small unconditional Sharpe ratio, but exhibits large swings in expected returns. Both pieces of evidence suggest that the economic risks investors worry about conditionally are often very different from those they worry about on average.

Our results contrast with the weaker out-of-sample predictability obtained when using approaches not disciplined by our economic restrictions. For instance one could separately estimate expected returns for each portfolio without recognizing the factor structure in returns and imposing the absence of near-arbitrage. We find this approach not to be fruitful: most of our individual portfolios are not predictable out-of-sample by their own book-to-market ratio or characteristic spread. Some purely statistical discipline—e.g., assuming that their own book-to-market ratio predicts each anomaly in the same way—helps, but generates less than half of the predictability our economic exercise uncovers. The out-of-sample robustness of our approach supports the validity of our restrictions.

We use our results to characterize the properties of the optimal factor timing portfolio under our economic assumptions and quantify the investment benefits to factor timing. More importantly, we also compare the properties of the conditional SDF with its unconditional counterpart which ignores predictability. Timing expected returns provides substantial investment gains; a pure timing portfolio achieves an average conditional Sharpe ratio of 0.85. This implies that the conditional variance of the SDF is substantially larger than that implied by unconditional measures. Most of the benefits of timing accrue from timing the anomaly factors rather than the aggregate market return. In addition, the composition of risks affecting the SDF changes strongly over time, at about business-cycle frequency. In other words, changes over time in the pricing of risks are not driven by one common source of variation, in contrast to many theories.[5] Rather, the SDF has independently varying loadings on a few weakly correlated sources of risks.

Finally, we assess the external validity of our approach by following its guidance to study two other asset classes: Treasury bonds and foreign exchange. The returns in these asset classes exhibits a very strong factor structure, much more so than stocks. For each of them, two PC factors explain more than 90% of the variance in realized returns, so we focus on the predictability of these components. We use forward spreads, the standard counterpart to valuation ratios in these markets, as predictors. Just like for stocks, we find that there is substantial predictability, but that this predictability is concentrated in the second principal component, an index-neutral long-short strategy. For bonds, a duration-neutral portfolio of long minus short maturity bonds is strongly predictable, more so than the average bond return often studied in the previous literature. For currencies, a relative carry portfolio is more predictable than the dollar-carry strategy.

To summarize, factor timing is very valuable, above and beyond market timing and factor investing taken separately. The changing conditional properties of the pricing kernel are mostly driven by market-neutral factors. To robustly reach these quantitative conclusions, it is crucial to acknowledge the economic nature of factor timing. The methods and facts we study in this paper are only the beginning of the economic enterprise of understanding the evolution of drivers of risk premia. They constitute a set of facts that future theories, which will name these sources of variation, will have to reflect.

The remainder of the paper is organized as follows. After a brief review of the related literature, Section 2 explains our methodology. Section 3 constructs the PC portfolios and documents their predictability. Section 4 considers the implications of this result for the predictability of individual anomalies. Section 5 considers the implications for the optimal timing portfolio and the properties of the conditional SDF. Finally, Section 6 extends our analysis to Treasuries and currencies.

[5] See for instance Campbell and Kyle (1993), Campbell and Cochrane (1999), Bansal and Yaron (2004), Barberis et al. (2015).

Related literature

This paper builds on the long literature which studies time series predictability of returns, starting from Shiller (1981) and Fama and French (1988) for stocks, or Fama and Bliss (1987) for bonds.[6] While the early evidence is mostly about aggregate returns, our main focus is on understanding predictability of cross-sections of returns. Various papers, such as Cochrane and Piazzesi (2005), Cieslak and Povala (2015) for bonds and Stambaugh et al. (2012), Akbas et al. (2015) for stocks, examine the ability of a single variable to forecast all returns in an asset class. Hence, these papers implicitly or explicitly assume a single source of time-varying risk premia.

A separate literature studies assets (or individual anomaly strategy returns) in isolation, forecasting each asset's return with asset specific predictors. Recent prominent examples are Campbell et al. (2009) and Lochstoer and Tetlock (2016), who use a bottom-up approach of aggregating firm-level estimates into portfolios in order to decompose variation in returns into discount rate and cash-flow news. Asness et al. (2000), Cohen et al. (2003), Arnott et al. (2016a,b) and others use valuation ratios to forecast anomaly returns. Greenwood and Hanson (2012) forecast characteristics based anomalies using their "issuer-purchaser" spread—the difference in the average characteristic for net equity issuers vs repurchasers. Their analysis, like the papers which use valuation ratios, is fundamentally at the asset (or individual anomaly) level, ignoring the potential correlation across anomaly strategy returns. Implicitly, these papers assume there are as many independent sources of time-varying risk premia as there are assets.

Our approach is, in a sense, an intermediate view between these two extremes. To focus directly on common variation in risk premia, we use a more top-down methodology. We directly measure the predictability of a few substantial components of returns with a few predictors and then project the forecasts back onto individual assets. The intuition underlying our method is that if two assets have highly correlated returns, they should also have highly correlated expected returns. Otherwise, it would be possible to achieve very high Sharpe ratios by exploiting the differential predictability of the two assets. We do, however, allow for "a few" sources of time-varying risk premia. Essentially, we impose a conditional version of the idea of no near-arbitrage. (Ross, 1976, Kozak et al., 2017, 2018)

The paper closest to ours is Brooks and Moskowitz (2017), which studies the predictability of bond returns across countries, summarizing them by factors similar to principal components. Finally, predictability arising from a few factors is also sometimes a part of rich

[6] See Koijen and Van Nieuwerburgh (2011) for a survey of recent work on the topic.

fully-specified models of SDFs. Dynamic term structure models—e.g., Joslin et al. (2014)—or dynamic asset pricing models—e.g., Adrian et al. (2015)—can include and estimate this predictability. We focus on a more reduced-form approach highlighting how economic restrictions discipline factor timing. In this way, our paper is similar in spirit to prior work which imposes economically motivated constraints to generate robust predictability. Campbell and Thompson (2007) use equilibrium and present value constraints to improve out-of-sample predictability of the aggregate stock market. Closely related to our framework, Kozak et al. (2017, 2018) use no near-arbitrage arguments to reduce the dimensionality of the set of factors in an unconditional setting. We use similar ideas to obtain robust predictability of the dynamics, rather than the unconditional levels, of factor risk premia.

There is another related literature which argues that cross-sectional long-short factors are not very predictable. Ilmanen and Nielsen (2015) argue that though valuation ratios (such as book-to-market) are useful for predicting asset class index returns, "this relationship turns out to be weaker for long/short factor portfolios." Asness (2016) claims that predicting anomaly strategy returns using valuation ratios is fruitless since the resulting timing strategies "aren't exceptionally strong, and they are way too correlated with the basic value strategy itself to make great impact on a portfolio." Asness et al. (2017) use book-to-market ratios to construct market timing strategies for the size, value and betting-against-beta strategies and conclude "there is near-unanimous support in the literature for the efficacy of value investing, and thus the value (or value spread) of anything should have some predictive power for that thing. The question is the strength of this relationship. Here we find that strength rather lacking." In contrast to these papers, we find that valuation ratios are extremely powerful for predicting long-short factor portfolios, once we focus on the common variation in expected returns uncovered through principal components analysis.

2 Methodology

We are interested in measuring the dynamics of risk premia for a cross-section of excess returns $\{R_{i,t}\}_{i \in I}$; our main empirical setting is the cross-section of stock returns. Studying these dynamics is important for the purpose of optimal portfolio choice, but also to understand the economic forces shaping equilibrium prices. Without any structure, it is challenging to create robust forecasts for all returns. We show how two plausible economic restrictions help address this issue.

First, we assume that assets are conditionally priced by a factor model, the main motivation behind factor timing portfolio strategies. Second, we assume that prices feature no near-arbitrage opportunities. These assumptions imply that measuring the predictability of the largest principal components of the set of factors is enough to characterize expected returns.[7] This strong dimension reduction allows us to use the standard tools for forecasting single return series to measure this predictability.

2.1 Factor Model, Factor Investing, and Factor Timing

Our first assumption is that all excess returns are conditionally priced by a finite set of traded factors $\{F_{j,t+1}\}_{j \in J}$, which we compile in a vector F_{t+1}.

Assumption 1. *(Conditional Factor Model) A conditional factor model holds:*

$$\mathbb{E}_t \left[R_{j,t+1} \right] = \beta_{jt}' \mathbb{E}_t \left[F_{t+1} \right]. \tag{1}$$

Factor models naturally emerge from economic models. In the ICAPM of Merton (1973), the factors capture the market return as well as sources of risks investors want to hedge. In the APT of Ross (1976), the factors are, rather, the sources of common variation in asset returns. More generally, factor models produce a useful way to reduce the large dimension of asset families to a smaller stable family of sources of priced risk.

Under the factor model, the conditional risk premium of each asset is characterized by the product of its conditional exposures to the factors, β_{jt}, and the conditional risk premia on the factors $\mathbb{E}_t \left[F_{t+1} \right]$. The first one, a conditional covariance of realized returns has been extensively studied; we focus on characterizing the conditional risk premia of the factors.

Another reason to focus solely on these premia is that factor models motivate factor investing, a portfolio construction using only the factors. Because the factors completely capture the sources of risks of concern to investors, optimal portfolios can be constructed from only these few factors— the so-called mutual fund theorem. Factor timing strategies are the dynamic counterpart of this observation; as the properties of the factors change, an investor should adjust her portfolio weights accordingly. For example, the maximum conditional Sharpe ratio return is obtained by:

$$R_{t+1}^{opt} = \mathbb{E}_t \left[F_{t+1} \right]' \Sigma_{F,t}^{-1} F_{t+1}, \tag{2}$$

[7]Specifically, the expected return on any asset is the product of the asset's conditional loading on, and the expected return of these few components.

where $\Sigma_{F,t}$ is the conditional covariance matrix of the factors. Knowledge of the conditional risk premia of the factors is crucial to form this and other dynamic strategies.

An equivalent formulation of the factor model is through an SDF, a central object of interest in economic models.

Corollary 1. *Under Assumption 1, the excess returns are conditionally priced by an SDF which is linear in the factors:*

$$0 = \mathbb{E}_t \left[m_{t+1} R_{j,t+1} \right], \tag{3}$$

$$m_{t+1} = a_t - \mathbb{E}_t \left[F_{t+1} \right]' \Sigma_{F,t}^{-1} F_{t+1}. \tag{4}$$

Interestingly, one can recognize that the loadings of the SDF on the factors correspond exactly to the weights of the optimal factor timing portfolio, a result from Hansen and Jagannathan (1991). This equivalence implies that reasonable economic restrictions on the SDF will also discipline factor timing strategies.

While going from individual assets to factors provides some useful dimension reduction and stabilization of covariance structure, we are still left with many return series to forecast.[8] A multitude of empirical results and theoretical motivations have put forward a large number of potential factors, leading to the emergence of what Cochrane (2011) calls the "factor zoo". Including potentially irrelevant factors does not affect the theoretical performance of a factor model; the SDF would just have no loading on these factors. However, including too many factors leads to greater probability of estimating spurious return predictability in finite samples. We now turn to a second assumption which helps discipline our empirical analysis.

2.2 Absence of Near-Arbitrage

We further impose the absence of near-arbitrage.

Assumption 2. *(Absence of near-arbitrage) There are no near-arbitrage opportunities: conditional Sharpe ratios are bounded above by a constant.*

This absence of "good deals" was for instance proposed by Cochrane and Saa-Requejo (2000). Ross (1976) also proposes a bound on the squared Sharpe ratio for an empirical implementation of his APT in a finite-asset economy. He suggests ruling out squared Sharpe

[8] This stabilization role is discussed for example in Brandt et al. (2009), Cochrane (2011), Kozak et al. (2017).

ratios greater than two times the squared Sharpe ratio of the market portfolio. Such a bound on the maximum squared Sharpe ratio is immediately equivalent to an upper bound on the variance of the SDF, m_{t+1}, that resides in the span of excess returns (Hansen and Jagannathan, 1991). This assumption holds in most commonly used completely specified equilibrium models of prices and therefore can be imposed at no cost when evaluating them. Kozak et al. (2017) show such a bound arises even in "behavioral" models due to the presence of arbitrageurs, whose marginal utility determines the SDF.

Under this assumption, Kozak et al. (2017) show that the SDF must load primarily on the dominant factors driving variation in the realized factors F_{t+1}. We define Z_{t+1} as the vector of the largest principal components of F_{t+1}.

Proposition 1. *(Kozak et al., 2017) Under Assumptions 1 and 2, the SDF can be approximated by a combination of the dominant factors:*

$$m_{t+1} \approx a_t - \mathbb{E}_t\left[Z_{t+1}\right]' \Sigma_{Z,t}^{-1} Z_{t+1}. \tag{5}$$

Equivalently, the maximum Sharpe ratio factor timing portfolio can be approximated by

$$R_{t+1}^{opt} = \mathbb{E}_t\left[Z_{t+1}\right]' \Sigma_{Z,t}^{-1} Z_{t+1}. \tag{6}$$

Intuitively, while it is entirely possible that each of the many proposed factors commands a risk premium, it is unlikely they all capture independent sources of risk. Otherwise, investors would be able to diversify across them and obtain implausibly large Sharpe ratios. Using similar arguments, Giglio and Xiu (2017) impose a low-dimensional factor model and use panel asymptotics to identify the SDF. Kelly et al. (2017) also focus on dominant components of the universe of characteristic-managed portfolios. We complement this work by studying how such an approach helps in studying *conditional* properties of expected returns.

Our two assumptions are complementary in the following sense. Assumption 1 delivers the conclusion that factor timing is sufficient; one need not time individual stocks. Even without this assumption, Assumption 2 provides a useful way to time a given set of factors. However, the two together allows us to measure properties of the SDF, shedding light on a fundamental economic quantity.

In practice, the second step achieves much further dimension reduction. As we will show in our empirical applications, no more than a few principal components explain a sizable fraction of the variation in the factor returns. In addition, because factors are often chosen to offer a stable correlation structure, the extraction of dominant components is readily implementable using the standard unconditional method.[9] We are left with estimating the conditional means and variances of these large principal components. In this paper, we concentrate on estimating the mean forecasts $\mathbb{E}_t[Z_{t+1}]$, that is produce return forecasts for a low-dimensional set of portfolio returns, typically two or three. The estimation of volatility is typically more straightforward.[10] Because we are only focusing on few components, we use standard forecasting methods for individual returns.

To summarize, our economic restrictions lead to the following approach to measure conditional expected returns and engage in factor timing:

1. Start from a set of pricing factors F_{t+1}.
2. Reduce this set of factors to a few dominant components, Z_{t+1}, using principal components analysis.
3. Produce separate individual forecasts of each of the Z_{t+1}, that is measures of $\mathbb{E}_t[Z_{t+1}]$.
4. To measure the conditional expected factor returns, apply these forecasts to factors using their loadings on the dominant components.
5. To engage in factor timing, use these forecasts to construct the portfolio given in Equation 6.

In the remainder of this paper, we implement this approach in the context of the cross-section of stock returns. We show how our method allows one to obtain robust measures of expected returns, is useful for market timing, but also generates novel empirical facts to discipline economic models. We also show our method produces strong results when applied to Treasury bonds or exchange rate strategies. Before doing so, we provide an alternative, more statistical motivation for our methodology.

2.3 Alternative Statistical Motivation

Another way to approach our empirical exercise is to look for common sources of variation in risk premia across base assets or factors. For example, starting from a vector of candidate

[9]Stock and Watson (2002) provide conditions under which unconditional principal components analysis identifies important components even in the presence of time-varying parameters.

[10]Moreira and Muir (2017) and Engle and Kelly (2012) are examples of work that point to methods for timing volatility and the benefits it provides.

predictors X_t, we want to assess their usefulness to forecast the returns.[11] In a linear setting, this corresponds to studying the vector of coefficients b_i' in the panel regression:

$$R_{i,t+1} = a_i + b_i' X_t + \varepsilon_{i,t+1}, \tag{7}$$

where one can replace $R_{i,t+1}$ by $F_{j,t+1}$ if focusing on factors. There are multiple ways to aggregate the information in the estimated coefficients of interest, b_i, to judge the success of X_t as a predictor.

One can ask if X_t predicts "something": is there a linear combination of the coefficients $b = [b_1 \cdots b_n]$ that is statistically distinct from zero? This corresponds exactly to a standard Wald test. This notion of predictability, is intuitively too lax. For instance, our conclusion about the predictive value of X_t could be driven by its ability to predict only a few assets or the lowest variance PC portfolios. We show in Section A.3 that a small amount of noise in measured returns can lead to significant spurious predictability of the smallest PC portfolios, even in population. This issue is exacerbated in small samples.

The other extreme is to ask whether elements of X_t predicts "everything", or that all coefficients in a row of b are statistically distinct from zero. For instance Cochrane and Piazzesi (2005) obtain such a pattern predicting Treasury bond returns of various maturities using the cross-section of yields, concluding to the presence of a single common factor in expected returns. While this approach can uncover interesting patterns, it is likely to be too stringent. We show in Section A.1 that such a test is often equivalent to testing whether X_t predicts the first principal component of realized returns. In other words, finding uniform predictability across all assets simply finds predictability of the "level" factor in returns. In contrast, we show in Section A.2 that if a predictor is useful for forecasting index-neutral factor returns, captured by a long-short portfolio, but not for aggregate returns, individual asset predictive regressions are unlikely to uncover such predictability.

Our approach strikes a balance between these two extremes by asking whether X_t predicts the largest principal components of returns. In other words, we focus on common predictability along the few dimensions explaining a large fraction of realized returns. Focusing on components with a large explanatory power avoids the issue of the Wald test. Entertaining multiple dimensions avoids the other extreme of only focusing on the first component of returns, and allows us to study time-series predictability of cross-sectional strategies.

[11]We come back to other types of predictors studied in the literature, in particular those varying in the cross-section, in Section 4.

