A study of data-driven momentum and disposition effects in the Chinese stock market by functional data analysis

Ruanmin Cao, Lajos Horváth, Zhenya Liu, Yuqian Zhao

A Study of Data-driven Momentum and Disposition Effects in the Chinese Stock Market by Functional Data Analysis

Ruanmin Cao* Lajos Horváth† Zhenya Liu‡ Yuqian Zhao§

Abstract

We apply a functional data analysis approach to decompose the cross-sectional Fama-French three-factor model residuals in the Chinese stock market. Our results indicate that other than Fama-French three factors, there are two orthonormal asset pricing factors describing the behavioral biases in their historical performances: between winner and loser stocks, and extreme and mediocre-performing stocks, respectively. We explain these two factors through investors' overreaction, overconfidence and the lead-lag effect. These findings empirically show the existence of momentum and disposition effects in the Chinese stock market. A buy-and-hold mean-variance optimized portfolio incorporating these two market anomalies boosts the Sharpe ratio to 1.27.

Keywords: Momentum effect, Disposition effect, Functional principal component analysis, Portfolio selection, Chinese stock market

JEL Classification: C58, G12, G15, G40

*CITIC Securities Co., China
†Mathematics Department, University of Utah, U.S.A
‡Corresponding author: School of Finance, Renmin University of China, Beijing, 100872, China. Email: zhenya_liu@hotmail.com
§Department of Statistics and Actuarial Science, University of Waterloo, Canada

1. Introduction

Understanding market anomalies in the Chinese stock market gives great value to both investors and market regulators. As the largest emerging market, the Chinese stock market is of enormous interest among international investors, especially after the MSCI added A shares[1] to its benchmark emerging market index in June 2017. Meanwhile, even though considerable progresses in financial liberalization and market regulation have been made since the 1990s, the Chinese stock market still shows differences from other developed markets. One typical feature of the Chinese stock market, similar to many other emerging markets, is its high volatility. Consequently, it is questionable whether the market anomalies discovered in the developed markets are still feasible in the Chinese stock market.

One controversial topic in the literature is the momentum effect in the Chinese stock market (see, Kang et al., 2002; Wong et al., 2006; Wu, 2011; Cheema and Nartea, 2014 and Cakici et al., 2015). Results suggest that the momentum anomaly in China gives a relatively reduced predictive power. However, these results contradict the knowledge that the momentum factor is the most pervasive and well-documented market anomaly (Fama and French, 2012; Asness et al., 2013). The standard approach to find market anomalies is to sort cross-sectional returns according to firm-specific historical characteristics, such as size, value, momentum, and dividends (see, Fama and French, 1992, 2012, 2015; Carhart, 1997; Daniel and Titman, 1997; amongst others), and one expects to obtain abnormal returns if anomalies exist. As a result, we can use multi-factor models, such as the Fama-French (F-F) three-, five-factor and the Carhart four-factor models (Fama and French, 1992, 1993, 2015; Carhart, 1997) to explain cross-sectional returns. Thus, due to these empirical contradictions, it is natural to ask if these traditional methodologies are still appropriate in the highly volatile Chinese stock market? Our impetus is to use updated statistical techniques to investigate market anomalies in Chinese A-shares.

[1] In this paper, we term the China A-shares market as the Chinese stock market, which is dominated in Chinese Yuan/CNY, and is mainly accessible for local investors. Compared with the developed markets, individual investors in this market outweigh institutions regarding the stock capitalization held.

Briefly, we apply a recently developed functional data analysis technique to decompose the total variation of the cross-sectional residuals. The intuition is that in a highly volatile market, some behavioral biases are hard to explain with the well-known asset pricing factors, and the residuals might play a more critical role. Blitz et al. (2011) constructed a residual-weighted portfolio and argued that the volatilities of total returns are too high to propose a well-performing trading strategy. Thus, we explore underlying anomalies based on the F-F three-factor residuals instead of the total stock returns. In Gandhi and Lustig (2015), the authors focused on a relatively chaotic market of U.S. bank shares and applied principal component analysis (PCA) to decompose a size-sorted cross-section of U.S. bank residuals, and they found that the second principal component suggests a size factor in highly leveraged bank shares.

Enlightened by their methodology, we propose a two-step functional data analysis approach to decompose the residual returns. In the first step, we smooth the cross-sectional residuals into functional curves in the Hilbert $L^2[0,1]$ space by using B-spline orthonormal basis functions. Secondly, we use the functional principal component analysis (FPCA) to obtain data-driven eigenfunctions. These eigenfunctions exploit the underlying risk patterns and suggest asset pricing factors. For a related work to study the cross-sectional returns under the framework of FDA we refer to Kokoszka et al. (2018).