3 Predicting Stock Returns

Our primary estimation centers on predicting equity returns. As factors, we focus on a broad set of fifty "anomaly" portfolios which effectively summarize the heterogeneity in expected returns, following the logic in Kozak et al. (2018) . The anomaly returns have a moderately strong factor structure, which, however, is substantially weaker than other asset classes such as Treasury bonds or foreign exchange. We proceed with predicting the first two PCs of long-short equity anomaly strategies and show that their predictability is stronger and more robust than the predictability of the aggregate market return. This result is the main building block of our analysis of the predictability of individual anomalies, the optimal factor timing portfolio and the conditional SDF.

3.1 Factor Data

Step 1 of our approach is to start with a set of pricing factors. We construct them as follows. We use the universe of CRSP and COMPUSTAT stocks and sort them into 10 value-weighted portfolios for each of the 50 characteristics studied in Kozak et al. (2018) and listed in Table 13 in the Appendix. Portfolios include all NYSE, AMEX, and NASDAQ firms; however, the breakpoints use only NYSE firms as in Fama and French (2016). Our sample consists of monthly data from November 1973 to December 2015.

We construct long-short anomalies as differences between each anomaly's return on portfolio 10 minus the return on portfolio 1.[12] For each anomaly strategy we also construct its corresponding measure of relative valuation based on book-to-market ratios of the underlying stocks. We define this measure as the difference in log book-to-market ratios of portfolio 10 and portfolio 1.[13]

Most of these portfolio sorts exhibit a significant spread in average returns and CAPM alphas. This finding has been documented in the vast literature on the cross-section of returns and can be verified in Table 13 in Appendix (B). In our sample, most anomalies show a large, nearly monotonic pattern in average returns across decile portfolios, consistent with prior

[12] We perform further transformation as follows: first, we orthogonalize each anomaly strategy with respect to the market portfolio by subtracting its market beta times the return on the market each period, with beta estimated in the full sample. Next, we rescale each anomaly to have equal variance, using variances estimated in full sample.

[13] The book-to-market ratio, bm, of a portfolio is defined as the sum of book equity relative to the total market capitalization of all firms in that portfolio. Equivalently, it is the market-capitalization weighted average of individual stocks' bm ratios. We standardize each anomaly's bm using the mean and standard deviation estimated in full sample.

Table 1: Percentage of variance explained by anomaly PCs

Percentage of variance explained by each PC of pooled anomaly portfolio returns. The top panel shows PC1-PC10 of both long and short ends of anomalies (100 portfolios). The bottom panel focuses on 50 long-short market-neutral strategies.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
	PCs of pooled P1, P10 portfolios									
% var. explained	80.4	3.3	3.1	2.2	0.9	0.7	0.6	0.6	0.5	0.5
Cumulative	80.4	83.7	86.8	89.0	89.9	90.6	91.1	91.7	92.2	92.6
	PCs of Long-Short strategies									
% var. explained	19.2	16.9	10.7	5.7	4.8	3.7	3.3	3.2	2.3	2.0
Cumulative	19.2	36.1	46.8	52.6	57.4	61.1	64.4	67.6	69.9	72.0

research. Rather than studying unconditional mean returns, our primary focus in this paper is on time variation in conditional expected returns, which has received considerably less attention in prior work.

3.2 Dominant Components of the Factors

Step 2 of our approach is to reduce this set of factors to a few dominant components, its largest PCs. We are interested in the joint predictability of anomaly portfolio returns. We construct PCs from the 50 anomaly long-short (decile 10 - decile 1) portfolios and study their predictability. Formally, consider the eigenvalue decomposition of anomaly excess returns, $\text{cov}\left(F_{t+1}\right) = Q\Lambda Q'$, where Q is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues.[14] The ith PC portfolio is formed as $PC_{i,t+1} = q_i' F_{t+1}$ where q_i is the ith column of Q.

Table 1 shows that anomaly portfolio returns exhibit a moderate factor structure. In the top panel we pool all long and short ends of each strategy (portfolios 1 and 10, that is, 100 portfolios in total). The first PC in the top panel, thus, roughly corresponds to the aggregate market. We see that it accounts for about 80% of the total return variation. The second and third principal components account for a much smaller but similar percentage of total variance.

The bottom panel focuses on our primary dataset, the 50 long-short anomaly strategies. The factor structure is weaker: the first two PCs are about equally important and together

[14]We compute $\text{cov}\left(F_{t+1}\right)$ using monthly returns.

account for about 40% of the total variation. In our analysis we focus on these two *relatively* large PC portfolios as well as the aggregate market portfolio. In other words, we study $Z_{t+1} = (R_{mkt,t+1}, PC_{1,t+1}, PC_{2,t+1})$.

In Figure 8 in the Appendix, we explore the eigenvectors underlying the PC factors, showing their composition in terms of the anomaly factors. The loadings have a natural interpretation. We can broadly view the PC1 portfolio as long half of the anomalies and short the other half. PC2 is essentially long most of the anomalies besides momentum and profitability-like strategies. Notably, the two PCs span 82% of the return variation of the average anomaly strategy that equal-weights all anomalies. PC2 alone is responsible for 77% of that variation, while PC1 explains an additional 5%.

3.3 Predicting the Large PCs of Anomaly Returns

Step 3 of our approach is to produce individual forecasts of the dominant components of factor returns. We obtain these forecasts using standard predictive regressions on valuation ratios.

Predictors. Following our broad goal of dimension reduction, we construct a single predictor for each portfolio: we use its net book-to-market ratio. For predicting $PC_{i,t+1}$, we construct its log book-to-market ratio $bm_{i,t}$ by combining the anomaly log book-to-market ratios according to portfolio weights: $bm_{i,t} = q_i' bm_t^F$. Valuation ratios are natural candidate predictors for future returns. Log-linearizing the clean surplus accounting relation of Ohlson (1995), Vuolteenaho (2002) shows that the log book-to-market ratio of a long-only strategy is a discounted sum of all future expected returns for this strategy minus future earning growth. We use the difference between this quantity for the long and short leg of our PCs, thereby capturing information about future expected returns.

Valuation ratios are the most commonly used forecaster for the market return, going back to Shiller (1981), Fama and French (1988), and Campbell and Shiller (1988). They have also been used at the individual stock level, by Vuolteenaho (2002), Cohen et al. (2003), Lochstoer and Tetlock (2016). Our conclusions are therefore readily comparable to this seminal work. However, we are not arguing that other predictors, perhaps motivated by specific economic theories, could not find additional sources of predictability.

Table 2: Predicting Dominant Equity Components with BE/ME ratios

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of excess market returns and two PCs of long-short anomaly returns on three predictors: (i) log of the aggregate book-to-market ratio ($\overline{bm}$), (ii) a restricted linear combination of anomalies' log book-to-market ratios with weights given by the first eigenvector of pooled long-short strategy returns (bm_1); and (iii) a restricted linear combination of anomalies' log book-to-market ratios with weights given by the second eigenvector of pooled long-short strategy returns (bm_2). The first three columns show results from an unrestricted regression where each return is forecast using all three predictors. The next three columns show results from imposing a diagonal structure; that is, each return in only forecast only with its own bm ratio. The last three rows show regression R^2, out-of-sample R^2, and p-value of the Wald test of joint significance of all regression coefficients. Circular block bootstrapped standard errors in parentheses.

	MKT	PC1	PC2	MKT	PC1	PC2
$\overline{bm}$	0.025	-0.043	0.017	0.027	-	-
	(1.39)	(2.71)	(1.05)	(1.64)		
bm_1	-0.003	0.030	-0.002	-	0.024	-
	(0.42)	(4.92)	(0.24)		(3.40)	
bm_2	-0.014	0.009	0.027	-	-	0.027
	(1.79)	(1.18)	(3.49)			(3.42)
R^2	0.097	0.374	0.226	0.044	0.240	0.209
OOS R^2	-0.334	0.235	0.217	0.136	0.245	0.208
Wald test p-value	0.214	0.000	0.009	0.259	0.003	0.003

Predictability results. We analyze the predictability of anomaly PC portfolios and the aggregate market using an annual holding period.[15] For ease of comparison, we normalize each portfolio return to have 5% standard deviation in the full sample.

Table 2 shows the results of the three predictive regressions. The first three columns show results from an unrestricted regression where each return is forecast using all three predictors. Column 1 shows that the market is slightly predictable in sample, with an R^2 of 10%, consistent with prior evidence (e.g., see Cochrane, 2008, 2011). The two restricted linear combinations of book-to-market ratios are insignificant in predicting the aggregate market. These variables, however, are highly significant in forecasting PC1 and PC2 of anomalies, respectively. The PC1-restricted linear combination of log book-to-market ratios, bm_1, forecasts PC1 with a t-statistic of 4.92 and contributes much of the total R^2 of 37%. Similarly, the PC2-restricted linear combination of book-to-market ratios, bm_2, has a t-

[15] We construct overlapping annual holding period returns using a monthly observation frequency.

statistic of 3.49 in forecasting PC2 of long-short anomaly returns and obtains a R^2 of 23%. The Wald test of joint significance of the three predictors (last row of the table) rejects the null of no predictability when predicting returns on PC1 and PC2, but not the aggregate market. To assess the robustness of the predictability evidence, we perform an out of sample (OOS) exercise. We estimate the regression parameters using the first half of the sample and use these parameters to construct forecasts in the second half.[16] The OOS R^2 for the aggregate market is negative whereas it is large for the two PCs, 24% and 22% respectively.[17]

We repeat the forecasting exercise but restrict the off-diagonal elements to zero. That is, we predict each portfolio only with its own *bm* ratio; the last three columns of Table 2 show the results. The full sample R^2 are mechanically lower than before, but surprisingly, the OOS R^2 is higher for the market and approximately unchanged for PC1 and PC2. As before, we find substantially higher in- and out-of-sample predictability for the PCs of anomalies as compared to the market. Based on this evidence we conclude that PC1 and PC2 are highly forecastable, more so than the aggregate market.[18]

Figure 1 shows the estimated expected returns on the market and the two anomaly PCs, in and out of sample.[19] These forecasts for the two principal components are remarkably similar in and out of sample. The OOS prediction for the market, however, substantially deviates from the IS prediction.

The correlation of the estimated expected returns on PC1 and PC2 with the aggregate market are 0.44 and 0.61, respectively. This suggests that more than 75% of variance of expected returns of either of the two PCs is unexplained by market expected returns. Expected returns on the two PCs are nearly uncorrelated with each other; their correlation is 0.05. Overall, this evidence indicates that there are multiple sources of time-varying expected returns in equities.

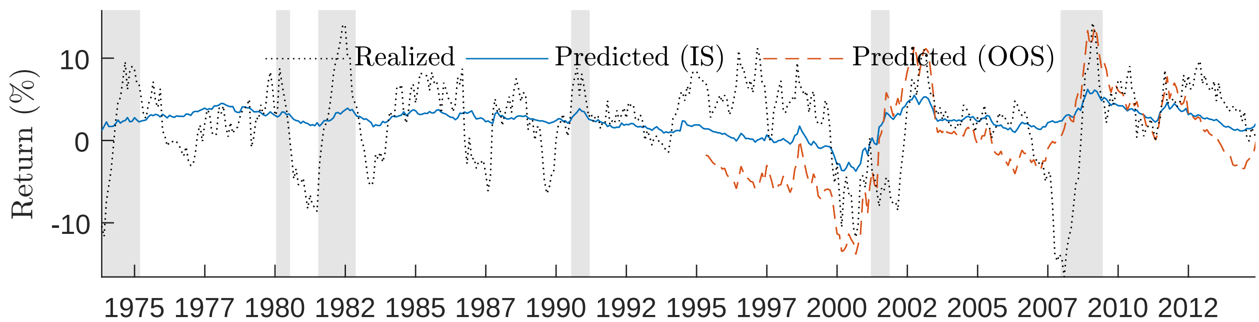
Importance of economic restrictions. Our emphasis on the largest principal components is predicated on the idea that they should capture most of the variation in expected returns. In Figure 2 we report predictability for all the principal components of our factors. Only large PCs are strongly predictable in- and out-of-sample by their own *bm* ratios. In

[16]In Table 12 we show estimates using principal components estimated using only the first half of the data; the results are similar.

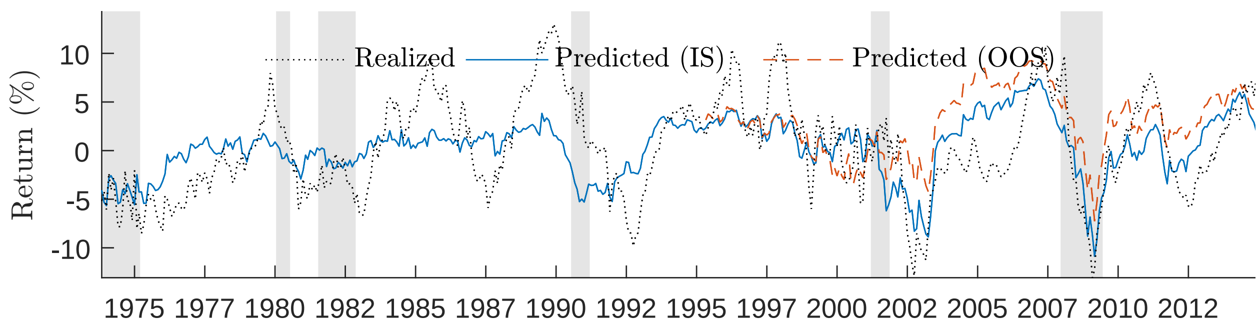
[17]We define the OOS R^2 as $1 - \frac{\mathrm{var}(r-\hat{r})}{\mathrm{var}(r)}$ where $\hat{r}$ is the forecast formed using parameters estimated in the first half.

[18]Imposing further economic restrictions may improve the OOS predictability of aggregate returns (Campbell and Thompson, 2007), but rarely to the high levels we obtain for the PC portfolios.

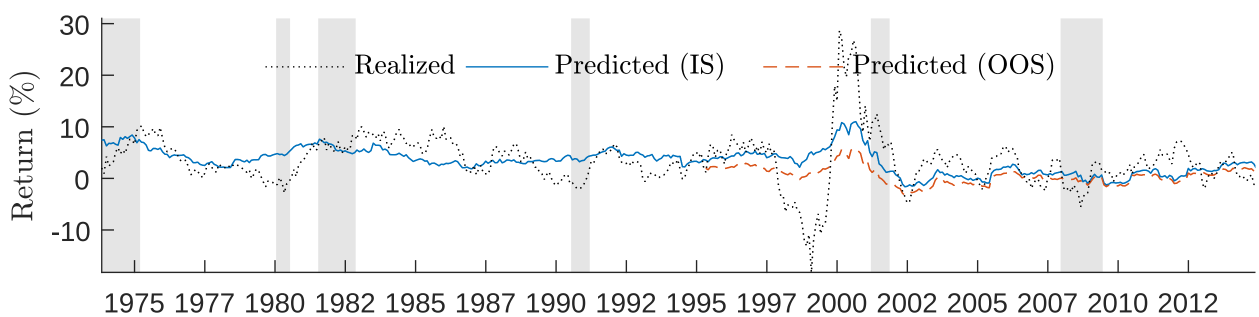
[19]We compute expected returns using the unrestricted estimates (first three columns) of Table 2.



(a) Market



(b) PC1



(c) PC2

Figure 1: Equities Predicted and Realized Returns. The plots show predicted and realized returns of the market (Panel a), PC1 (Panel b), and PC2 (Panel c). The solid line represents forecasts using parameters estimated using the full sample. The dashed line give forecasts using parameters estimated in the first half (pre-1995) and the dotted line gives realized returns.

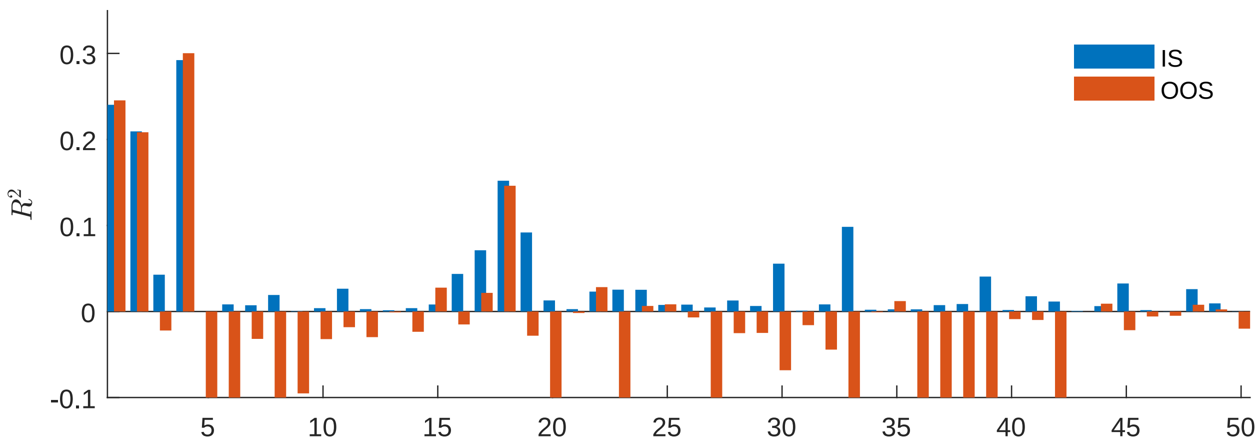


Figure 2: Predictability of Equity PCs with Own *bm* Ratios. The plot shows the in-sample (IS) and out-of-sample R^2 of predicting each PC of anomaly portfolio returns with its own *bm* ratio.

this sense, our focus on predicting only the first few PCs is not only motivated by theory, but finds strong support in the data. In addition, by focusing on predicting only these dominant components we uncover robust predictability in the data and largely ignore spurious predictability that stems from small PCs. This latter result echoes the more statistical concern we put forward as well.

We now turn to forecasting individual anomalies to consider how the restrictions we impose help in this exercise.