To understand the differences between our method and the method used by Gandhi and Lustig (2015), we can compare FPCA with PCA. In the context of FDA, since discrete high dimensional data are treated as continuous functional curves, FPCA extracts eigenfunctions rather than eigenvectors, so in some sense, FPCA is a continuous version of the PCA. In equity markets, although the cross-sectional returns are discretely observed, the risk patterns on the cross-sections are naturally continuous, which makes FPCA more appropriate in this situation. An additional benefit from using the smoothing technique is that the FPCA controls the level of roughness by filtering noise while preserving main features. For details, we refer to the monographs of Ramsay and Silverman (2006) and Horváth and Kokoszka (2012).

In the empirical study, we apply the two-step method to explore the F-F three-factor residuals. Because

our interest lies in the momentum effect, we first sort the cross-sectional returns based on 12-month firm-specific accumulated past returns. The cross-sections are ranked based on their historical performances, and we form ten equal-weighted portfolios from a loser group to a winner group. Next, we decompose the variability of the residual returns to estimate eigenfunctions. In our case, we estimated 13 eigenfunctions of which the first four eigenfunctions explain 87.77% of the total variation of the F-F three-factor residuals, with the first one already explaining 64.62%, and the second one 14.73%. We then construct two asset pricing factors by using the standardized weights suggested by the first two eigenfunctions, leaving the remaining eigenfunctions as noise.

These two data-driven risk factors help us to understand the Chinese stock market better. The first factor is a data-driven momentum winner-minus-loser (wml) factor. Note that the behavioral bias toward losers and winners can either be explained by the winner-minus-loser effect (Jegadeesh and Titman, 1993) or the contrarian loser-minus-winner effect (Bondt and Thaler, 1985). Both scenarios reveal the momentum effect. We construct relative strength portfolios (Jegadeesh and Titman, 1993) and find that the data-driven *wml* factor is belonging to the latter case. From the perspective of behavioral finance, the momentum effect may originate from three sources: (1) investor sentiments of underreacting short-run information and overreacting information in the long-run (Chan et al., 1995; Barberis et al., 1998; Jegadeesh and Titman, 2001); (2) investor overconfidence (Bondt and Thaler, 1985; Kang et al., 2002); (3) the lead-lag effect (Lo and Mackinlay, 1990). Based on an analysis of market data, our results show that investors in the Chinese stock market are overreact to short-run information. The lead-lag effect also exists, which is consistent with the conclusion made by Kang et al. (2002).

The second factor, indicating the distinction between the extreme-performing and mediocre-performing groups, can be explained by the asymmetric disposition effect. As an extension of the prospect theory, the disposition effect (Shefrin and Statman, 1985) describes the investors' risk attitude toward stock performances, which indicates that investors sell winning stock too early and ride losers too long (also see Fazzini, 2006; Barberis and Xiong, 2009). Some developments in the disposition effect were discussed by Ben-David

and Hirshleifer (2012; also see An, 2015), who argued that the selling/buying function of the disposition effect for past returns is asymmetric U-shaped and explained this with investor overconfidence. Our findings are also in agreement with Chen et al. (2007), who documented the disposition effect by using brokerage account data in the Chinese stock market. Overall, these two new asset pricing factors increase the adjusted R^2 in the F-F three-factor model from 0.79 to 0.89 and bring less significant adjusted returns (alpha). The evidence of the robustness of these two factors is shown by using the Fama-Macbeth regression.

Our paper contributes to the literature in several ways. We suggest a functional data analysis approach to decompose cross-sectional returns and calibrate market anomalies from samples. Our findings show that the momentum effect in the Chinese stock market is not as weak as mentioned in the literature. We also complement the literature by empirically showing the disposition effect in the Chinese stock market. Instead of using individual investor-level data, we show that the asymmetric U-shape disposition effect can be uncovered from market data. The well-known finding that disposition effect is a source of momentum effect in Grinblatt and Han (2005) has been challenged recently. Birru (2015) did not find disposition effect after a stock split and he argued that the disposition effect is not able to explain the momentum effect. We complement this vein of research by showing that momentum and disposition effects give two orthonormal risk premiums in the Chinese stock market. Additionally, we use behavioral finance theories to explain our results, this providing a better understanding of the impact of investors' behavior and regulation rules on the pricing equilibrium in the Chinese stock market. This provides a deeper market insight for international investment practitioners. Our results can be easily applied to portfolio selection in practice. We incorporate the exploited risk factors into a buy-and-hold portfolio, and the Sharpe ratio of the mean-variance optimization increased to 1.27.

The outline of the paper is as follows. In Section 2, we discuss the dataset and apply conventional multi-factor models to study the market anomalies in the Chinese stock market. In Section 3 we introduce a two-step functional data analysis approach to decompose cross-sectional residuals. Section 4 exploits the residuals in the Fama-French three-factor model and introduces two data-driven common risk factors, which

are further elaborated in Section 5. Section 6 incorporates them into eigenfunction portfolios, and concluding remarks are made in Section 7.