4 Predicting Individual Factors

4.1 Individual Factor Timing

Step 4 of our approach is to infer expected return forecasts for the individual factors from the forecasts of our dominant components. Since factors are known linear combinations of the PC portfolios, we can use the estimates in Table 2 to generate forecasts for each anomaly. Notably, each anomaly return is implicitly predicted by the whole cross-section of *bm* ratios. Table 3 shows the in- and out-of-sample R^2 for each anomaly return using our method. Many of the anomalies are highly predictable; roughly half have OOS R^2 greater than 10% and only eleven have negative OOS R^2. Our economically-motivated approach allows us to uncover these patterns in a robust way.

The substantial anomaly predictability we document in Table 2 also contributes to the recent debate on whether these strategies represent actual investment opportunities or are statistical artifacts which are largely data-mined. For example, Hou et al. (2017) claim that most anomalies are not robust to small methodological or sample variations and conclude there is "widespread p-hacking in the anomalies literature." Using a different methodology, Harvey and Liu (2016) argue that the three-factor model of Fama and French (1993) provides a good description of expected returns, and hence, most CAPM anomalies are spurious. Interestingly, we find that some anomalies such as size and leverage which have low unconditional Sharpe ratios are nonetheless highly predictable. This indicates that these strategies at least sometimes represent "important" deviations from the CAPM. This echoes the importance of conditioning information, as emphasized by Nagel and Singleton (2011) and others. More generally these results highlight that a lack of unconditional risk premium does not necessarily imply a lack of conditional risk premium. This insight is maybe best exemplified by PC1, which exhibits a low unconditional Sharpe ratio but a high degree of predictability.

Table 3: Predicting individual anomaly returns: R^2 (%)

Predictive R^2 of individual anomalies returns using implied fitted values based on PC forecasts. Column 1 (IS) provides estimates in full sample. Column 2 (OOS) shows out-of-sample R^2.

	IS	OOS
1. Size	22.6	33.8
2. Value (A)	20.2	11.6
3. Gross Profitability	-2.8	-45.1
4. Value-Profitablity	12.1	20.8
5. F-score	6.4	-5.9
6. Debt Issuance	10.6	12.0
7. Share Repurchases	25.6	11.6
8. Net Issuance (A)	17.7	8.6
9. Accruals	0.3	-0.1
10. Asset Growth	14.4	21.9
11. Asset Turnover	2.9	-3.1
12. Gross Margins	6.8	-11.4
13. Dividend/Price	13.5	14.7
14. Earnings/Price	12.3	15.9
15. Cash Flows/Price	11.4	3.9
16. Net Operating Assets	-1.7	-13.9
17. Investment/Assets	15.7	17.0
18. Investment/Capital	-1.3	5.0
19. Investment Growth	14.9	16.2
20. Sales Growth	5.8	10.0
21. Leverage	23.7	27.5
22. Return on Assets (A)	24.1	3.9
23. Return on Book Equity (A)	21.7	1.9
24. Sales/Price	18.2	18.4
25. Growth in LTNOA	-7.4	-0.5
26. Momentum (6m)	13.4	5.3
27. Value-Momentum	3.4	4.5
28. Value-Momentum-Prof.	-0.2	0.3
29. Short Interest	2.6	10.1
30. Momentum (12m)	9.4	7.3
31. Industry Momentum	7.4	-7.7
32. Momentum-Reversals	9.4	20.7
33. Long Run Reversals	29.9	30.6
34. Value (M)	32.7	26.3
35. Net Issuance (M)	12.2	7.5
36. Earnings Surprises	-8.8	-8.1
37. Return on Book Equity (Q)	17.1	4.0
38. Return on Market Equity	15.5	9.6
39. Return on Assets (Q)	20.8	0.0
40. Short-Term Reversals	6.5	13.6
41. Idiosyncratic Volatility	28.1	23.7
42. Beta Arbitrage	7.4	9.0
43. Seasonality	-2.3	-4.2
44. Industry Rel. Reversals	8.2	16.8
45. Industry Rel. Rev. (L.V.)	14.9	23.0
46. Ind. Mom-Reversals	-2.0	-6.6
47. Composite Issuance	2.0	7.7
48. Price	35.9	35.8
49. Age	16.5	18.2
50. Share Volume	4.4	12.3

4.2 Comparison to Alternative Approaches

We compare our methodology to a variety of alternatives. Table 4 gives the cross-sectional mean, median, and standard deviation of the predictive R^2 for each method, both in- and out-of-sample. The first column reports results from our method, with individual anomaly results already presented in Table 3. The first alternative we consider is predicting each anomaly with the three *bm* ratios from Table 2 with no other restrictions imposed. This fixes the conditioning information but increases the degrees of freedom, allowing each anomaly unrestricted loadings on the predictors. Mechanically, forecasting using the same predictors without cross-sectional restrictions produces higher full sample R^2 than our method, as shown in the second column. Out of sample, however, this method performs quite poorly, with negative mean and median R^2 and a large standard deviation. Again, this result supports the importance of imposing economic restrictions on the joint dynamics of risk premia.

A second alternative is to forecast each anomaly using its own *bm* ratio with no restrictions imposed. The second column shows that this method achieves substantial in-sample predictive R^2, comparable to our method, but is much worse out-of-sample. This suggests that in-sample overfitting is substantial. There are a number of previous papers using a similar methodology. Asness et al. (2000), Cohen et al. (2003) find that the value anomaly itself is forecastable by its own *bm* ratio.[20] Arnott et al. (2016b) use various valuation measures (price ratios) to forecast eight anomaly returns; they forecast each anomaly return with its own valuation measures and find statistically significant results. Asness et al. (2017) use each anomaly's *bm* ratio to construct timing strategies for the value, momentum, and low β anomalies based on the idea that "there is near-unanimous support in the literature for the efficacy of value investing, and thus the value (or value spread) of anything should have some predictive power for that thing. The question is the strength of this relationship." Using their methodology, however, they find "here...strength [is] rather lacking". Implicitly, this method allows for as many independent sources of time-varying expected returns as there are assets or factors to predict.[21] This framework is optimal only if one thinks the anomalies and their expected returns are independent, at odds with our Assumption 2.

The fourth column shows results from a third, natural alternative: predicting each anomaly with its own characteristic spread. For example, when sorting firms into deciles

[20] Asness et al. (2000) uses the ratio of the book-to-market of value stocks to that of growth stocks. Cohen et al. (2003) uses the log difference, as we do.

[21] Formally, the covariance matrix of expected returns is allowed to have full rank.

based on log market capitalization, the characteristic spread at time t is $\log(ME_{1,t}) - \log(ME_{10,t})$ where $ME_{i,t}$ is the weighted average market capitalization of firms in decile i at time t. As above with bm ratios, each anomaly is forecast using its own anomaly specific variable. Somewhat surprisingly, this approach generates very little predictability, both in full sample and out of sample. The performance could possibly be improved through judicious transformation of the sorting variable, though without theoretical guidance on the functional form such an exercise is prone to data mining concerns. In a related paper, Greenwood and Hanson (2012) forecast characteristic-based anomalies using their "issuer-purchaser" spread, or the difference in characteristic (say ME or bm) for net equity issuers vs repurchasers. As above, using an anomaly-specific forecasting variable implicitly allows for for as many independent sources of time-varying expected returns as there are assets or factors to predict.

In the last column we show results from repeating the exercise of forecasting each anomaly with its own bm ratio, but now estimate the predictive coefficients from a panel regression which imposes that the coefficient on bm is the same for all anomalies. Formally, we estimate $\mathbb{E}_t(F_{i,t+1}) = a_{0,i} + a_1 bm_{i,t}$, allowing each anomaly factor to have a different intercept, thereby allowing different unconditional means, but imposing uniform predictability of returns using bm. The predictive R^2 are higher than with other methods, but still substantially lower than using our restricted approach, both in- and out-of-sample. In related work, Campbell et al. (2009), Lochstoer and Tetlock (2016) use a bottom-up approach of aggregating firm-level estimates to portfolios in order to decompose variation in returns into discount rate and cash-flow news. They estimate a panel VAR in which they forecast each stock's return using its own bm ratio, additional stock characteristics, and some aggregate variables.[22] Unlike the previous studies, they impose that the coefficients in the predictive regression are the same for all stocks. While imposing this equality is a form of statistical regularization, it still allows for as many independent sources of time-varying expected returns as there are stocks.

There is another literature quite unlike the above alternatives, focusing on one or a few return predictors. Stambaugh et al. (2012) forecast twelve anomaly strategy returns using the aggregate sentiment index of Baker and Wurgler (2006) and find statistically significant predictability for all of the anomalies they consider. Further, the predictive coefficients are of similar magnitude across anomalies.[23] This is reminiscent of the tradition in bond return

[22] Campbell et al. (2009) do not include any aggregate variables and further cross-sectionally demean stock returns each period. Since most stocks have market β close to unity, their VAR approximately forecasts market-neutral stock returns using each stock's own bm ratio and other characteristics.

[23] They do not report R^2 values, so it is difficult to determine the economic significance of their findings.

Table 4: Predicting individual anomaly returns: comparison to other approaches

The table reports cross-sectional mean, median, and standard deviation of R^2s for our restricted method (from Table 3) and four alternative methods: (i) unrestricted predictability regression of each anomaly on the three *bm* ratios from Table 2; (ii) unrestricted regression of each anomaly on its own *bm* ratio; (iii) unrestricted regression of each anomaly on its own characteristic spread; and (iv) pooled (panel) regression of each anomaly on its own *bm* ratio which imposes the same slope across all anomalies. Two panels show in-sample and out-of-sample R^2 in %.

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
			In-sample R^2, %		
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
			Out-of-sample R^2, %		
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

forecasting which seeks a single variable that is a significant predictor of excess bond returns of all maturities Cochrane and Piazzesi (2005, 2008), Cieslak and Povala (2015). Using a single predictor for all assets implicitly or explicitly assumes there is only a single source of time-varying expected returns.[24] Akbas et al. (2015) start similarly, forecasting eleven anomaly returns individually using aggregate mutual fund and hedge fund flows. Based on the pattern of coefficients, they divide the anomalies into two groups: “real investment factors” and “others”. They then form an aggregate portfolio return for each group and forecast these two returns and find substantial R^2 for the “other group” and nearly zero for the investment group. This is similar in spirit to our construction to two PCs of anomalies as dependent variables, though determined ex-post rather than ex-ante.

[24]Formally, the covariance matrix of expected returns has at most rank one.

Table 5: Equity summary statistics

The table shows mean, standard deviation, skewness and Sharpe ratio of returns on the dominant components of equities (the market and first two PCs of anomalies). The first three columns are for static investment strategies and next three are for pure timing strategies.

	Static portfolios			Pure timing portfolios		
	MKT	PC1	PC2	MKT	PC1	PC2
Mean (%)	2.34	0.35	3.38	0.02	0.09	0.06
Std. Dev. (%)	5.00	5.00	5.00	0.10	0.18	0.25
Skewness	-0.87	0.08	0.82	2.49	2.84	5.23
Sharpe Ratio	0.47	0.07	0.68	0.24	0.51	0.22

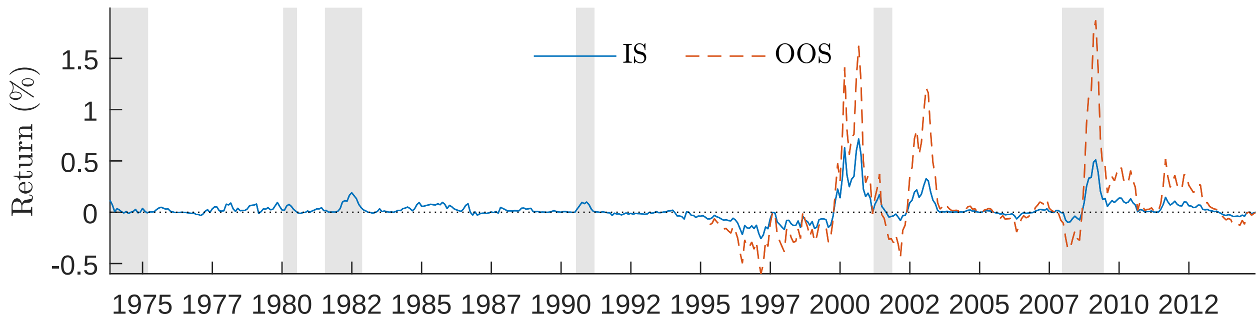
5 Optimal Factor Timing Portfolio and Conditional SDF

We turn to step 5 of our approach: using our forecasts to form an optimal factor timing portfolio. Of more interest economically, this optimal portfolio also informs the properties of the conditional SDF. We first quantify the gains to timing each of our three dominant components individually. Then we combine them to get an overall measure of the benefits to factor timing, or equivalently of the extra variability of the conditional SDF. Finally, we study how the composition of this portfolio and discount factor varies over time.

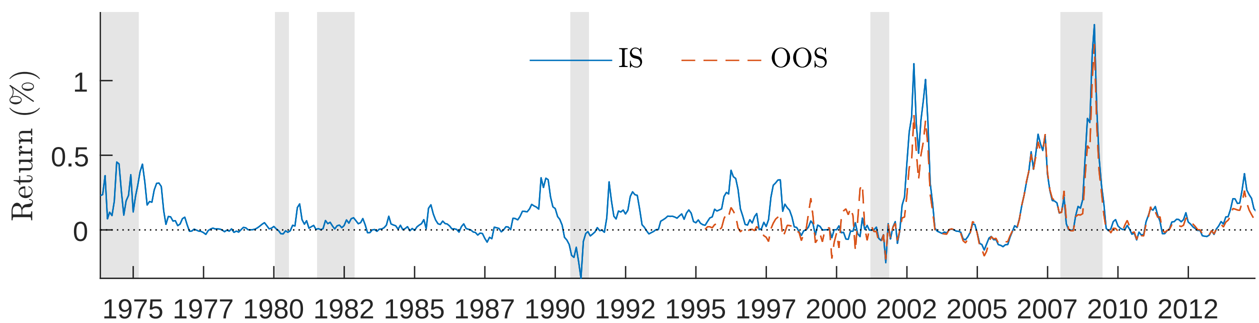
5.1 Timing Dominant Components Individually

We now assess the benefits to timing each of our dominant components by comparing two strategies: (1) a static strategy which invests a constant fraction of wealth proportional to an asset's unconditional mean and (2) a dynamic strategy which invests proportional to the conditional mean of the portfolio. For full sample results, we construct conditional means using parameter estimates from Table 2. For OOS, we use parameters estimated using the first half of the data. For ease of comparison, we report the difference between the returns on the static and dynamic strategies. This is the return on a "pure timing" strategy with zero weight on average, but which is sometimes long and other times short the asset. Hence, it measures the relative gains and losses from using conditioning information.

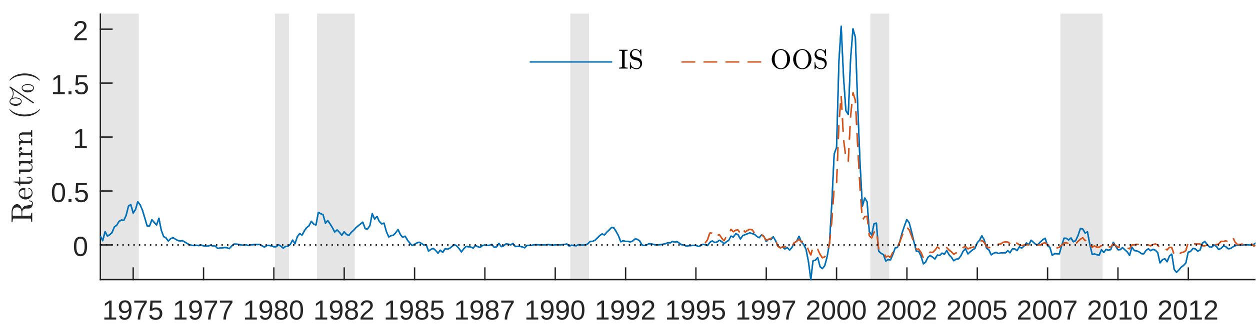
Table 5 shows summary statistics of the pure timing strategies (right panel) which exploit factor timing and compares them to the summary statistics of the PC portfolios (left panel).



(a) Market



(b) PC1



(c) PC2

Figure 3: Equities Timing Strategy Returns. The figure shows realized returns on pure timing strategies of the market (Panel a), PC1 (Panel b), and PC2 (Panel c).

Interestingly, PC1 is essentially unpriced unconditionally yet the timing strategy that exploits time-variation in PC1's expected returns has a Sharpe ratio of 0.51. This evidence suggests that the economic risks investors worry about conditionally are often very different from those they worry about on average. Timing PC2 also provides a sizable Sharpe ratio, 0.22.[25]

Figure 3 shows the time series of realized returns on dynamic strategies for the market, PC1, and PC2. The full-sample Sharpe ratio for the market is 0.24 compared to Sharpe ratios of 0.51 and 0.22 for PC1 and PC2, respectively. Out-of-sample Sharpe ratios of PCs are similar to their in-sample counterparts. The OOS Sharpe ratio of the aggregate market is 0.25. OOS Sharpe ratios of PC1 and PC2 are 0.44 and 0.26, respectively. These numbers suggest there is a substantial amount of time-series variation in expected relative returns, independently of any unconditional premium. An investor can therefore substantially enhance her Sharpe ratio by varying exposures to anomaly strategies over time compared to simply passively holding the anomalies and collecting their unconditional risk premia.

We now discuss the economic implications of this predictability and factor timing in general through the lens of optimal portfolios and SDFs.

5.2 The Optimal Factor Timing Portfolio

We showed that the anomaly PC returns are highly predictable, in and out of sample, and generate large Sharpe ratios when timed. We now combine these and the aggregate market return into an optimal factor timing portfolio and consider the implied SDF. We construct this portfolio based on the SDF in Equation 5, which we denote as m^c_{t+1}. Specifically, we use our methodology to predict factor means, $\mathbb{E}_t\left[Z_{t+1}\right]$. We then use these forecasts to construct forecast errors and compute an estimate of the conditional covariance matrix of the market and PC returns, $\Sigma_{Z,t}$, which we assume is homoskedastic. In Table 11 we empirically test this assumption by forecasting squared errors from the regressions in Table 2 using bm ratios and find no robust predictability.[26] Finally, we construct portfolio weights $\omega_t = \Sigma_{Z,t}^{-1}\mathbb{E}_t\left[Z_{t+1}\right]$. As a benchmark we use the the unconditional SDF and the corresponding optimal static portfolio, $m^u_{t+1} = a_t - \mathbb{E}\left[Z_t\right]' \Sigma_Z^{-1} Z_{t+1}$, with weights given by $\omega^u = \Sigma_Z^{-1}\mathbb{E}\left[Z_{t+1}\right]$.