2. Data and benchmark models

In this section we introduce the dataset and revisit the results of classic multi-factor models in Chinese A-shares. We regress past return-sorted cross-sectional returns onto the classical asset pricing factors (Fama and French, 1992; Carhart, 1997). The F-F three-factor residuals are kept for the use of later sections.

2.1. Dataset

We collect monthly adjusted closing prices [2] for all Chinese A-shares to avoid erratic behavior in daily and weekly frequency. To maintain the consistency with relevant studies, we exclude the financial sector due to its high leverages, as well as listings without 12-month past returns information. The dataset ranges between January 2007 and June 2017 in order to eliminate the influence of the structural reform in the Chinese stock market during the year of 2006[3]. Note that the sample period over the last decade contains two business cycles. There are 1470 shares listed in 2007, and this number increases to 3226 in June 2017. We use the log return transformations of the raw price data.

We use the Shanghai Stock Exchange Composite index and three-month Treasury bill rate as the market index r_t^m and the risk-free rate r_t^f, respectively. For constructing the size and value factors, we collect firm-specific market capitalization and book-to-market ratios. The short-term and long-term firm-specific past returns are computed by the equations

$$PR_{i,t} = \sum_{j=t-K}^{t-2} r_{i,j}, \qquad LPR_{i,t} = \sum_{j=t-60}^{t-48} r_{i,j}, \tag{1}$$

[2]The closing price is adjusted by its stock splits and dividend payments.

[3]There are non-tradable shares in the Chinese stock market before 2005 because of the liquidity shortage. In order to reconstruct the ownership structure of listed companies, the authority launched a structural reform to eliminate the non-tradable shares during 2005-2006.

where the subscript i denotes stock i. Unless otherwise stated, $K = 12$. The long-term $LPR_{i,t}$ is only used to assess the performances of relative strength portfolios in section 5.1.

The market risk premium is computed as $rmrf_t = r_t^m - r_t^f$. We first sort cross-sectional returns into deciles according to firm-specific size, value and past returns in ascending order, and then construct equal-weighted small-minus-big (smb), high-minus-low (hml) and wml common risk factors. Additionally, we collect the monthly firm-level turnovers to analyze their trading activities in section 5.2. All data was obtained from the Wind database.

2.2. Multi-factor models

We start to build portfolios according to the standard portfolio strategy of Fama and French (1993). Cross-sectional returns are ranked by their 12-month past returns. We group stocks into ten portfolios ranging from "loser" to "winner" according to their firm-specific $PR_{i,t}$. The portfolio returns are computed with equal weights. For completeness, we also consider a "winner-minus-loser" portfolio. We first assess the explanatory power of the F-F three-factor model,

$$r_{i,t} - r_t^f = \alpha_i + \beta_{1,i} rmrf_t + \beta_{2,i} smb_t + \beta_{3,i} hml_t + \epsilon_{i,t}, \tag{2}$$

where $r_{i,t}$ denotes the ith decile portfolio return for $i \in [1, 10]$. The F-F three-factor model studies the cross-sectional market anomalies through firm-level fundamental information (also see the five-factor model, Fama and French, 2015). Meanwhile, in order to verify existing results about the momentum effect, we examine the Carhart (1997) four-factor model,

$$r_{i,t} - r_t^f = \alpha_i + \beta_{1,i} rmrf_t + \beta_{2,i} smb_t + \beta_{3,i} hml_t + \beta_{4,i} wml_t + \epsilon_{i,t}. \tag{3}$$

Once all multi-factor models are estimated, the next issue is to assess their performance. Since a better factor model always delivers less significant risk-adjusted returns (intercept α_i), the GRS test (Gibbons et al., 1989) is used here to test the joint significance of intercept coefficients. We aim to test the null hypothesis $H_0 : \alpha_i = 0$, for all $i = 1, 2, \ldots, 10$. When comparing two multi-factor models $\mathcal{M}_1$ and $\mathcal{M}_2$, if $\mathcal{M}_1$ rejects H_0 at a lower significance level, we then conclude that $\mathcal{M}_1$ better predicts cross-sectional returns than $\mathcal{M}_2$.

[Insert Table 1 Here]

Table 1 displays the estimators of the regression models specified in Equations (2) and (3). Panel A shows the estimation results in the F-F three-factor model, and we observe three remarkable facts. First, the GRS test does not find an overall statistically significant risk-adjusted return, and the estimated intercepts in the majority of deciles are not statistically significant. The only exceptions are the two winner (9th and 10th) groups and the winner-minus-loser portfolio, where these portfolios show significant risk-adjusted returns with negative signs. Second, the model exhibits better explanatory power to 4th-6th groups with adjusted $R^2 = 0.86$, while it becomes less efficient for the loser and winner groups, e.g., adjusted $R^2 = 0.79$ in the winner group. Lastly, we notice that the coefficients of F-F three factors are statistically significant. The market factor and size factor positively explain the excess returns, while the value factor shows negative coefficients to all of the deciles under consideration. Similar results have been discovered in the literature (see Kang et al., 2002; Wong et al., 2006; Cakici, 2015). The size effect is significant because initial public offerings (IPO) in China are regulated by the China Securities Regulatory Commission (CRSC) rather than based on a system of registration like in the developed markets. This barrier makes newly listed small shares more likely to be over-valued in the stock market. The negative signs of the value factor are possible because retail investors pay less attention to the value factor; they mainly make their investment decisions based on market rumors.