To illustrate the benefits of factor timing, we consider a "pure timing" portfolio with weights given by $\omega_{pt} = \Sigma_{Z,t}^{-1}\left(\mathbb{E}_t\left[Z_{t+1}\right] - \mathbb{E}\left[Z_t\right]\right)$. Such a portfolio has zero average weight in

[25] The lower Sharpe ratio from timing PC2 is a manifestation of the heteroskedasticity we document in Table 11.

[26] It is definitely possible that using other conditioning information could uncover predictability of conditional variances.

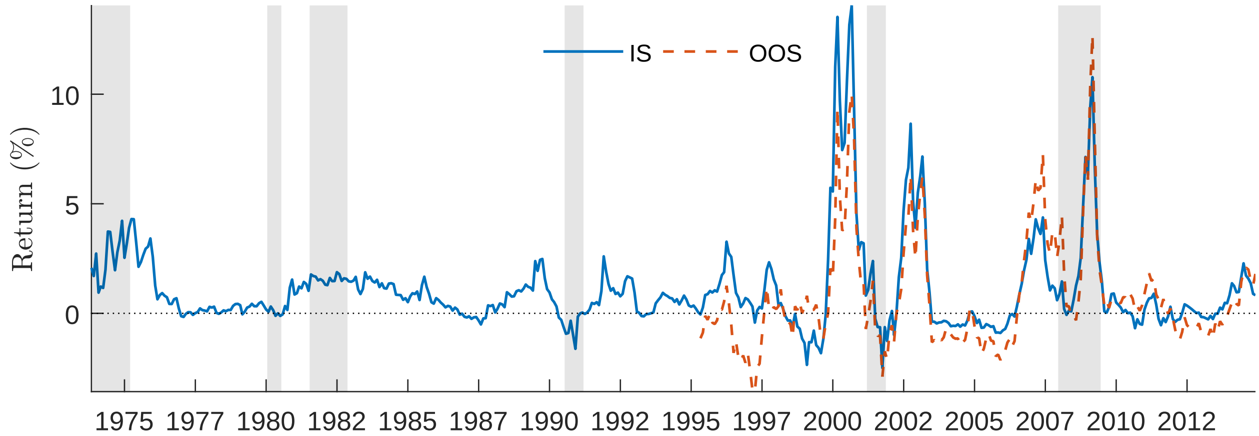


Figure 4: Pure Timing Portfolio Returns (Relative to Static). This figure plots the return on the optimal pure factor timing portfolio, with portfolio weights $\omega_{pt} = \Sigma_{Z,t}^{-1} (\mathbb{E}_t [Z_{t+1}] - \mathbb{E} [Z_t])$. The blue solid line depicts this difference in the full sample; the dashed red line uses only the first half of the sample to estimate predictive regression parameters.

Table 6: Conditional and unconditional Sharpe ratios of various timing strategies

The table reports conditional and average unconditional Sharpe ratios of three strategies: (i) static strategy, based on unconditional estimates of Σ_Z and $\mathbb{E} [Z_t]$ with weights given by $\omega_s = \Sigma_Z^{-1} \mathbb{E} [Z_t]$; (ii) expected factor timing strategy based on the mean forecasts $\mathbb{E}_t [Z_{t+1}]$ but with homoskedastic conditional covariance matrix of residuals $\Sigma_{Z,t}$, with portfolio weights $\omega_t = \Sigma_{Z,t}^{-1} \mathbb{E}_t [Z_{t+1}]$; and (iii) pure timing strategy constructed using weights $\omega_{pt} = \Sigma_{Z,t}^{-1} (\mathbb{E}_t [Z_{t+1}] - \mathbb{E} [Z_t])$.

	Static portfolio	Dynamic portfolio	Pure timing portfolio
Unconditional SR	0.91	0.76	0.53
Average conditional SR	0.96	1.33	0.85

all risky assets, and hence, its performance comes purely from using conditioning information. Figure 4 plots the portfolio's returns. The blue solid line depicts this difference in the full sample; the dashed red line uses information only the first half of the sample to estimate predictive regression parameters. The figure illustrates that the pure timing portfolio delivers high average returns and rarely loses money, both in the full sample and out of sample. This finding suggests that the implied trading strategy is highly profitable.

We explore the performance of various timing portfolios in more detail in Table 6. We consider three strategies corresponding to the optimal static portfolio, the optimal factor timing portfolio, and the pure timing portfolio defined above. The table reports unconditional and average conditional Sharpe ratios for these strategies. For any portfolio ω_t, the conditional Sharpe ratio is $S_t = (\omega_t' \Sigma_{Z,t} \omega_t)^{-\frac{1}{2}} \omega_t' \mathbb{E}_t [Z_{t+1}]$. Applying this formula for the optimal static portfolio, we obtain its expected conditional Sharpe ratio:

$$S_t^u = \left(\mu' \Sigma_Z^{-1} \Sigma_{Z,t} \Sigma_Z^{-1} \mu\right)^{-\frac{1}{2}} \mu' \Sigma_Z^{-1} \mathbb{E}_t [Z_{t+1}]. \tag{8}$$

where $\mu \equiv \mathbb{E}[Z_{t+1}]$. Similarly, by applying this formula to the optimal timing portfolio, we can write its expected conditional Sharpe ratio as

$$S_t^c = \left(\mathbb{E}_t [Z_{t+1}]' \Sigma_{Z,t}^{-1} \mathbb{E}_t [Z_{t+1}]\right)^{\frac{1}{2}}. \tag{9}$$

The table shows that the optimal factor timing strategy delivers a substantially higher average conditional Sharpe ratio than the strategy that ignores the predictability of returns: 1.33 versus 0.96. Factor timing, however, does not necessarily improve the unconditional Sharpe ratio, as can be seen from the table. The reason for this is that the unconditional variance of the timing portfolio becomes large due to strong time-variation in expected returns. The pure timing strategy, which is a difference of returns on the optimal factor timing and optimal static strategies, delivers a high average conditional (and unconditional) Sharpe ratio of 0.85 (0.53) as well. Clearly the benefits of timing anomaly returns are sizable, suggesting that the conditional SDF often differs substantially from its unconditional counterpart, and that the time-series predictability of returns is important both statistically and economically.

The conditional volatility of an SDF is equal to the conditional Sharpe ratio of the corresponding optimal portfolio. Therefore, the sizable benefits of factor timing suggests that the conditional SDF is substantially more volatile than its unconditional counterpart. This finding further deepens the longstanding economic puzzle of the high volatility of the SDF.[27] We now explore another difference between the conditional and unconditional SDFs: the dramatic time-varying difference in their composition.

5.3 The Changing Composition of the Conditional SDF

Large differences in conditional Sharpe ratios in Table 6 and highly positive returns on the pure timing portfolio in Figure 4 suggest that the conditional and unconditional SDFs are quite different from each other. We further assess the economic magnitude of such differences using two separate metrics.

[27]This puzzle was first pointed out in an unconditional context by Hansen and Jagannathan (1991).

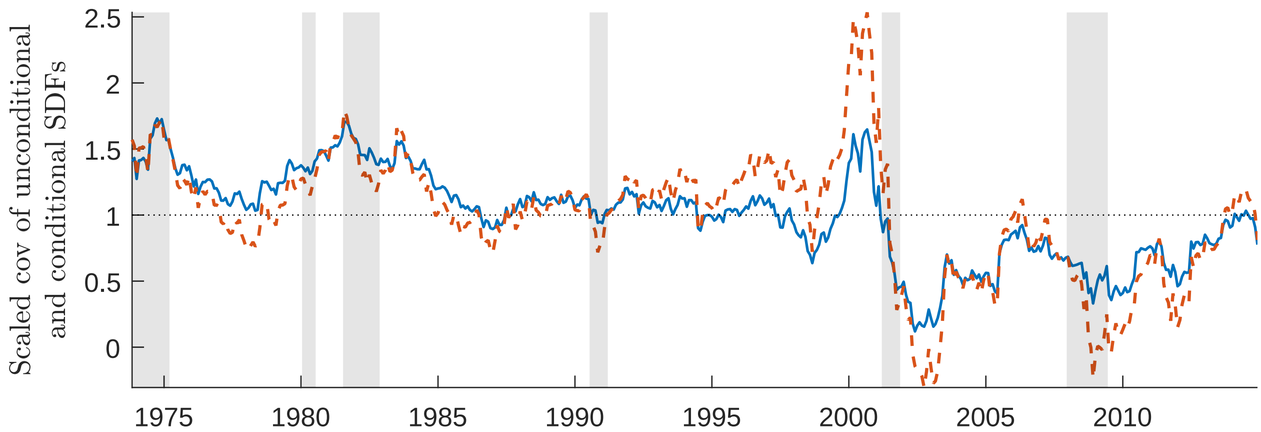


Figure 5: Covariance of Unconditional and Conditional SDFs. This figure plots the conditional covariance of the unconditional and unconditional SDFs, $\mathrm{cov}_t\left(m^u_{t+1}, m^c_{t+1}\right)$, from Equation 10, rescaled by the unconditional variance of m^u_{t+1}, $\mathbb{E}\left[Z_{t+1}\right]' \Sigma_Z^{-1} \mathbb{E}\left[Z_{t+1}\right]$ (solid blue line). The red dashed line ignores predictability of the aggregate market, that is, sets the first element of $\mathbb{E}_t\left(R_{mkt,t+1}\right)$ equal to $\mathbb{E}\left(R_{nkt,t+1}\right)$.

Our first metric is the conditional covariance between the two SDFs,

$$\mathrm{cov}_t\left(m^u_{t+1}, m^c_{t+1}\right) = \mathbb{E}\left[Z_{t+1}\right]' \Sigma_Z^{-1} \mathbb{E}_t\left[Z_{t+1}\right], \tag{10}$$

rescaled by the unconditional variance of m^u_{t+1}, $\mathrm{var}\left(m^u_{t+1}\right) = \mathbb{E}\left[Z_{t+1}\right]' \Sigma_Z^{-1} \mathbb{E}\left[Z_{t+1}\right]$, which can be roughly interpreted as the conditional regression coefficient of m^c_{t+1} on m^u_{t+1}. Note that this statistic does not depend on the conditional covariance matrix $\Sigma_{Z,t}$, and hence, all time variation is due to predictability encoded in the vector of expected returns $\mathbb{E}_t\left[Z_{t+1}\right]$.

In Figure 5 we plot this rescaled covariance (blue solid line). The red dashed line shows the resulting statistic ignoring predictability of the aggregate market; that is, setting $\mathbb{E}_t\left(R_{mkt,t+1}\right)$ equal to its unconditional mean $\mathbb{E}\left(R_{mkt,t+1}\right)$. The resulting covariance is very similar but slightly more volatile. The figure shows that predictability of PC1 and PC2 of anomalies primarily drives the disconnect between unconditional and conditional SDFs; factor timing, rather than market timing, drives the difference between the two SDFs. This variation is sizable—the statistic ranges from around 0 to above 1.5—suggesting that the conditional SDF is often very different from its unconditional counterpart. Therefore, the predictability that our method uncovers constitutes an important economic source of variation in the composition of the conditional SDF.

Our second metric focuses on the magnitude of variation of risk prices through time. We plot the time series of conditional SDF coefficients, $\omega_t = \Sigma_{Z,t}^{-1} \mathbb{E}_t\left[Z_{t+1}\right]$, for the market

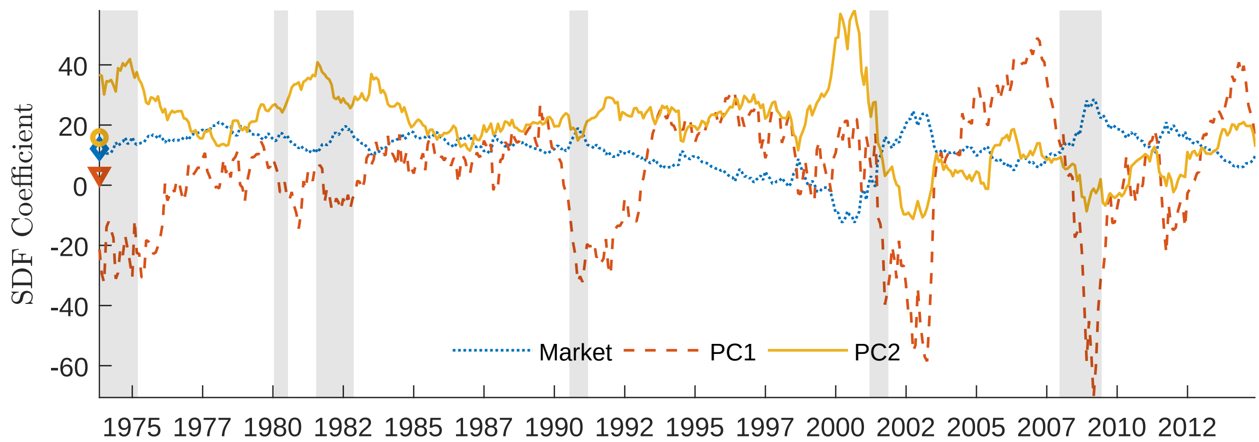


Figure 6: Time variation in risk prices. The figure plots the time series of conditional SDF coefficients, $\mathbb{E}_t\left[Z_{t+1}\right]' \Sigma_{Z,t}^{-1}$, for the market (dotted blue), PC1 (dashed red), and PC2 (solid yellow) factors. Markers on the y-axis depict the point estimates of the unconditional SDF coefficients (same colors; diamond, triangle, and circle, respectively). Factors are rescaled to have the same standard deviation of 5%.

(dotted blue), PC1 (dashed red), and PC2 (solid yellow) factors. Markers on the y-axis depict the point estimates of the unconditional SDF coefficients (same colors; diamond, triangle, and circle, respectively). The three SDF component returns are rescaled to have the same unconditional standard deviation of 5%.

The figure shows that prices of risk vary substantially through time. Notably, the price of market risk tends to increase during times of economic recessions. Similarly, the coefficient on the PC1 factor seem to contain some business-cycle-related variation as well. Importantly, SDF coefficients on PC1 and PC2 factors vary substantially more than the coefficient on the market factor, which is a consequence of the high predictive R^2 observed in Table 2.

Overall, the evidence in this section strongly suggests that (i) the benefits of timing anomaly returns are large, , (ii) the predictability that our method uncovers constitutes an important economic source of variation in the composition of the conditional SDF, and, hence, (iii) the conditional SDF differs substantially from its unconditional counterpart

6 Other Asset Classes

To assess the validity of our approach and conclusions beyond the setting of equities, we now turn to other asset classes. Specifically, we analyze Treasury bonds and currency carry trades. The remarkable finding of Section 3 is that index-neutral long-short portfolios are highly predictable, more so than the index itself. This result leads to a number of our conclusions so we ask whether it applies to these other asset classes. Treasuries and currencies exhibit a very strong factor structure; their returns are almost perfectly spanned by two PCs. Therefore steps 1 and 2 of our approach are combined and we directly study the conditional returns of these two PCs. As predictors, we use forward spreads, the natural counterpart to valuation ratios. For both bonds and currencies, we find substantial predictability of the second, index-neutral, PC portfolio which is more robust OOS than predictability of the first PC, the index.

6.1 Treasury Bonds

We study the predictability of Treasury bond returns. We find substantial predictability of relative bond returns across maturities in addition to the existing evidence on aggregate predictability. We show this relative predictability is masked when forecasting individual bond returns. In contrast, it is substantial and clearly visible when directly forecasting the second principal component.

We obtain yields on zero-coupon Treasury bonds with maturities from 1 to 15 years from Gürkaynak et al. (2006).[28] As is common, we calculate annual log excess returns from these log yields.[29] As in Cieslak and Povala (2015), we then compute rescaled excess returns of bonds of maturity n, $rx_{t+1}^{(n)}$, by dividing excess return by $n-1$:[30]

$$rx_{t+1}^{(n)} \equiv -y_{t+1}^{(n-1)} + \frac{n}{n-1} y_t^{(n)} - \frac{1}{n-1} y_t^{(1)}.$$

The first two principal components, which we term LevelR, SlopeR, account for nearly 100% of the variance of the fourteen individual zero-coupon bond returns.[31] Hence, we restrict our analysis to these PCs, treating LevelR as the "index".

We use forward spreads as the counterpart of the valuation ratio for stocks. Following

[28]Following Joslin et al. (2014), our sample is 1985-2015.

[29]As in Cochrane and Piazzesi (2005), we use the one-year zero-coupon yield as the risk-free rate. See Cochrane and Piazzesi (2008) for a comprehensive exposition of bond yields, forward rates, and returns.

[30]Rescaling largely eliminates scale effects across bond returns. Scaled returns have approximately equal standard deviation whereas this statistic varies by a factor of 10 for unscaled returns. The transformed excess returns all have modified duration equal to unity.

[31]These are principal components of *returns*, unlike the usual Level and Slope which are principal components of *yields*. The two are however related: LevelR is strongly correlated with changes in Level, and a similar relation holds for SlopeR and Slope.

Cochrane and Piazzesi (2008), we define log forward rates as

$$f_t^{(n)} \equiv p_t^{(n-1)} - p_t^{(n)},$$

where $p_t^{(n)}$ denotes the time t log price of an n-year bond. Likewise, log forward spreads are

$$s_t^{(n)} \equiv f_t^{(n)} - y_t^{(1)},$$

where $y_t^{(1)}$ is the time t yield on a one-year zero-coupon bond.[32] Given the strong factor structure in forward spreads, we use only the first three principal components of spreads as predictive variables; we denote these by FS1, FS2, and FS3.[33]

In addition, there is a long literature arguing for a relationship between macroeconomic variables and bond returns, with a number of other papers highlighting issues related to these findings.[34] We revisit this issue, illustrating the importance to impose our economic restrictions to understand this predictability. Following Joslin et al. (2014) we proxy for macroeconomic conditions with GRO, the Chicago Fed National Activity Index.[35] We consider other macroeconomic variables in Appendix B.2.1.