Panel B of Table 1 shows that, not surprisingly, the Carhart four-factor model outperforms the F-F three-factor model with a higher adjusted R^2, especially in the loser and winner groups. The coefficients

of the *wml* factor are negative and positive at the losers and winners groups, respectively. The result of the GRS test implies that the Carhart four-factor model generates less significant risk-adjusted returns than the F-F three-factor model. Nevertheless, the variation in the cross-sectional returns is still not to be fully explained, given the fact that risk-adjusted returns are jointly significant at the 1% significance level in both of the models. Therefore, more risk patterns are hidden in the residuals and worth decomposing.

3. Functional data decomposition approach

In this section, we introduce a functional data analysis approach to decompose cross-sections of portfolio residuals. Converting the panel data $\boldsymbol{y}_{i,t} = [y_{1,t}, y_{2,t}, \ldots, y_{N,t}]$, $1 \leq t \leq T$, $1 \leq i \leq N$ to functional observations, we smooth the cross-sectional entities into the $[0, 1]$ interval without loss of generality and obtain $y_t(u)$ for $1 \leq t \leq T$ and $0 \leq u \leq 1$. We assume that the random curve $y_t(u)$ is square integrable and that it is in the Hilbert space $L^2[0, 1]$ for all $1 \leq t \leq T$. The space $L^2[0, 1]$ is equipped with inner product $\langle x(u), y(u) \rangle = \int_0^1 x(u)y(u)du$, and the corresponding norm is $\|x(u)\| = (\int_0^1 x^2(u)du)^{1/2}$. The functional data $y_t(u)$ can be expanded by the Karhunen-Loéve Theorem (Horváth and Kokoszka, 2012),

$$y_t(u) = \sum_{m=1}^{\infty} \eta_{t,m}\phi_m(u) \approx \sum_{m=1}^{M} \eta_{t,m}\phi_m(u), \tag{4}$$

where $\{\phi_m(u), 0 \leq u \leq 1\}$ are orthonormal functions on $[0, 1]$, i.e., $\langle \phi_i, \phi_j \rangle = 0$ if $i \neq j$ and $\langle \phi_i, \phi_i \rangle = ||\phi_i||^2 = 1$. The most commonly used basis functions $\phi_m(u)$ are Fourier basis and B-splines basis. The former is mainly used for cases in which data contains periodical or nearly-periodical patterns, and the latter usually adapts to non-periodic data. The number of bases M can be crucial, and it is determined by optimizing the trade-off between fitting features and noises in the smoothing process. The coefficients $\eta_{t,m}$ $1 \leq m \leq M$, $1 \leq t \leq T$ are given by the inner product of $\phi_m(u)$ and $y_t(u)$,

$$\eta_{t,m} = \langle \phi_m(u), y_t(u) \rangle = \int_0^1 \phi_m(u)y_t(u)du. \tag{5}$$

We assume that

Assumption 3.1 *$\{y_t(u), 0 \leq u \leq 1\}$ is a stationary sequence with sample paths in $L^2[0,1]$, and also the moment condition $E||y_t(u)||^4 < \infty$ is satisfied.*

Assumption 3.1 is very mild as it is fairly acceptable that financial data follows autoregressive and GARCH-type processes in $L^2[0,1]$ (Hörmann and Kokoszka, 2010). Due to stationarity, the functional mean $Ey_t(u) = \mu(u)$ does not depend on time. The covariance function of $\{y_t(u), 0 \leq u \leq 1\}$ is defined by

$$c(u,v) = E[(y_t(u) - \mu(u))(y_t(v) - \mu(v))],$$

where $c(u,v)$ is a symmetric, non-negative definite function. We aim to decompose $c(u,v)$. There are eigenvalues and corresponding eigenfunctions as $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ and $\psi_1(u)$, $\psi_2(u)$, ... satisfying

$$\lambda_i \psi_i(u) = \int_0^1 c(u,v)\psi_i(v)dv, \quad i = 1, 2, \ldots. \tag{6}$$

In principal component analysis as well as in functional data analysis, $\phi_m = \psi_m$ is a suitable choice in (4) because $\sum_{m=1}^{M} \langle y_t(u), \psi_m \rangle \psi_m(u)$ is the best approximation for $y_t(u)$ in $L^2[0,1]$, i.e., when $y_t(u)$ is approximated with $1, 2, \ldots, M$ orthogonal functions in $\mathcal{H}$, we get the smallest $E \left\| y_t(u) - \sum_{m=1}^{M} \langle y_t(u), , \psi_m \rangle \psi_m \right\|^2$. However, $c(u,v)$ is unknown and it should be estimated with

$$\hat{c}_T(u,v) = \frac{1}{T}\sum_{t=1}^{T}(y_t(u) - \bar{y}_T(u))(y_t(v) - \bar{y}_T(v)), \text{ where } \bar{y}_T(u) = \frac{1}{T}\sum_{t=1}^{T} y_t(u).$$