Table 7 shows the estimates from predicting LevelR and SlopeR using only the PCs of forward spreads and also including GRO as a further predictive variable. Using only forward spreads, LevelR and SlopeR exhibit similar full sample predictability, but SlopeR has dramatically higher OOS R^2. The negative OOS R^2 for the index (LevelR) echoes our finding for equities. Adding GRO as an additional predictive variable substantially improves the full and OOS R^2 for SlopeR but is nearly irrelevant for LevelR. Finally, the expected returns on LevelR and SlopeR are nearly uncorrelated, whether or not we include GRO. This finding demonstrates that in Treasury bonds, there are multiple important sources of time-varying risk premia.[36]

[32]As in Cochrane and Piazzesi (2008), we compute three month moving averages of forward spreads, which reduces the effect of measurement error.

[33]The eigenvectors of returns almost perfectly coincide with the eigenvectors of forward spreads.

[34]See Estrella and Mishkin (1997), Evans and Marshall (2001, 1998), Ang and Piazzesi (2003), Cooper and Priestley (2008), Ludvigson and Ng (2009), Bikbov and Chernov (2010), Joslin et al. (2014), Cieslak and Povala (2015) among others. Issues are pointed out for instance in Duffee (2013), Bauer and Rudebusch (2016).

[35]For ease of comparison we rescale GRO to have 1% standard deviation.

[36]This low correlation could obtain due to a particular pattern of time variation in risk premia combined with time variation in factor loadings, even with a conditional single-factor model. In unreported results we formally reject such a model with LevelR as the single factor.

Table 7: Predicting PC Returns with Forward Spreads and GRO

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of the first two principal components of zero-coupon bond excess returns on the the first three principal components of lagged forward spreads and GRO (Chicago Fed National Activity Index).

	LevelR			SlopeR		
FS1	0.31	0.28	-	0.53	0.65	-
	(1.72)	(1.56)		(2.65)	(4.04)	
FS2	-1.76	-1.91	-	0.12	0.83	-
	(2.85)	(2.93)		(0.18)	(1.44)	
FS3	-3.66	-2.98	-	3.96	0.76	-
	(1.97)	(1.48)		(1.95)	(0.43)	
GRO	-	-0.50	-0.55	-	2.34	1.78
		(0.70)	(0.82)		(3.60)	(2.51)
R^2	0.22	0.23	0.01	0.22	0.37	0.13
OOS R^2	-1.74	-1.97	-0.01	0.11	0.18	-0.03
Wald test p-value	0.00	0.01	0.71	0.03	0.00	0.04

The importance of economic restrictions for predictability of macroeconomic variables. The predictive behavior of GRO highlights the importance of focusing on dominant components rather than individual assets. Lack of predictive power for LevelR but significant incremental predictability of SlopeR suggests GRO would not be statistically significant in individual bond regressions. Figure 7 shows the estimated coefficient on GRO by bond maturity with ± 2 standard errors. The coefficients show an upward sloping pattern across maturity, as expected based on the results in Table 7, but none of the coefficients are statistically significant at conventional levels, except for the 2-year bond. This demonstrates that individual bond forecasting regressions often miss predictability across maturities which is statistically and economically significant, but only detectable by directly forecasting PCs of returns.

6.2 Foreign Exchange

We study the predictability of foreign exchange (FX) returns. Similarly to the other two asset classes, we find that a relative carry portfolio is more predictable than a basket of all currencies against the dollar.

We construct a panel of annual holding period excess return to investing in each currency

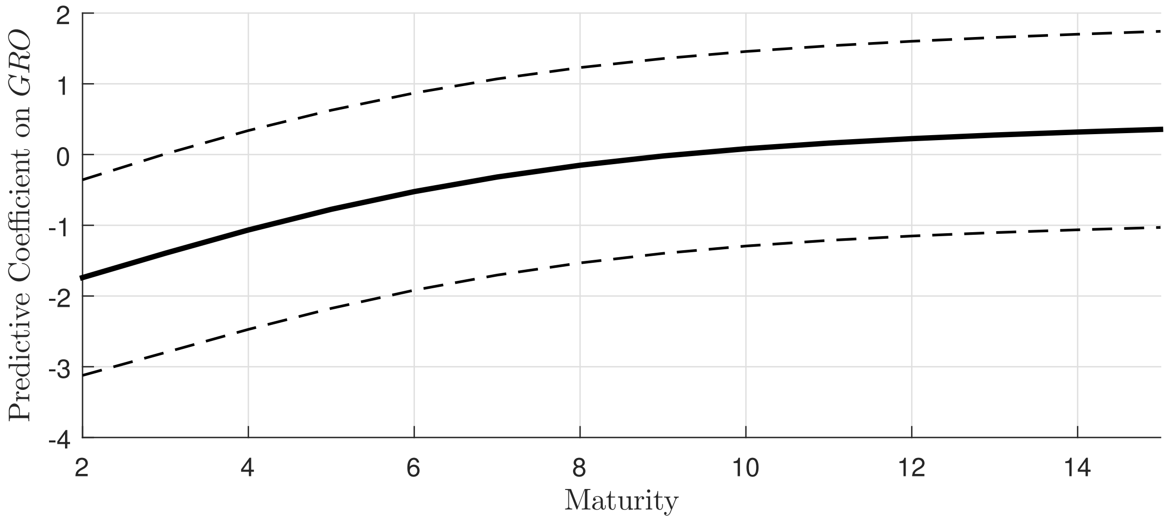


Figure 7: Predicting Bond Returns with Economic Activity. We plot predictive coefficients on the Chicago Fed National Activity Index GRO (and ± 2 standard error bands) by maturity from predictive regressions of zero-coupon bond excess returns on GRO and the first three principal components of lagged forward spreads.

j as follows:

$$r_{j,t+1} = f_{j,t} - s_{j,t} - \Delta s_{j,t+1},$$

where $s_{j,t}$ is the spot exchange rate between currency j and the dollar and $f_{j,t}$ is the corresponding forward exchange rate.[37] Spots, forwards and interest rates come from various datasets: WM/Reuters dataset for spots; WM/Reuters dataset for forwards; BBI/Reuters dataset for developed countries (spots and forwards); Financial Times Eurocurrency for interest rates.[38] We combine all series country by country in order to obtain the longest time series possible. Each country is added to the sample as its data become available. Euro area countries are removed at the dates when the Euro is adopted in each individual country.

As factors, we sort the individual currencies into five portfolios based on their forward spread with the US — or equivalently their interest rate differential by covered interest parity — following Lustig et al. (2011). Portfolios are rebalanced daily based on the average of the forward spread in the recent month. We add and drop countries to portfolios as new data becomes available. The sample is from January 1985 until January 2017. In the beginning

[37]We use annual holding period returns measured at a daily frequency.

[38]We also use Eurocurrency Financial Times interest rates and substitute them in place of forward rates according to covered interest rate parity when longer data series are available. Covered interest parity gives: $f_{j,t} - s_{j,t} = i_{j,t} - i_{\$,t}$, the difference between forward and spot is equal to the difference between foreign and domestic rates.

Table 8: Predicting PC returns of Foreign Exchange rates

Forecasting regression coefficient estimates. We forecast the first two PC portfolios of currencies sorted on forward spreads (Dollar-Carry and Relative-Carry) using the first two PCs of forward spreads (FS1 and FS2). Circular block bootstrapped standard errors in paratheses. The sample is daily from January 1985 until January 2017.

	Dollar-Carry	Relative-Carry
FS1	0.52	-0.89
	(1.24)	(2.76)
FS2	0.07	1.30
	(0.11)	(2.48)
R^2	0.04	0.19
OOS R^2	-0.06	0.05
Wald test p-value	0.67	0.00

of the sample we have about three countries per portfolio and about ten per portfolio in the second half of sample.[39] As with bonds, there are only two principal components with meaningful variance, which we denote as Dollar-Carry and Relative-Carry. Dollar-Carry is effectively the return of investing in a basket of all currencies against the U.S. Dollar. Relative-Carry is the return of a portfolio long currencies with high forward spread and short currencies with low forward spread.[40]

We use again valuation ratios, log forward-spreads as forecasting variables. First we average the forward spread across portfolio components at each point in time; this generates five time-series predictors. Since there are only two PCs of forward spreads with meaningful variance, we use only these linear combinations as predictors; we denote these as FS1 and FS2.[41] Lustig et al. (2014) predict Dollar-Carry returns using the average forward differential. Since FS1 has nearly perfect correlation with the average forward differential, their method is essentially equivalent to predicting the first PC of currency returns with FS1.

[39]Our sample contains the following countries: Australia, Austria, Belgium, Bulgaria, Canada, Chile, Colombia, Croatia, Czech Rep., Denmark, EU, Egypt, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Jordan, Latvia, Lithuania, Malta, Mexico, Morocco, Netherlands, New Zealand, Norway, Pakistan, Philippines, Poland, Portugal, Romania, Russia, Singapore, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Tunisia, UK.

[40]Lustig et al. (2014) study time-series predictability of the Dollar-Carry factor. Lustig et al. (2011), Lettau et al. (2014) study the Relative-Carry ("Slope") factor in the cross-sectional setting, and Bakshi and Panayotov (2013), Ready et al. (2017) present some evidence of predictability of Relative-Carry.

[41]As we found with bonds, the eigenvectors of forward spreads are very similar to the eigenvectors of returns.

Estimation results are shown in Table 8. For Dollar-Carry, the R^2 is small, neither predictive variable is individually significant, and a joint Wald test fails to reject the null of no predictability. For Relative-Carry, in contrast, the R^2 is economically large, both predictors are individually significant, and a joint test convincingly rejects the null. This evidence is in line with our previous findings in other asset classes: cross-sectional returns are more predictable in the time series than aggregate returns.

The correlation of expected returns on Dollar-Carry and Relative-Carry is 0.69, suggesting that more than 50% of variance of expected returns of each of the two PCs is unexplained by the other PC. Again, we conclude that there are multiple sizable sources of time-varying priced risk.

7 Concluding Remarks

In this paper we study factor timing, which combines the ideas of long-short factor investing and market timing. Unlike much previous work, we find that factor timing can be very valuable, generating superior performance relative to market timing and factor investing alone. To recover robust predictability of the multitude of market-neutral factors, it is crucial to acknowledge the economic nature of factor timing. Since the stochastic discount factor is affine in the return on the optimal factor timing portfolio, economic restrictions on the composition of the SDF map directly into restrictions on the construction of the optimal portfolio. The idea of no near-arbitrage (Ross, 1976, Cochrane and Saa-Requejo, 2000, Kozak et al., 2017) implies that the SDF can be well-approximated as a linear combination of a few principal components which are the largest sources of variation in realized factor returns. This implies that the predictability of any factor stems from predictability of these few PCs. Hence, rather than forecasting each factor in isolation, we can recover nearly all relevant predictability by forecasting only these few SDF components. This insight dramatically reduces the dimensionality of the forecasting problem, leading to significant and robust return predictability. Given return forecasts for these few dominant PCs, we construct the optimal factor timing portfolio.

The same link we exploit to discipline our forecasting exercise also allows us to learn about the SDF from the resulting optimal portfolio. First, we find that the conditional SDF which accounts for the substantial time-variation in expected returns is often quite different from its unconditional counterpart. Specifically, there are times when their conditional correlation is nearly zero. This results from the substantial time-variation in optimal factor portfolio weights due to return predictability. Moreover, this wedge is mostly driven by market-neutral factors which exhibit significantly more predictability than the market index.

Finally, we show that substantial, robust predictability of index-neutral long-short factors is a pervasive phenomenon not specific to equities. Specifically, we study the predictability of Treasury bonds and foreign exchange rates, asset classes which exhibit a very strong factor structure featuring a dominant "level" factor and a second "slope" factor. For bonds, a duration-neutral portfolio of long minus short maturity bonds ("slope") is strongly predictable, more so than the average bond return ("level") often studied in the previous literature. For currencies, a relative carry portfolio ("slope") is more predictable than the dollar-carry strategy ("level").

The methods and facts we study in this paper are only the beginning of the economic enterprise of understanding the evolution of drivers of risk premia. They constitute a set of facts that future theories, which will name these sources of variation, will have to reflect.

References

Adrian, T., R. K. Crump, and E. Moench (2015). Regression-based estimation of dynamic asset pricing models. *Journal of Financial Economics 118*(2), 211–244.

Akbas, F., W. J. Armstrong, S. Sorescu, and A. Subrahmanyam (2015). Smart money, dumb money, and capital market anomalies. *Journal of Financial Economics 118*(2), 355–382.

Ang, A. and M. Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics 50*, 745–787.

Arnott, R. D., N. Beck, and V. Kalesnik (2016a). Timing "smart beta" strategies? of course! buy low, sell high!

Arnott, R. D., N. Beck, and V. Kalesnik (2016b). To win with "smart beta" ask if the price is right.

Asness, C., S. Chandra, A. Ilmanen, and R. Israel (2017). Contrarian factor timing is deceptively difficult. *The Journal of Portfolio Management 43*(5), 72–87.

Asness, C. S. (2016). The siren song of factor timing aka "smart beta timing" aka "style timing". *The Journal of Portfolio Management 42*(5), 1–6.

Asness, C. S., J. A. Friedman, R. J. Krail, and J. M. Liew (2000). Style timing: Value versus growth. *Journal of Portfolio Management 26*(3), 50–60.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance 61*(4), 1645–1680.

Bakshi, G. and G. Panayotov (2013). Predictability of currency carry trades and asset pricing implications. *Journal of Financial Economics 110*(1), 139–163.

Bansal, R. and A. Yaron (2004). Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance 59*(4), 1481–1509.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics 115*(1), 1–24.

Bauer, M. D. and J. D. Hamilton (2017). Robust bond risk premia. Technical report, National Bureau of Economic Research.

Bauer, M. D. and G. D. Rudebusch (2016). Resolving the spanning puzzle in macro-finance term structure models. *Review of Finance*, rfw044.

Bikbov, R. and M. Chernov (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics 159*(1), 166–182.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies 22*(9), 3411–3447.

Brooks, J. and T. J. Moskowitz (2017). The cross-section of government bond returns.

Campbell, J. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies 1*(3), 195–228.

Campbell, J. Y. and J. H. Cochrane (1999). By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy 107*(2), pp. 205–251.

Campbell, J. Y. and A. S. Kyle (1993). Smart money, Noise Trading and Stock Price Behaviour. *The Review of Economic Studies 60*(1), 1–34.

Campbell, J. Y., C. Polk, and T. Vuolteenaho (2009). Growth or glamour? fundamentals and systematic risk in stock returns. *The Review of Financial Studies 23*(1), 305–344.

Campbell, J. Y. and S. B. Thompson (2007). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies 21*(4), 1509–1531.

Cieslak, A. and P. Povala (2015). Expected returns in Treasury bonds. *Review of Financial Studies*, hhv032.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies 21*(4), 1533–1575.

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

Cochrane, J. H. and M. Piazzesi (2005). Bond Risk Premia. *American Economic Review*, 138–160.

Cochrane, J. H. and M. Piazzesi (2008). Decomposing the yield curve. *Graduate School of Business, University of Chicago, Working Paper*.

Cochrane, J. H. and J. Saa-Requejo (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy 108*(1), 79–119.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003). The value spread. *The Journal of Finance 58*(2), 609–641.

Cooper, I. and R. Priestley (2008). Time-varying risk premiums and the output gap. *The Review of Financial Studies 22*(7), 2801–2833.

Duffee, G. R. (2013). *Bond Pricing and the Macroeconomy*, Volume 2, Chapter Vol. 2, Part B, pp. 907–967. Elsevier.

Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics 30*(2), 212–228.

Estrella, A. and F. S. Mishkin (1997). The predictive power of the term structure of interest rates in europe and the united states: Implications for the european central bank. *European economic review 41*(7), 1375–1401.

Evans, C. L. and D. A. Marshall (1998). Monetary policy and the term structure of nominal interest rates: evidence and theory. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 49, pp. 53–111. Elsevier.

Evans, C. L. and D. A. Marshall (2001). Economic determinants of the term structure of nominal interest rates. *manuscript, Federal Reserve Bank of Chicago*.

Fama, E. and R. Bliss (1987). The information in long-maturity forward rates. *The American Economic Review*, 680–692.

Fama, E. and K. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22*(1), 3–25.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stock and bonds. *Journal of Financial Economics 33*, 3–56.

Fama, E. F. and K. R. French (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies 29*(1), 69–103.

Giglio, S. and D. Xiu (2017). Inference on risk premia in the presence of omitted factors. Technical report, National Bureau of Economic Research.

Greenwood, R. and S. G. Hanson (2012). Share issuance and factor timing. *The Journal of Finance 67*(2), 761–798.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2006, June). The U.S. Treasury Yield Curve: 1961 to the Present.

Hansen, L. P. and R. Jagannathan (1991). Implications of Security Market Data for Models of Dynamic Economies. *Journal of political economy 99*(2), 225–262.

Harvey, C. R. and Y. Liu (2016). Lucky factors. Technical report.

Hou, K., C. Xue, and L. Zhang (2017). Replicating anomalies. Technical report, National Bureau of Economic Research.

Ilmanen, A. and L. N. Nielsen (2015). Are defensive stocks expensive? a closer look at value spreads.

Joslin, S., M. Priebsch, and K. J. Singleton (2014). Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. *The Journal of Finance 69*(3), 1197–1233.

Kelly, B. T., S. Pruitt, and Y. Su (2017). Some characteristics are risk exposures, and the rest are irrelevant.

Koijen, R. S. and S. Van Nieuwerburgh (2011). Predictability of returns and cash flows. *Annu. Rev. Financ. Econ. 3*(1), 467–491.