Now the eigenvalues and eigenfunctions of the theoretical $c(u,v)$ are estimated by $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq 0$ and $\hat{\psi}_1(u)$, $\hat{\psi}_2(u)$, ... satisfying

$$\hat{\lambda}_i \hat{\psi}_i(u) = \int_0^1 \hat{c}_T(u,v)\hat{\psi}_i(v)dv, \quad i = 1, 2, \ldots, \tag{7}$$

so due to the optimality properties of the eigenfunctions of $\hat{c}_T(u,v)$, a popular way to decompose functional

observations $y_t(u)$ is

$$y_t(u) \approx \hat{y}_t(u) = \sum_{m=1}^{M} \hat{\eta}_{t,m}\hat{\psi}_m(u), \tag{8}$$

with $\hat{\eta}_{t,m} = \left\langle \hat{\psi}_m(u), y_t(u) \right\rangle = \int_0^1 \hat{\psi}_m(u) y_t(u) du$.

This means that $y_t(u)$ is projected into the finite dimensional space spanned by $\{\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_M\}$. This projection works well if $y_t(u)$ is a weakly dependent sequence. Then,

Theorem 3.1 *If Assumptions 3.1 holds, then we have that for each* $1 \leq i \leq M$

$$|\hat{\lambda}_i - \lambda_i| \xrightarrow{p} 0,$$

$$||\hat{\mathfrak{c}}_i\hat{\psi}_i - \psi_i|| \xrightarrow{p} 0,$$

where $\xrightarrow{p}$ *denotes convergence in probability and* $\hat{\mathfrak{c}}_1$, $\hat{\mathfrak{c}}_2$, ... *are random signs defined by* $\hat{\mathfrak{c}}_i = sign(\left\langle \hat{\psi}_i, \psi_i \right\rangle)$.

Theorem 3.1 is an immediate consequence of the ergodic theorem in Hilbert spaces, i.e., $||\hat{c}_T - c|| \xrightarrow{p} 0$ (see Hörmann and Kokoszka, 2010 for the proof). Another consequence of the theorem is that $\hat{y}_t(u)$ in (8) is close to $y_t(u)$ if M and T are both large. Therefore, the cross-sectional patterns in the variation of $y_t(u)$ are decomposed and expressed by a finite number of estimated eigenfunctions $\hat{\psi}_i$, and their variance contributions are captured by their corresponding $\hat{\lambda}_i$.

From the computational point of view, the integral is approximated by Riemann sums; for instance, $\hat{\eta}_{t,m}$ in (8) is approximated by

$$\hat{\eta}_{t,m} = \frac{1}{N}\sum_{i=1}^{N} y_t(u_i)\hat{\psi}(u_i),$$

where $y_t(u_i)$ is $y_t(u)$ on the regular grid u_i, $1 \leq i \leq N$.

4. Empirical results: decomposition of Chinese A-share cross-sectional returns

This section uses the method proposed above to decompose the F-F three-factor residuals, i.e., $\hat{\epsilon}_{i,t}$ in Equation (2). In order to accommodate the notation in the last section, we rewrite $\hat{\epsilon}_{i,t}$ into $\hat{\epsilon}_t(u_i)$. The reason for omitting the Carhart four-factor residuals is because our dependent variables are past return-sorted cross-sectional returns, the *wml* factor is suspected to contain overlaps and should not be controlled here.

4.1. Decomposing variations in residuals

We explore the estimated residuals $\hat{\epsilon}_t(u_i)$, $i \in [1, N]$, where N is the total number of sorted groups. As a pre-analysis, we again sort cross-sectional returns into 100 groups ($N = 100$) using the algorithm mentioned in section 2.2, and this is because the previous ten groups are not enough for implementing the functional smoothing technique. In each of the groups, we regress excess portfolio returns on the F-F three factors and save the residuals $\hat{\epsilon}_t(u_i)$, $1 \leq i \leq 100$. Before adopting the two-step decomposition method, the estimated residual vectors are demeaned and standardized for emphasizing variations and reducing noise as suggested by Blitz et al. (2011). In the first step, we smooth these residuals into functional curves. At each time t, we set 11 knots on $\hat{\boldsymbol{\epsilon}}_t(u_i)$. Considering that there is hardly any periodical pattern in the cross-sectional returns, we choose 13 B-spline functions as the smoothing bases as shown in Figure 1. The smoothing technique leads to N=126 functional curves $\hat{\epsilon}_t(u)$ spanning over 2007-2017, appearing in the upper panel of Figure 2. By smoothing residual observations to curves, Assumption 3.1 is satisfied.

[Insert Figure 1, Figure 2 and Figure 3 Here]

Next, we apply the FPCA method to decompose the functional residual curves $\hat{\epsilon}_t(u)$ and find 13 empirical eigenfunctions (ef), $\psi(u) = \{\hat{\psi}_j(u) | \hat{\lambda}_j > 0, 1 \leq j \leq 13\}$, where $\hat{\psi}_j(u)$ is the jth eigenfunction and $\hat{\lambda}_j$ is the corresponding eigenvalue. The lower panel of Figure 2 illustrates that the first four eigenfunctions explain 87.77% of the total variation. The first eigenfunction $\hat{\psi}_1(u)$ explains 64.62% and displays an upward trend.

This pattern implies a risk factor that loads negatively on the loser groups and positively on the winner groups. It is similar but not identical to the momentum *wml* factor. The second eigenfunction $\hat{\psi}_2(u)$, plotted as an asymmetric U/V shaped curve, accounts for 14.73% of the total variation. This risk pattern suggests that investors behave differently toward extreme and mediocre stocks, thereby suggesting a risk factor that loads positively on the extreme-performing groups and negatively on the mediocre-performing groups. We compute the 95% confidence envelopes of the first four eigenfunctions via a bootstrapping approach (5,000 iterations) (Hall and Hosseini-Nasab, 2006), as shown in Figure 3. Table 2 displays the percentages of variations explained by the first four eigenfunctions, where the results exploited from the real data are consistent with the bootstrapped ones. The eigenfunctions from the third and below are hard to align with cross-sectional portfolio returns. Also because they account for minor variations, we treat them as noise.

[Insert Table 2 Here]

Another method to select useful eigenfunctions is by using of the information coefficient (IC). Figure 4 plots ICs between the past returns-sorted cross-sectional returns and eigenfunctions. Over the entire period, $\hat{\psi}_1(u)$ shows the highest average IC, followed by $\hat{\psi}_2(u)$, while the average IC of $\hat{\psi}_3(u)$ and $\hat{\psi}_4(u)$ are flatter and close to zero. Therefore, we use $\hat{\psi}_1(u)$ and $\hat{\psi}_2(u)$ to construct asset pricing factors.

[Insert Figure 4 Here]

4.2. Data-driven asset pricing factors: $FPC1_t$ and $FPC2_t$

In order to check whether the uncovered asset pricing factors actually explain the cross-sectional returns, we build two data-driven variables: $FPC1_t$ ($\mathbf{r}_t(u_i) \times \hat{\psi}_1(u_i)'$), and $FPC2_t$ ($\mathbf{r}_t(u_i) \times \hat{\psi}_2(u_i)'$), where $\mathbf{r}_t(u_i) = [r_t(u_1), r_t(u_2), \ldots, r_t(u_N)]$ is a 126×100 matrix, and $\hat{\psi}_j(u_i)'$ is a 100×1 weights vector with the grid values of $\hat{\psi}_j(u)$. The first variable is a portfolio that shorts loser shares and longs winner shares, while the second variable is a portfolio that longs extreme-performing shares and shorts mediocre-performing shares. Note

that the weights are standardized and their sum is equal to one.

We apply the Fama-MacBeth regression (Fama and MacBeth, 1973) on $FPC1_t$ and $FPC2_t$, which is the standard method to estimate the coefficients of asset pricing factors. Table 3 shows the results: $FPC1_t$ has significant risk premium on past returns-sorted portfolio returns with a strong January effect; $FPC2_t$ is statistically significant at the 10% level, and surprisingly it is less significant over the course of January. The final column shows that, due to orthogonality, the coefficients of risk premiums for $FPC1_t$ and $FPC2_t$ remain unchanged and the adjusted R^2 equals the sum of single-factor regressions. This finding reveals that both $FPC1_t$ and $FPC2_t$ are strong asset pricing factors in the Chinese stock market.

[Insert Table 3 Here]

In order to have a contest with the benchmark factor models, we again compute the multi-factor regression models while adding the $FPC1_t$ and $FPC2_t$ factors. Table 4 reports the results of the regression models specified in Equations (9) and (10). Compared with Table 1, adding $FPC1_t$ and $FPC2_t$ improves the explanatory power of the models, which we expected given the higher adjusted R^2s and the statistically less significant risk-adjusted returns are reported.

$$r_{i,t} - r_t^f = \alpha_i + \beta_{1,i} rmrf_t + \beta_{2,i} smb_t + \beta_{3,i} hml_t + \beta_{4,i} FPCX_t + \epsilon_{i,t}, \quad X = 1, 2., \tag{9}$$

$$r_{i,t} - r_t^f = \alpha_i + \beta_{1,i} rmrf_t + \beta_{2,i} smb_t + \beta_{3,i} hml_t + \beta_{4,i} FPC1_t + \beta_{5,i} FPC2_t + \epsilon_{i,t}. \tag{10}$$

Remarkably, panel A of Table 4 shows that the estimated coefficients of the F-F three factors remain almost the same as in Table 1. The estimated coefficients of $FPC1_t$ are similar to the *wml* factor in panel B of Table 1. Also, we note that the sign of the coefficients of $FPC1_t$ are also identical to the sign of the *wml* factor, and the adjusted R^2 is equal to 0.95 when we regress $FPC1_t$ on *wml*. Hence, it is not hard to see that $FPC1_t$ is a data-driven form of the *wml* factor.