Kozak, S., S. Nagel, and S. Santosh (2017). Interpreting factor models. *The Journal of Finance, forthcoming.*

Kozak, S., S. Nagel, and S. Santosh (2018). Shrinking the cross-section.

Lettau, M. and S. Ludvigson (2001). Consumption, Aggregate Wealth, and Expected Stock Returns. *Journal of Finance 56*, 815–849.

Lettau, M., M. Maggiori, and M. Weber (2014). Conditional risk premia in currency markets and other asset classes. *Journal of Financial Economics 114*(2), 197–225.

Lochstoer, L. A. and P. C. Tetlock (2016). What drives anomaly returns?

Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. *The Review of Financial Studies 22*(12), 5027–5067.

Lustig, H., N. Roussanov, and A. Verdelhan (2011). Common risk factors in currency markets. *Review of Financial Studies 24*(11), 3731–3777.

Lustig, H., N. Roussanov, and A. Verdelhan (2014). Countercyclical currency risk premia. *Journal of Financial Economics 111*(3), 527–553.

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

Moreira, A. and T. Muir (2017). Volatility-managed portfolios. *The Journal of Finance 72*(4), 1611–1644.

Nagel, S. and K. J. Singleton (2011). Estimation and Evaluation of Conditional Asset Pricing Models. *The Journal of Finance 66*(3), 873–909.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary accounting research 11*(2), 661–687.

Ready, R., N. Roussanov, and C. Ward (2017). Commodity trade and the carry trade: A tale of two countries. *The Journal of Finance*.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341–360.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review 71*, 321 – 436.

Stambaugh, R. F., J. Yu, and Y. Yuan (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics 104*(2), 288–302.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*(460), 1167–1179.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance 57*(1), 233–264.

Appendix

A Statistical Properties

Our approach to the predictability of cross-sections of returns is focused on predicting important dimensions of the data rather than considering regressions at the individual asset level. In this section, we study more systematically the relation between predicting important components of returns and predicting individual returns.

We consider three features that were relevant in our empirical applications and provide ways to quantify them more generally. First, there is a strong link between predicting the first principal component of returns and predicting each individual return. Second, it is difficult to detect predictability of the second or higher components of returns in individual regressions when the first component is large. Third, joint tests of significance in individual regressions are susceptible to picking up small unimportant patterns of predictability. All derivations are in Appendix **??**.

A.1 First Principal Component and Individual Regressions

A common empirical situation is that a family of returns $\{R_{i,t+1}\}_{i\in I}$ has a strong common component F_{t+1}. When this component is predictable by a variable X_t, does this imply that the individual returns are predictable by X_t? We answer this question quantitatively by deriving a series of bounds linking the predictability of F_{t+1} with the individual predictability of asset returns. We first zoom in on one particular return before considering properties for an entire family of returns.

One individual return: a purely statistical bound. Define $\mathrm{R}^2_{1,i}$ as the population R-squared of the contemporaneous regression of an individual asset on the common component,

$$R_{i,t+1} = \lambda_i F_{t+1} + \varepsilon_{i,t+1}, \tag{11}$$

and R^2_X as the R-squared of the predictive regression of the factor,

$$F_{t+1} = \beta_1 X_t + u_{t+1}. \tag{12}$$

We are interested in $\mathrm{R}^2_{X,i}$, the R-squared of the predictive regression

$$R_{i,t+1} = b_i X_t + v_{t+1}. \tag{13}$$

The following proposition characterizes a lower bound on this quantity.[42]

Proposition 2. *If a variable X_t predicts a factor F_{t+1} with R-squared R^2_X and an individual return is explained by this factor with R-squared $R^2_{1,i}$, then a lower bound for the R-squared $R^2_{X,i}$ of predicting this return using X_t is given by:*

$$R^2_{X,i} \geq \max\left(\sqrt{R^2_{1,i}R^2_X} - \sqrt{\left(1-R^2_{1,i}\right)\left(1-R^2_X\right)}, 0\right)^2. \tag{14}$$

[42] Without loss of generality, we assume that the predictor X_t has unit variance.

Proof. By the definition of a regression R^2 we have $\mathrm{R}^2_{1,i} = \frac{\lambda_i^2 \mathrm{var}(F_{t+1})}{\mathrm{var}(R_{i,t+1})}$, $\mathrm{R}^2_X = \frac{\beta_1^2}{\mathrm{var}(F_{t+1})}$, and $\mathrm{R}^2_{X,i} = \frac{b_i^2}{\mathrm{var}(R_{i,t+1})}$. The the linearity of regression we have $b_i = \lambda_i \beta_1 + \mathrm{cov}\left(X_t, u_{i,t+1}\right)$. We can bound the second term in this expression:

$$\begin{aligned} \left|\mathrm{cov}\left(X_t, u_{i,t+1}\right)\right| &= \left|\mathrm{corr}\left(X_t, u_{i,t+1}\right)\right| \sqrt{\mathrm{var}\left(u_{i,t+1}\right)} \\ &\leq \sqrt{1 - \mathrm{R}^2_{1,i}} \sqrt{\mathrm{var}\left(u_{i,t+1}\right)}, \end{aligned}$$

where the bound comes from the fact that the correlation matrix of $u_{i,t+1}$, F_{t+1} and X_{t+1} has to be semidefinite positive and therefore have a positive determinant.

If $|\lambda_i \beta_i| \leq \sqrt{1 - \mathrm{R}^2_{1,i}} \sqrt{\mathrm{var}\left(u_{i,t+1}\right)}$, then 0 is a lower bound for $\mathrm{R}^2_{X,i}$. In the other case, we obtain the following bound:

$$\begin{aligned} \mathrm{R}^2_{X,i} &\geq \frac{\left(\lambda_i \beta_1 - \sqrt{1 - \mathrm{R}^2_{1,i}} \mathrm{var}\left(u_{i,t+1}\right)\right)^2}{\mathrm{var}\left(R_{i,t+1}\right)} \\ &\geq \left(\sqrt{\frac{\lambda_i^2 \beta_1^2}{\mathrm{var}\left(R_{i,t+1}\right)}} - \sqrt{1 - \mathrm{R}^2_{1,i}} \sqrt{\frac{\mathrm{var}\left(u_{i,t+1}\right)}{\mathrm{var}\left(R_{i,t+1}\right)}}\right)^2 \\ &\geq \left(\sqrt{\mathrm{R}^2_{1,i} \mathrm{R}^2_X} - \sqrt{\left(1 - \mathrm{R}^2_{1,i}\right)\left(1 - \mathrm{R}^2_X\right)}\right)^2 \end{aligned}$$

Putting the two cases together gives Equation 14. □

Intuitively, if X_t strongly predicts the common factor, and the factor has high explanatory power for individual returns, then X_t should predict the individual returns as well. The bound is indeed increasing in the R-squared of these two steps. However, it is lower than the product of the two R-squared — a naive guess that assumes "transitivity" of predictability. This is because the predictor X_t might also predict the residual $\varepsilon_{i,t+1}$ in a way that offsets the predictability coming from the factor. The orthogonality of F_{t+1} and $\varepsilon_{i,t+1}$ limits this force, but does not eliminate it.

To get a quantitative sense of the tightness of this bound, consider the case of bond returns. The level factor explains about 90% of the variation in individual returns, and it can be predicted with an R-squared around 25%. Plugging into our bound, this implies a predictive R-squared of at least 4% for a typical individual bond return. This is a sizable number, but also much less than the 22.5% implied by a naive approach.

One individual return: a bound with an economic restriction. One reason this bound is relatively lax is that it does not take into account the nature of the variable $\varepsilon_{i,t+1}$. Indeed, if, as is the case in our setting, the component F_{t+1} is itself an excess return, the residual $\varepsilon_{i,t+1}$ is one too. It is therefore natural to make the economic assumption that it cannot be "too"' predictable by the variable X_t. This corresponds to imposing an upper bound $\mathrm{R}^2_{\max}$ on the R-squared of the predictive regression of $\varepsilon_{i,t+1}$ by X_{t+1}.[43] In this case, our bound becomes:

$$\mathrm{R}^2_{X,i} \geq \max\left(\sqrt{\mathrm{R}^2_{1,i}\mathrm{R}^2_X} - \sqrt{\mathrm{R}^2_{\max}\left(1-\mathrm{R}^2_X\right)}, 0\right)^2. \quad (15)$$

Such an approach can considerably tighten the bound. For instance, in our example for treasuries, one could impose an upper bound of 25% for predicting the residual. This yields a lower bound on predicting the return $R_{i,t+1}$ of 10%, a much larger number, statistically and economically.

Family of returns: the symmetric case. Another reason that predictability of the common factor must transmit to predictability of individual returns is that by design it absorbs common variation across all those returns. To highlight this point, we consider the following simple symmetric case. We assume that the factor is the average of all the individual returns, $F_{t+1} = \frac{1}{N}\sum_i R_{i,t+1}$. We further assume that all assets have the same loading on the factor and the factor has the same explanatory power for each return. This corresponds to constant λ_i, and R^2_{1i} across assets. We then immediately have:

$$\sum_i u_{i,t+1} = 0$$
$$\sum_i \mathrm{cov}\left(X_t, u_{i,t+1}\right) = 0.$$

Letting $\gamma_i = \mathrm{cov}\left(X_t, u_{i,t+1}\right)$ we then obtain an expression for an individual asset:

$$\begin{aligned}\mathrm{R}^2_{X,i} &= \frac{\left(\lambda_i\beta_1 + \gamma_i\right)^2}{\mathrm{var}\left(R_{i,t+1}\right)} \\ &= \mathrm{R}^2_1\mathrm{R}^2_X + \frac{\gamma_i^2}{\mathrm{var}\left(R_{i,t+1}\right)} + 2\gamma_i\frac{\lambda_i\beta_1}{\mathrm{var}\left(R_{i,t+1}\right)}\end{aligned}$$

Finally, taking averages across assets we have:

$$\mathbb{E}_i\left[\mathrm{R}^2_{X,i}\right] = \mathrm{R}^2_1\mathrm{R}^2_X + \mathrm{var}_i\left(\mathrm{R}^2_{X,i}\right), \quad (16)$$

where $\mathbb{E}_i(\cdot)$ and $\mathrm{var}_i(\cdot)$ are the mean and variance in the cross section of individual returns and we use the fact that we use the fact that $\mathbb{E}_i\left[\gamma_i\right] = 0$. This formula implies that the average explanatory power is now at least as large as given by the transitive formula. This would correspond to 22.5% in our example, almost the same value as the predictive R-squared for the common factor. Furthermore, the more unequal this predictive power is across assets, the stronger it must be on average. That is, if the variable X_t does less well than the transitive R-squared for some particular returns, it must compensate more than one-to-one for the other assets.

From predicting "everything" to aggregate returns. Maintaining the same assumptions, we can rearrange Equation 16 to see what the predictability of "everything" implies for

[43]One way to determine a reasonable bound on $\mathrm{R}^2_{\max}$ is to note that the standard deviation of an asset's conditional Sharpe ratio equals $\sqrt{\frac{\mathrm{R}^2_{\mathrm{x,i}}}{1-\mathrm{R}^2_{x,i}}}$.

predictability of the common factor. We have:

$$\mathrm{R}_X^2 = \frac{\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right] - \mathrm{var}_i\left(\mathrm{R}_{X,i}^2\right)}{\mathrm{R}_1^2}.$$

At first this may not seem very powerful since $\mathrm{var}_i\left(\mathrm{R}_{X,i}^2\right)$ could be large. This maximal variance, however, is related to the average $\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right]$. Consider the simple example of only two assets. Then, if the average $\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right]$ is 10%, the maximal variance is only 1%, which obtains when $\mathrm{R}_{X,1}^2 = 0\%$ and $\mathrm{R}_{X,2}^2 = 20\%$. In general with two assets we have

$$\mathrm{var}_i\left(\mathrm{R}_{X,i}^2\right) \leq \left(0.5 - \left|\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right] - 0.5\right|\right)^2$$

which gives the bound

$$\mathrm{R}_X^2 \geq \frac{\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right] - \left(0.5 - \left|\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right] - 0.5\right|\right)^2}{\mathrm{R}_1^2}.$$

For large N, the Bhatia-Davis inequality gives:

$$\mathrm{R}_X^2 \geq \frac{\left(1 - \mathrm{R}_{\max}^2\right)\mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right] + \mathbb{E}_i\left[\mathrm{R}_{X,i}^2\right]^2}{\mathrm{R}_1^2},$$

where $\mathrm{R}_{\max}^2$, as before, is the maximum $\mathrm{R}_{X,i}^2$ from any individual asset forecasting regression. For reasonable values of $\mathrm{R}_{\max}^2$, such as 0.5 or less, the bound implies that ~22% average R^2 we obtain for individual bonds implies at least 18% R_X^2, the R-squared when predicting the aggregate portfolio return.

A.2 Low Power of Individual Tests

While individual regressions are strongly related to predicting the first common component of returns, they can face challenges in detecting predictability of other factors. We provide a way to quantify this issue by characterizing the statistical power of a test of significance for a predictor that only predicts one particular component of returns.

I.i.d. predictor. Consider first the case where the forecasting variable X_{t+1} has i.i.d. draws.[44] Suppose that X_t forecasts only one particular principal component j with population R-squared R_X^2 and the remaining principal component returns are i.i.d. Gaussian with known mean.[45] For power analysis, we consider repeated samples of length T.[46]

[44]The formulas hereafter admit simple generalizations to multivariate prediction.

[45]More generally, the components need not be principal components. They must be uncorrelated and only one particular component must be forecastable by our predictor. If the mean is unknown, the results below are unchanged except that the degrees of freedom are $T-1$ instead of T.

[46]The analysis treats X as stochastic. With fixed X the distribution is normal instead of a Student t.

When directly forecasting the principal component return, $F_{j,t+1}$, the power to correctly reject the null with test of nominal size α is

$$\text{power}\left(F_2\right) = G\left(-t_{\alpha/2,T} - z\right) + \left[1 - G\left(t_{\alpha/2,T} - z\right)\right], \tag{17}$$

where G is the CDF of a t-distribution with T degrees of freedom, $z = \sqrt{\mathrm{R}_X^2}\sqrt{T}\left(1 - \mathrm{R}_X^2\right)^{-\frac{1}{2}}$, and $t_{\alpha/2,T}$ is the $\frac{\alpha}{2}$ critical value from the t-distribution.

In contrast, when directly forecasting an individual return, $R_{i,t+1}$, the power is

$$\text{power}\left(R_i\right) = G\left(-t_{\alpha/2,T} - \zeta\right) + \left(1 - G\left(t_{\alpha/2,T} - \zeta\right)\right), \tag{18}$$

where $\zeta = \sqrt{\mathrm{R}_X^2}\sqrt{T}\left(\left(1 - \mathrm{R}_X^2\right) + \frac{1-\mathrm{R}_{j,i}^2}{\mathrm{R}_{ji}^2}\right)^{-\frac{1}{2}}$. By symmetry of the t-distribution and because $\zeta \leq z$, we immediately obtain that power (F_2) is larger than power (R_i) for all assets. Therefore, there is always more information about predictability of the important component by studying it directly.

Persistent predictor. To understand whether these results are a useful approximation for the more general case of a persistent predictor, we turn to simulation. We model X as an $AR(1)$ process,

$$X_{t+1} = \phi X_t + \nu_{t+1},$$

with the normalization $\text{var}_t\left[\nu_{t+1}\right] = \frac{1}{1-\phi^2}$, so that X has unit unconditional variance. Simulated returns have the same unconditional covariance as in the bond data, but we assume the second principal component is forecastable by X. We simulate 30-year histories and forecast returns of PCs and individual bonds in each simulated sample. We compute the sampling variance of an estimator from the simulated distribution and construct relevant t-statistics.[47]

Table 9 shows the probability of rejecting the null of no predictability under the true distribution. Panels A, B, and C correspond to a persistence of the predictor, ϕ, equal to 0, 0.3 and 0.6, respectively. Each row corresponds to a simulations with the indicated value of R_X^2, the population R-squared obtained when forecasting the second PC of returns with X_t. The first column shows the probability of rejecting the null when forecasting the first principal component of returns. Since it is not predictable by construction, the rejection probability should be 5% for a test with that nominal size. The simulated rejection probabilities are all close to 5%, indicating the t-test has approximately correct rejection probability even when the predictor is persistent. In fact, the values are very similar across panels: the t-distribution provides a close approximation even when not exact. For individual bonds, we see that power is much lower than for the second principal component for all values of the persistence parameters.[48] These results suggest that the closed form formulas for the i.i.d. case (panel A) constitute a good approximation for settings with persistent predictors.

[47]We assume the researcher knows the sampling variance. This circumvents known small-sample issues with HAC variance estimators.

[48]Across simulations, the power is U-shaped in maturity. This is to be expected for our particular setting since squared loadings on the second principal component are also u-shaped.

Table 9: Power of Predictive Regressions when Only Slope is Predictable

The table gives probability (in %) of rejecting the null hypothesis of no predictability given the alternative hypothesis. ϕ is the annual auto-correlation of the predictive variable. Each row gives results for the indicated theoretical R^2 when forecasting Slope. Level and Slope are the first and second principal components, respectively, of bond returns. The remaining columns are for zero-coupon bonds with the indicated maturity (in years). We compute the probabilities from 100,000 simulations, each of thirty years.

(a) No Persistence, $\phi = 0$

R^2	Level	Slope	3Y	6Y	9Y	12Y	15Y
10%	5	39	9	5	6	8	10
20%	5	74	14	5	7	12	16
30%	5	93	19	5	8	15	21
40%	5	99	25	6	9	18	26

(b) $\phi = 0.3$

R^2	Level	Slope	3Y	6Y	9Y	12Y	15Y
10%	6	40	10	6	7	9	11
20%	6	73	15	6	8	12	16
30%	6	92	20	6	8	15	21
40%	6	99	25	6	9	19	27

(c) $\phi = 0.6$

R^2	Level	Slope	3Y	6Y	9Y	12Y	15Y
10%	7	41	12	7	8	10	11
20%	7	74	16	7	9	14	17
30%	7	91	21	8	10	17	23
40%	7	98	27	8	11	20	29

Table 10: Predicting PC Returns with Noise

The table gives the population R^2 and probability (in %) of rejecting the null of no predictability when forecasting principal components of bond returns with lagged bond yields. The last row reports the size of a Wald test over the first few principal components. We compute these values from 100,000 simulations of 30 years, with yields contaminated by i.i.d. noise with 5bp standard deviation.