Panel B of Table 4 shows that $FPC2_t$ significantly explains excess returns in the mediocre and loser groups. These findings are reasonable because $FPC2_t$ affirms the behavioral biases toward mediocre and extreme-performing shares. The coefficients of $FPC2_t$ in the mediocre groups are around -0.14, and this coefficient changes to 0.08 in the loser group. Another remarkable fact is that $FPC2_t$ cannot explain the *wml* portfolio return (adjusted $R^2 = 0.15$), which suggests that $FPC2_t$ explains phenomena which can not be explained by the *wml* or $FPC1_t$ factor.

Due to orthogonality, panel C of Table 4 shows the enhanced explanatory power of the multi-factor models when adding $FPC1_t$ and $FPC2_t$, with adjusted R^2 over 0.87 for all groups. This model takes advantage of the observation that the risk-adjusted returns in all of the groups under study become insignificant. The relatively lower P-values of the GRS tests indicate that model (10) outperforms the other factor models discussed above. Finally, in order to assess the influence of market state on our asset pricing factors, we follow Cooper et al. (2004) to inspect the role of market states on the FPC factors. By defining an upward market (downward market) that accumulated historical 6-month or 12-month market returns as positive (negative), our results rarely show a notable impact on $FPC1_t$ and a slightly significant impact on $FPC2_t$. We omit these results here.

[Insert Table 4 Here]

5. Explanations of data-driven asset pricing factors

We now explain the new asset pricing factors from the perspective of behavioral finance. Since the FPC factors are constructed based on exploited cross-sectional risk patterns, together with the fact that cross-sectional returns are sorted by past return, these factors give insight to investors' behavioral biases regarding persistence of equities' performances.

5.1. How to explain $FPC1_t$?

The previous section showed that $FPC1_t$ is a data-driven *wml* factor describing the investors' behavioral bias toward the winner and loser shares. Here we investigate the underlying psychological structures of the risk factor $FPC1_t$. We first follow Jegadeesh and Titman (1993) in constructing relative strength portfolios. We sort cross-sectional returns in ascending order by firm-specific past returns with different look-back periods K, where K varies from short-term (3, 6, 9 and 12 months) to long-term (60 months, cf. Equation (1)). The holding periods are 3, 6, 9, 12 and 24 months. In total, there are 50 portfolios (25 momentum winner-minus-loser and 25 contrarian loser-minus-winner). The average monthly returns are reported in Table 5. The momentum strategy generates no profit while the portfolio with the contrarian strategy earns positive profits under any formation and holding periods. Our results are consistent with the findings of Kang et al. (2002), Griffin et al. (2003) and Wu (2011).

[Insert Table 5 Here]

One reason for observing profitable contrarian portfolios, as suggested by Bondt and Thaler (1985), is investors' overreaction, as evidenced by calculating autocorrelations of portfolio returns with negative signs. Lo and Mackinlay (1990) observed that the cross lead-lag effect is another reason for a profitable contrarian strategy. The lead-lag effect is observed when a higher return on stock i at period $t-1$ leads to a higher return on stock j at period t, i.e., calculating positive cross-serial correlations in portfolio returns. We thus expect to get negative autocorrelations and positive cross-serial correlations when overreaction and lead-lag effects exist.

To study the dependence structure of past returns-sorted cross-sectional portfolio returns, we calculate serial correlations among these portfolios using lags ranging from 1 to 4. Panel A of Table 6 documents serial correlation coefficients, which shows that 33 out of 40 coefficients are negative. This result indicates a short-run overreaction effect in the Chinese stock market. Similar results were found by Tan et al. (2008)

and Ni et al. (2015). Panel B of Table 6 gives cross-autocorrelations with lags 1. Even though the cross-serial correlations vary between positive numbers and negative numbers, note that the historical returns of the loser group is positively correlated with the contemporary returns of the winner group (with correlation value 0.06). A similar situation occurs between the historical returns of the winner group and the contemporary returns of the loser group (with correlation value 0.02). This result implies that there is a lead-lag effect between winner and loser portfolios. We thus confirm that investors' overreaction and the lead-lag structure are the driving forces of the data-driven momentum factor.