	$PC1$	$PC2$	$PC3$	$PC4$
population R^2 (%)	0	4	40	46
5% Wald test size (%)	6	11	70	83
Joint 5% Wald test size (%)	6	10	51	75

A.3 Spurious Predictability in Joint Tests

Finally, we consider the issue of spurious predictability likely to be picked up by the Wald test. The Wald test assesses joint significance of the predictive coefficients for individual returns b_i, or equivalently joint significance of the predictive coefficients for all the principal components of the family β_i. Very small components are likely to pick up spurious patterns of predictability, or predictability originating from minor measurement error in prices. This would lead to a rejection of the Wald test, even though there is no economically interesting predictability.

To get a sense of the importance of this issue, we study a simple simulation. We present here an application to the case of Treasury bond returns, but this approach can be adapted to other settings.[49] In the simulation, we introduce a tiny amount of noise in prices, i.e. yields, then predict bond returns. We assume there is no true predictability but observed yields, $\widetilde{y}$, have i.i.d noise with standard deviation of σ_ε:

$$\widetilde{y}_{n,\tau} = y_{n,\tau} + \varepsilon_{n,\tau}.$$

We construct observed returns from observed yields,

$$\begin{aligned} \widetilde{r}_{n,\tau+1} &= -\left(n-1\right)\widetilde{y}_{n-1,\tau} + n\widetilde{y}_{n,\tau} - \widetilde{y}_{1,\tau} \\ &= r_{n,\tau+1} - \left(n-1\right)\varepsilon_{n-1,\tau+1} + n\varepsilon_{n,\tau} - \varepsilon_{1,\tau}, \end{aligned} \tag{19}$$

and simulate 30-year histories of "observed" returns based on the sample covariance matrix of realized bond returns and set $\sigma_\varepsilon = 5$bp. For context, annual yield changes have approximately 1% standard deviation. We then consider forecasting returns in the presence of these errors. For each PC portfolio we compute (via simulation) the true size of a nominal 5% Wald test of the null hypothesis of no predictability. Table 10 shows the population R^2 of the predictive regression and probability of rejecting the null for each principal component. The first two PC portfolios have R-squared close to zero, and the test size is somewhat higher than the nominal 5%. The second two PCs, on the other hand, have large R-squared and the size of the test is of an order of magnitude larger than 5%. The last row shows the rejection probability for a test that "large" PCs are not

[49]For comparison with previous papers such as Cochrane and Piazzesi (2005), we use yields on bonds with maturities from one to five years constructed as in Fama and Bliss (1987),

predictable, that is a Wald test over the first few principal components. For example, the second column gives the rejection probability for a joint test that PC1 and PC2 are not predictable. The high individual test size for the smaller PCs contaminates this joint Wald test. Indeed, while the size of the test is 10% for the first couple of PCs, it then jumps up and is as high as 75% for the first four PCs. This result exemplifies well the issue with the Wald test: very small, economically meaningless variation in prices tends to get captured in small principal components and generates uninteresting or spurious predictability. Because the Wald test puts them on the same footing as the larger, more interesting sources of variation, it tends to reject the null too often.

B Internet Appendix: Additional Results

B.1 Stocks

Table 11: Predicting conditional variances of dominant components with BE/ME ratios

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of squared errors from excess market returns and two PCs of long-short anomaly returns regressions on three predictors: (i) log of the aggregate BE/ME ($\overline{bm}$), (ii) a restricted linear combination of anomalies' log BE/ME ratios with weights given by the first eigenvector of pooled long-short strategy returns (bm_1); and (iii) a restricted linear combination of anomalies' log BE/ME ratios with weights given by the second eigenvector of pooled long-short strategy returns (bm_2). The first three columns show results from an unrestricted regression where each return is forecast using all three predictors. The next three columns show results from imposing a diagonal structure; that is, each return in only predicted by its own bm ratio. The last three rows show regression R^2, out-of-sample R^2, and p-value of the Wald test of joint significance of all regression coefficients. Circular block bootstrapped standard errors in parentheses.

	MKT	PC1	PC2	MKT	PC1	PC2
$\overline{bm}$	-0.009	-0.005	-0.044	-0.012	-	-
	(0.68)	(0.30)	(3.24)	(0.85)		
bm_1	-0.003	0.003	-0.009	-	-0.003	-
	(0.55)	(0.46)	(1.79)		(0.55)	
bm_2	0.000	-0.002	0.017	-	-	0.020
	(0.03)	(0.23)	(2.69)			(2.38)
R^2	0.012	0.006	0.315	0.008	0.003	0.119
OOS R^2	-0.115	-0.260	-0.019	-0.006	-0.011	-0.016
Wald test p-value	0.902	0.987	0.000	0.698	0.857	0.059

As discussed in Moreira and Muir (2017), optimal timing strategies rely not only on estimates of conditional expected returns, but also conditional volatilities. We now assess whether the return predictability we uncover using bm ratios carries over to predictability of variances. To measure this, we first forecast returns as in Table 2, compute squared forecast errors, and then forecast these using bm ratios.[50] Table 11 shows the estimation results; the first three columns allow returns and variances to depend on all three bm ratios and the next three impose a diagonal structure. For market and PC1, we find little evidence of variance predicatiblity, in- or out-of-sample. For PC2, the estimates show some heteroskedasticity captured by bm ratios; this is, however, not robust, as evidence by the nearly zero OOS R^2. Given this finding, we assume that $\Sigma_{Z,t}$ is constant.

[50]This second step include a generated regressand, which can induce spurious small sample predictability.

Table 12: Predicting PC returns of all anomalies with BE/ME ratios (OOS PCs)

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of excess market returns and two PCs of long-short anomaly returns on three predictors: (i) log of the aggregate BE/ME ($\overline{bm}$), (ii) a restricted linear combination of anomalies' log BE/ME ratios with weights given by the first eigenvector of pooled long-short strategy returns (bm_1); and (iii) a restricted linear combination of anomalies' log BE/ME ratios with weights given by the second eigenvector of pooled long-short strategy returns (bm_2). Eigenvectors are estimated in the first part of the sample. The first three columns show results from an unrestricted regression where each return is forecast using all three predictors. The next three columns show results from imposing a diagonal structure; that is, each return in only predicted by its own bm ratio. The last three rows show regression R^2, out-of-sample R^2, and p-value of the Wald test of joint significance of all regression coefficients. Circular block bootstrapped standard errors in parentheses.

	MKT	PC1	PC2	MKT	PC1	PC2
$\overline{bm}$	0.031	0.039	-0.017	0.027	-	-
	(1.75)	(2.27)	(1.31)	(1.66)		
bm_1	-0.002	0.027	-0.003	-	0.025	-
	(0.32)	(4.48)	(0.69)		(3.63)	
bm_2	-0.011	-0.001	0.030	-	-	0.028
	(1.41)	(0.15)	(4.26)			(4.01)
R^2	0.076	0.358	0.193	0.044	0.272	0.174
OOS R^2	0.086	0.117	0.083	0.136	0.233	0.210
Wald test p-value	0.324	0.000	0.001	0.251	0.001	0.000

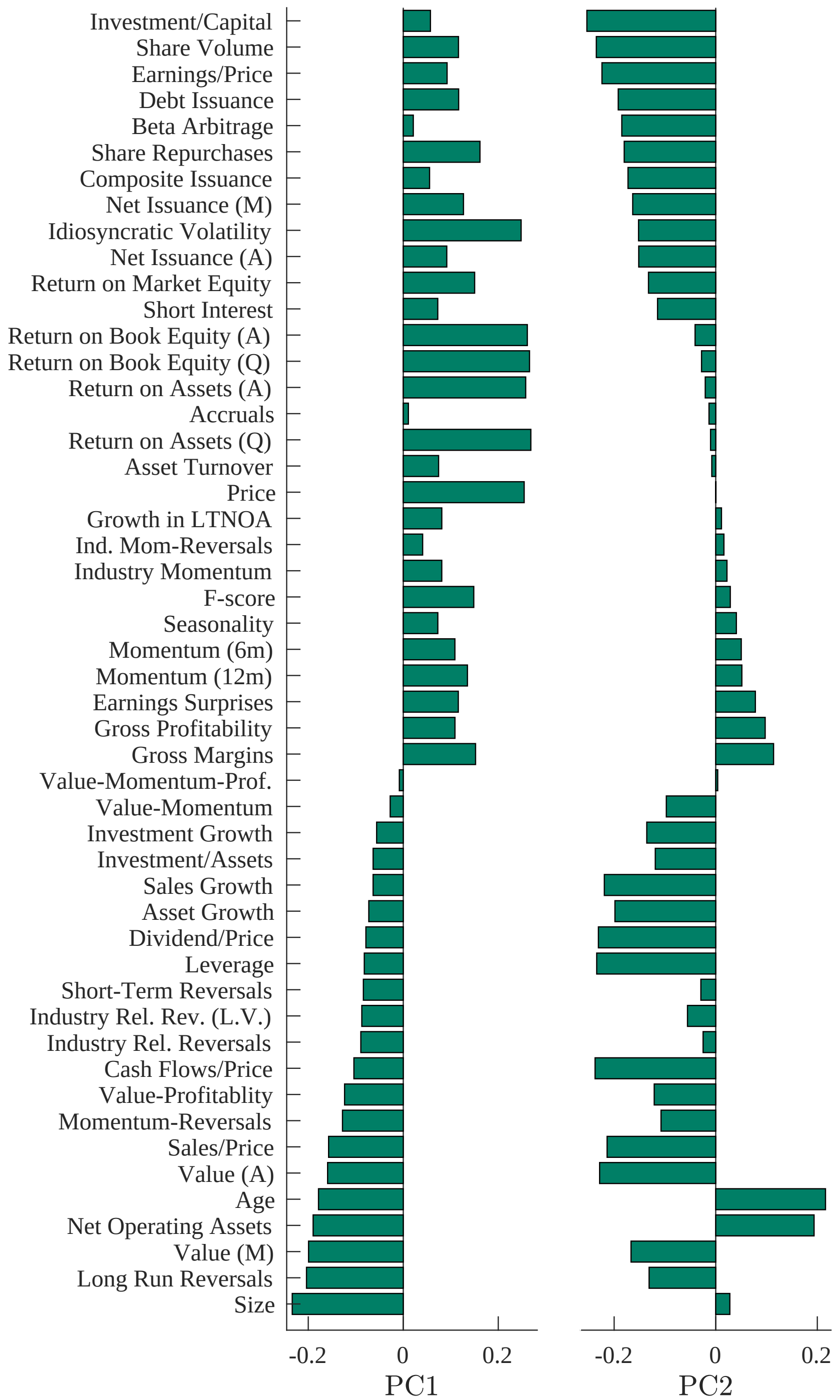


Figure 8: Anomalies Eigenvector Loadings. The figure plots eigenvector loadings of 50 long-short anomalies.

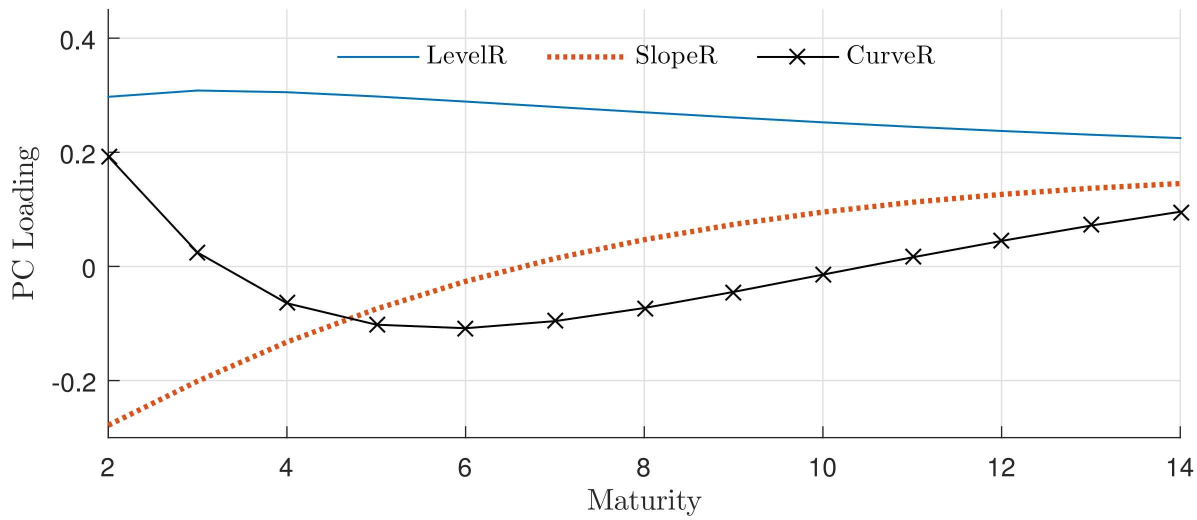
Table 13: Part I: Anomaly portfolios mean excess returns, %, annualized

Columns P1 through P10 show mean annualized returns (in %) on each anomaly portfolio net of risk-free rate. The column P10-P1 lists mean returns on the strategy which is long portfolio 10 and short portfolio 1. Excess returns on beta arbitrage portfolios are scaled by their respective betas. F-score, Debt Issuance, and Share Repurchases are binary sorts; therefore only returns on P1 and P10 are reported for these characteristics. Portfolios include all NYSE, AMEX, and NASDAQ firms; however, the breakpoints use only NYSE firms. Monthly data from November 1973 to December 2015.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-P1
1. Size	5.8	7.7	8.2	8.8	8.6	9.2	8.9	9.5	9.0	8.9	3.2
2. Value (A)	4.9	7.1	7.9	7.1	8.4	8.4	8.7	8.5	8.5	11.9	6.9
3. Gross Profitability	5.4	5.8	6.2	5.7	7.7	6.8	7.1	6.4	7.0	9.2	3.8
4. Value-Profitablity	4.1	5.6	4.2	6.3	8.5	8.0	10.4	11.2	11.3	13.2	9.1
5. F-score	6.1	-	-	-	-	-	-	-	-	7.1	0.9
6. Debt Issuance	6.2	-	-	-	-	-	-	-	-	7.9	1.6
7. Share Repurchases	6.2	-	-	-	-	-	-	-	-	7.7	1.5
8. Net Issuance (A)	2.6	5.0	8.4	7.8	7.8	7.2	6.3	8.7	8.1	11.3	8.7
9. Accruals	3.7	6.1	5.0	6.9	7.0	7.2	8.0	7.2	9.5	8.1	4.4
10. Asset Growth	4.9	6.6	7.1	7.2	6.9	7.6	7.3	8.7	10.0	9.9	4.9
11. Asset Turnover	4.5	6.8	6.0	6.2	7.5	7.8	8.5	6.7	9.5	9.2	4.7
12. Gross Margins	6.7	6.6	7.9	6.9	8.3	6.3	7.2	6.7	5.8	6.8	0.1
13. Dividend/Price	5.5	4.8	6.7	6.9	7.1	9.0	9.5	7.9	7.8	8.6	3.1
14. Earnings/Price	4.1	4.8	6.6	7.3	7.1	7.6	9.4	9.0	8.7	11.6	7.5
15. Cash Flows/Price	4.5	7.3	5.9	7.9	8.4	8.2	7.5	9.3	10.9	10.5	6.0
16. Net Operating Assets	3.1	6.5	6.7	3.5	7.4	7.6	8.2	7.7	8.7	7.9	4.8
17. Investment/Assets	4.2	5.1	7.4	6.3	8.0	6.1	7.9	8.5	8.6	10.2	6.0
18. Investment/Capital	5.8	6.7	6.0	7.1	6.9	8.3	7.8	7.6	8.7	9.1	3.3
19. Investment Growth	4.5	7.9	6.6	6.5	6.2	7.3	7.8	8.0	9.8	8.3	3.8
20. Sales Growth	6.9	6.7	7.0	6.6	7.6	8.7	6.6	7.6	8.8	6.8	-0.1
21. Leverage	5.3	6.4	6.6	10.3	7.3	8.3	8.6	8.6	8.7	8.1	2.8
22. Return on Assets (A)	4.0	8.5	7.3	7.1	7.3	7.1	7.1	7.7	6.1	6.8	2.8

Table 13: Part II: Anomaly portfolios mean excess returns, %, annualized

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-P1
23. Return on Book Equity (A)	5.9	6.9	6.7	7.5	6.3	7.2	6.4	7.0	6.3	7.5	1.6
24. Sales/Price	4.6	6.1	6.6	8.3	8.9	8.5	9.0	10.7	10.7	12.8	8.2
25. Growth in LTNOA	6.5	6.3	6.7	8.2	5.9	6.9	6.4	7.7	7.8	7.7	1.1
26. Momentum (6m)	8.9	8.8	8.3	8.5	7.4	8.1	6.6	5.3	7.2	10.6	1.6
27. Value-Momentum	6.1	7.6	6.7	7.2	8.4	9.2	9.7	8.3	8.2	11.0	4.9
28. Value-Momentum-Prof.	5.9	7.6	7.5	8.1	7.1	5.4	7.5	8.5	11.5	14.3	8.4
29. Short Interest	6.3	5.7	8.5	8.5	8.1	6.1	6.8	5.9	4.3	5.2	-1.1
30. Momentum (12m)	-2.2	4.2	5.6	6.8	5.5	6.9	6.9	9.1	9.0	12.4	14.6
31. Industry Momentum	5.7	5.2	7.0	6.0	7.1	9.6	7.6	6.7	8.7	9.4	3.7
32. Momentum-Reversals	5.1	6.8	7.2	6.7	7.5	8.6	7.2	8.9	8.7	11.4	6.3
33. Long Run Reversals	6.6	6.9	7.3	8.1	8.1	8.5	8.3	9.5	9.8	11.1	4.5
34. Value (M)	5.5	6.2	6.4	6.9	8.0	7.5	9.0	7.2	11.6	11.5	6.0
35. Net Issuance (M)	3.6	5.2	10.0	8.0	8.7	7.4	7.4	8.0	9.9	10.5	6.9
36. Earnings Surprises	4.1	3.9	4.8	7.4	6.7	7.6	7.2	7.4	8.1	10.8	6.7
37. Return on Book Equity (Q)	1.4	5.6	6.5	4.7	5.6	6.3	7.5	7.2	7.2	8.9	7.5
38. Return on Market Equity	-0.1	1.6	6.3	5.9	7.2	7.1	8.0	10.6	11.4	15.5	15.5
39. Return on Assets (Q)	1.8	4.5	6.8	7.2	7.2	6.7	8.0	7.4	6.8	7.5	5.8
40. Short-Term Reversals	3.0	4.3	6.8	6.5	6.9	7.6	8.5	9.1	9.6	7.2	4.2
41. Idiosyncratic Volatility	-0.5	7.8	10.6	7.6	9.7	8.5	7.6	7.3	7.2	6.9	7.4
42. Beta Arbitrage	3.1	3.1	4.1	6.6	8.0	9.4	10.7	11.4	14.2	16.7	13.6
43. Seasonality	3.0	3.4	5.7	5.2	7.6	6.7	7.6	7.1	9.3	12.8	9.8
44. Industry Rel. Reversals	1.5	3.1	4.1	6.0	6.2	7.5	8.6	11.1	12.5	12.3	10.8
45. Industry Rel. Rev. (L.V.)	0.6	4.2	4.5	6.7	5.8	6.5	9.0	10.5	13.2	14.9	14.3
46. Ind. Mom-Reversals	2.7	4.4	5.6	5.6	7.5	6.9	8.6	8.8	10.3	14.3	11.5
47. Composite Issuance	4.0	5.6	5.9	6.1	7.4	7.2	6.9	7.6	9.9	10.2	6.3
48. Price	5.5	8.1	8.4	9.7	8.3	8.2	7.1	7.1	7.3	5.8	0.3
49. Age	6.5	7.9	5.5	9.9	5.5	7.7	9.4	7.2	6.8	6.8	0.2
50. Share Volume	6.3	7.8	6.6	6.4	7.5	6.0	7.6	6.6	6.2	6.2	-0.1



	LevelR	SlopeR	CurveR
% of variance	90.1	9.3	0.5

Figure 9: Factor Structure in Realized Returns. The top panel plots the first three eigenvectors of realized zero-coupon bond excess returns, termed LevelR, SlopeR, and CurveR. The bottom panel shows the percent of total variance contributed by each factor.

B.2 Bonds

Table 15: Bond Summary Statistics

The table shows mean, standard deviation, skewness and Sharpe ratio of returns on the dominant components of bonds (LevelR and SlopeR). The first two columns are for static investment strategies and next three are for pure timing strategies.

	Static portfolios		Pure timing portfolios	
	LevelR	SlopeR	LevelR	SlopeR
Mean (%)	3.11	-0.73	0.06	0.09
Std. Dev. (%)	5.00	5.00	0.16	0.15
Skewness	0.16	-0.00	2.07	1.10
Sharpe Ratio	0.62	-0.15	0.35	0.60

Table 14: Part I: Implied anomaly returns by PC: R^2 (%)

Predictive R^2 of individual anomalies returns implied by PC forecasts. Columns labeled "Full" combine forecasts of both PCs; "PC1/PC2 only" focus only on predictability stemming from PC1/PC2.

	IS (Full)	IS (PC1 only)	IS (PC2 only)	OOS (Full)	OOS (PC1 only)	OOS (PC2 only)
1. Size	22.6	25.5	-3.5	33.8	31.2	2.0
2. Value (A)	20.2	16.7	4.7	11.6	16.7	-3.7
3. Gross Profitability	-2.8	-0.6	-1.9	-45.1	-39.3	-5.4
4. Value-Profitablity	12.1	5.4	7.0	20.8	16.6	4.6
5. F-score	6.4	5.7	0.5	-5.9	-9.1	2.9
6. Debt Issuance	10.6	1.8	8.4	12.0	0.0	11.5
7. Share Repurchases	25.6	1.7	23.1	11.6	-13.5	24.3
8. Net Issuance (A)	17.7	-0.9	18.3	8.6	-7.1	15.5
9. Accruals	0.3	-0.1	0.4	-0.1	-0.1	-0.0
10. Asset Growth	14.4	10.3	4.7	21.9	19.3	3.3
11. Asset Turnover	2.9	-0.6	3.4	-3.1	-7.0	3.8
12. Gross Margins	6.8	5.7	1.4	-11.4	-10.3	-0.6
13. Dividend/Price	13.5	4.5	10.0	14.7	8.5	7.5
14. Earnings/Price	12.3	0.1	12.1	15.9	-0.5	16.3
15. Cash Flows/Price	11.4	11.2	1.4	3.9	10.1	-4.5
16. Net Operating Assets	-1.7	1.3	-2.9	-13.9	-9.6	-4.2
17. Investment/Assets	15.7	11.3	4.7	17.0	21.9	-4.3
18. Investment/Capital	-1.3	0.4	-1.5	5.0	0.6	4.5
19. Investment Growth	14.9	10.5	4.8	16.2	15.8	0.8
20. Sales Growth	5.8	6.7	-0.1	10.0	9.5	1.4
21. Leverage	23.7	8.8	15.7	27.5	17.3	10.9
22. Return on Assets (A)	24.1	16.9	6.4	3.9	-10.1	13.1
23. Return on Book Equity (A)	21.7	15.9	4.9	1.9	-15.7	16.4
24. Sales/Price	18.2	9.2	9.8	18.4	12.7	6.7
25. Growth in LTNOA	-7.4	-6.8	-0.7	-0.5	0.3	-0.8
26. Momentum (6m)	13.4	14.0	-0.6	5.3	6.2	-0.8
27. Value-Momentum	3.4	-0.1	3.8	4.5	3.0	1.7
28. Value-Momentum-Prof.	-0.2	-0.1	-0.1	0.3	0.4	-0.1
29. Short Interest	2.6	3.6	-1.1	10.1	5.7	4.2
30. Momentum (12m)	9.4	9.9	-0.4	7.3	7.7	-0.3
31. Industry Momentum	7.4	6.4	1.0	-7.7	-9.4	1.7
32. Momentum-Reversals	9.4	0.5	9.3	20.7	15.2	5.8
33. Long Run Reversals	29.9	26.4	4.0	30.6	31.6	-0.4
34. Value (M)	32.7	29.9	3.5	26.3	29.6	-2.5
35. Net Issuance (M)	12.2	1.2	10.6	7.5	-11.0	18.0

Table 14: Part II: Implied anomaly returns by PC: R^2 (%)

	IS (Full)	IS (PC1 only)	IS (PC2 only)	OOS (Full)	OOS (PC1 only)	OOS (PC2 only)
36. Earnings Surprises	-8.8	-5.1	-3.5	-8.1	-4.1	-3.7
37. Return on Book Equity (Q)	17.1	13.0	3.1	4.0	-8.0	11.1
38. Return on Market Equity	15.5	-0.6	15.6	9.6	-4.1	13.2
39. Return on Assets (Q)	20.8	13.6	6.3	0.0	-11.8	11.0
40. Short-Term Reversals	6.5	6.7	-0.2	13.6	13.7	-0.0
41. Idiosyncratic Volatility	28.1	14.0	12.7	23.7	-1.1	23.2
42. Beta Arbitrage	7.4	-0.5	8.1	9.0	-1.0	10.2
43. Seasonality	-2.3	-1.5	-0.7	-4.2	-3.7	-0.5
44. Industry Rel. Reversals	8.2	9.7	-1.4	16.8	17.4	-0.6
45. Industry Rel. Rev. (L.V.)	14.9	10.9	4.1	23.0	21.6	1.4
46. Ind. Mom-Reversals	-2.0	-2.7	0.6	-6.6	-6.8	0.2
47. Composite Issuance	2.0	0.1	2.0	7.7	-0.3	7.9
48. Price	35.9	35.4	-0.3	35.8	30.6	4.4
49. Age	16.5	3.7	11.7	18.2	-5.5	22.7
50. Share Volume	4.4	0.6	3.4	12.3	-1.6	13.6
Mean R^2	11.4	7.0	4.4	8.7	3.5	5.3
Median R^2	11.8	5.6	3.5	9.3	0.2	3.1
Std. Dev. of R^2	10.3	8.9	5.9	13.9	14.2	7.7

Table 17: Predicting PC Returns with Forward Spreads and *GRO*

We report predictive coefficients, absolute t-statistics (in parentheses), and R^2s from predictive regressions of the first two principal components of zero-coupon bond excess returns on the the first three principal components of lagged forward spreads and *GRO* (Chicago Fed National Activity Index). Individual bond returns are computed from the Fama-Bliss zero-coupon yields (Fama and Bliss, 1987). Out-of-sample R^2s are calculated using the first half of the sample to estimate coefficients and using these estimates to forecast returns in the second half.

	FS1	FS2	FS3	GRO	R^2
LevelR	0.06	-2.14	-0.92	-1.29	0.26
	(0.10)	(3.37)	(1.45)	(1.80)	
SlopeR	2.68	0.84	-0.71	2.33	0.36
	(4.62)	(1.45)	(1.25)	(3.73)	

Table 16: Predicting Bond Returns with Forward Spreads

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of zero-coupon bond excess returns on the the first three principal components of lagged forward spreads. Each column presents results for the indicated maturity (in years).

	3Y	5Y	7Y	9Y	11Y	13Y	15Y
FS1	0.06	0.19	0.25	0.29	0.30	0.30	0.29
	(0.34)	(1.21)	(1.87)	(2.32)	(2.57)	(2.67)	(2.66)
FS2	-0.43	-0.40	-0.37	-0.34	-0.32	-0.30	-0.29
	(2.59)	(2.79)	(2.95)	(3.01)	(3.02)	(2.99)	(2.96)
FS3	-0.38	-0.32	-0.26	-0.20	-0.15	-0.10	-0.06
	(2.23)	(2.20)	(2.03)	(1.73)	(1.36)	(0.96)	(0.59)
R^2	0.20	0.21	0.23	0.23	0.23	0.23	0.22

B.2.1 The role of other macroeconomic variables

To further explore the link between macroeconomic conditions and bond risk premia, we consider a variety of business cycle variables in addition to GRO. Specifically, we use: (i) PCE, annual real change in per capital personal consumption expenditures; (ii) GDP, annual real change in gross domestic product; (iii) IND, annual change in real industrial production; (iv) LNF1, the first PC of 132 measures of economic activity constructed as in Ludvigson and Ng (2009); and (v) CAY, the consumption/wealth ratio constructed as in Lettau and Ludvigson (2001).[51] Since PCE, GDP, and CAY are only available quarterly, we limit all analysis to that frequency. As above for GRO, we forecast LevelR and SlopeR using the three PCs of forward spreads and each of these measures, one at a time. For ease of comparison, all predictive variables are normalized to have 1% standard deviation. Table 18 shows the estimated coefficients on the business cycle variables as well as in and out-of-sample R^2 statistics. As before, none of the business cycle variables is statistically or economically significant in predicting LevelR. In contrast, all variables besides CAY substantially predict SlopeR, with similar coefficients. Note that LNF1 is essentially a "real" factor that "loads heavily on measure of employment and production ... [but] displays little correlation with prices or financial variables" (Ludvigson and Ng, 2009). Hence, any of the real business cycle measures gives a similar result that SlopeR has low expected returns during downturns. CAY, which is a well-known predictor of equity risk premia, seems to have little power in forecasting either aggregate or relative bond excess returns. These results challenge the view that real variables are not useful to forecast returns across the yield curve, by demonstrating these variables explain a particular important component of this cross-section of returns.[52]

[51] PCE and GDP are from the Bureau of Economic Analysis. IND is from the Federal Reserve Board of Governors. LNF1 and CAY are available at Sydney Ludvigson's website, https://www.sydneyludvigson.com/data-and-appendixes.

[52] Interestingly, our explanation for lack of success in predicting LevelR or individual bond returns does not interact much with those centered around difficulties to draw statistical inference with persistent predictors — e.g. Bauer and Hamilton (2017) — because for SlopeR, both yield-based and macroeconomic predictors

Table 18: Predicting PC Returns with Business Cycle Indicators

We report predictive coefficients and absolute t-statistics (in parentheses) from predictive regressions of the first two principal components of zero-coupon bond excess returns on the the first three principal components of lagged forward spreads and various business cycle indicators. The data is quarterly from 1985-2015. The indicators are the Chicago Fed National Activity Index (GRO), real per-capita annual growth in personal consumption expenditures (PCE), real per-capita annual growth in gross domestic product (GDP), real annual growth in industrial production (IND), the first PC of a broad set of macro variables (LNF1), and the consumption/wealth ratio (CAY). Reported coefficients are β in the forecasting regression:

$$rx_{i,t+1} = a_i + b_i' \mathrm{FS}_t + \beta_i x_t + \varepsilon_{i,t+1},$$

where x_t is the business cycle indicator and FS_t are the first three PCs of forward spreads. Absolute circular block bootstrap t-statistics are in parentheses. R-square values are computed as before.

	LevelR			SlopeR		
	β	IS R^2	OOS R^2	β	IS R^2	OOS R^2
GRO	-0.45	0.22	-2.43	2.39	0.40	0.15
	(0.59)			(3.56)		
PCE	-0.62	0.22	-2.15	2.78	0.37	0.35
	(0.59)			(2.98)		
GDP	-0.51	0.22	-2.24	3.90	0.47	0.18
	(0.46)			(4.54)		
IND	-0.22	0.22	-2.19	2.19	0.33	-0.06
	(0.25)			(2.68)		
LNF1	0.02	0.21	-2.40	2.64	0.37	0.11
	(0.02)			(3.17)		
CAY	-0.83	0.23	-2.06	-0.30	0.23	0.01
	(1.00)			(0.34)		

appear to evolve mostly at the business cycle frequency.

B.3 Foreign Exchange

Table 19: Foreign exchange rates summary statistics

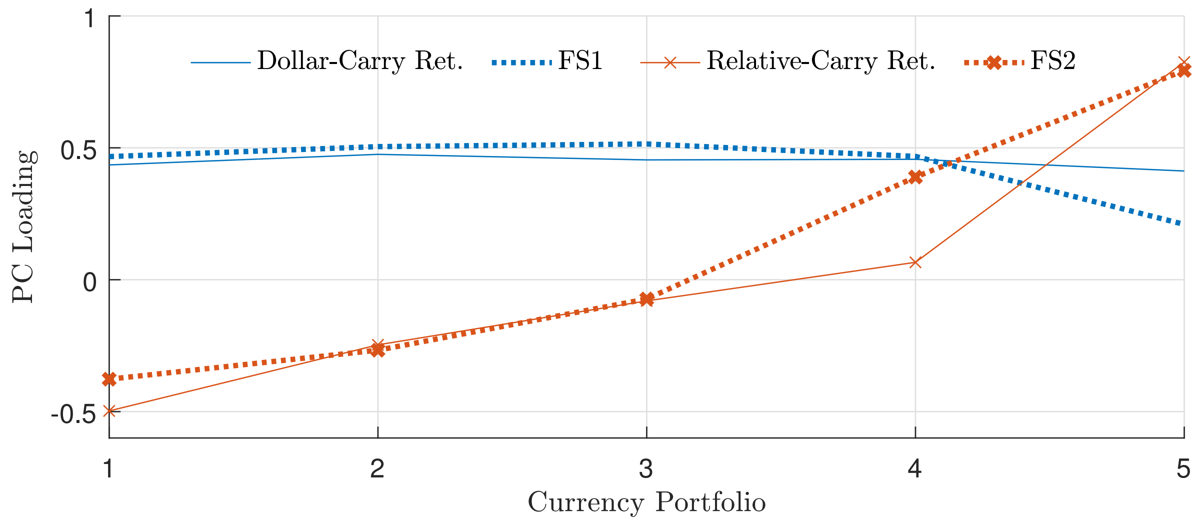
Average forward discounts (row 1), average changes in spot rates (row 2), and excess returns (row 3) for five portfolios (P1–P5) are shown. Hansen-Hodrick standard errors in paratheses. The sample is daily from December 1975 till December 2016.

	P1	P2	P3	P4	P5
Forward spread, $f_t - s_t$	-1.69	-0.12	1.25	2.92	6.42
	(0.35)	(0.32)	(0.30)	(0.31)	(0.53)
Change in spot rates, Δs_{t+1}	-2.10	-1.53	-1.22	-0.23	2.12
	(1.74)	(1.74)	(1.54)	(1.64)	(1.63)
Excess return, $f_t - s_{t+1}$	0.40	1.42	2.47	3.16	4.30
	(1.83)	(1.78)	(1.66)	(1.69)	(1.61)

Table 20: FX PC portfolios summary statistics

The table shows mean, standard deviation, skewness and Sharpe ratio of returns on the dominant components of bonds (Dollar-Carry and Relative-Carry). The first two columns are for static investment strategies and next three are for pure timing strategies.

	Static portfolios		Pure timing portfolios	
	Dollar-Carry	Relative-Carry	Dollar-Carry	Relative-Carry
Mean (%)	1.32	2.21	0.01	0.05
Std. Dev. (%)	5.00	5.00	0.05	0.13
Skewness	-0.05	-0.44	0.37	1.25
Sharpe Ratio	0.26	0.44	0.20	0.38



	Dollar-Carry	Relative-Carry	Curve
% of variance	82.5	10.0	3.4

Figure 10: Factor Structure in Realized Returns and Forward Spreads. The top panel plots the first two eigenvectors of realized currency portfolio returns, termed Dollar-Carry, Relative-Carry, and Curve as well as the first two eigenvectors of forward differentials (FS1 and FS2). The bottom panel shows the percent of total variance contributed by each factor.