[Insert Table 6 Here]

5.2. How to explain $FPC2_t$?

$FPC2_t$ captures investors' asymmetric investment behavior regarding extreme and mediocre-performing shares. This pattern is in accordance with the asymmetric U-shaped disposition effect observed by Ben-David and Hirshleifer (2012), see Panel B of Figure 2 in their paper. The authors claimed this effect reflects investors' overconfidence in the U.S. market, as investors always over-value their knowledge and are inclined to trade stocks with big news; these stocks are more likely to be extreme-performing stocks.

Moreover, another potential source of $FPC2_t$ is the selling pressure driven by the disposition effect on both outperforming and underperforming stocks. Deviations from the current price level increases liquidations in extreme-performing stocks. Hence, investors have to pay the risk premium to 'mediocre-performing stocks' for their safety. This argument parallels with the low-volatility anomaly. We note that our work, from this perspective, demonstrates another measure of "safety", in addition to idiosyncratic risk (Blitz and Vliet, 2007) and beta (Frazzini and Pederson, 2014). Regulation rules in the Chinese stock market restrict the daily price ceiling and floor, i.e., maximum 10% price changes, which have been widely adopted in Asian exchanges (Chan et al., 2005). A significant price change in extreme-performing shares thus leads to drainage of liquidity and severely exacerbates order imbalance, resulting in a higher return compensating

for the illiquidity.

Another notable pattern is that the $FPC2_t$ factor shows a right skewness toward losers. This asymmetric pattern can be interpreted by prospect theory. Prospect theory states that investors are more inclined to trade winning stocks and to hold losers. In another word, investors are more likely to re-examine positions or to update their beliefs about profitable stocks but are relatively inattentive to losers. Also see An (2015).

[Insert Figure 5 Here]

In order to empirically assess the disposition effect, we use firm-specific turnover to measure their trading activities. Table 7 shows that the turnover ratio between the extreme and mediocre groups is as high as 1.4124. Table 8 further reports turnover ratios between each of individual groups. The ratios between the loser groups and mediocre groups are more than 1.0 but less than 1.1, while these ratios rise to 1.8245 in cases of the winner and group 5. Figure 5 plots the turnover ratio for each group during the sample period. Although trading turnover in the loser group is not always higher than in the mediocre groups, the asymmetric U-shaped pattern is observable, particularly during the periods of 2009-2010, 2011-2014 and 2017.

[Insert Table 7 and Table 8 Here]

Meanwhile, since the eigenfunctions $\hat{\psi}_1(u)$ and $\hat{\psi}_2(u)$ are orthonormal, the behavioral biases explained by $FPC1_t$ and $FPC2_t$ cannot account for each other. In other words, this result indicates that the disposition effect can be separated from the momentum effect so that it deliveries a different risk premium[4].

6. The eigenfunction portfolio

In this section, we apply the $FPC1$ and $FPC2$ factors to develop a portfolio construction strategy. The relative strength portfolios have shown that the contrarian strategy earns positive returns while the momen-

[4]We also applied the same approach to the U.S. market including more than 7,000 shares in the NYSE, AMEX and NASDAQ from 2,000 to 2017. The result showed that the $FPC2_t$ factor still exists but is not as strong as in the Chinese stock market.

tum strategy does not. Thus, together with the market index, we treat the contrarian loser-minus-winner portfolio as the benchmark trading strategy. Considering a buy-and-hold strategy without any transaction cost, we conducted an in-sample simulation. Recall that cross-sectional returns were sorted into 100 groups from the loser to the winner. We constructed extreme spread portfolios, which portfolio weights are assigned by using the standardized values of $\hat{\psi}_1(u)$ and $\hat{\psi}_2(u)$. We name these portfolios as eigenfunction portfolios and denote them with $EP1$ and $EP2$, respectively. We also apply a mean-variance optimization on the $EP1$ and $EP2$, where the optimal weights are calculated by using historical three-month rolling windows, in order to avoid any future information.

Table 9 reports some popular statistics. The first eigenfunction strategy performs similarly to the contrarian-lmw strategy with a Sharpe ratio of around 0.48. The second eigenfunction portfolio is also profitable with Sharpe ratio of 0.33, which is lower than EP1 as fewer variations are explained. Lastly, benefiting from the orthogonality, the mean-variance optimized portfolio records a boosting of the Sharpe ratio to 1.27 and a much lower maximum draw-down of -0.02.

Figure 6 plots the cumulative returns for contrarian, $EP1$, $EP2$, and mean-variance optimized portfolios. The monthly performance shows that the mean-variance optimization obtains positive returns in all of the months under study. Therefore, the use of common risk patterns exploited by $FPC1_t$ and $FPC2_t$ leads to better portfolio selection. Note that as long as $EP1$ and $EP2$ persist out-of-sample, the momentum and disposition payoff structures should always be profitable in the active portfolio selection. Next, we discuss the profitability of eigenfunction portfolios under the framework of the FDA.

[Insert Table 9 and Figure 6 Here]

In order to discuss the profitability of eigenfunctions in the $L^2[0,1]$ space, we consider the recalling functional objective $r_t(u)$. Recalling Equation (7), where the empirical eigenfunctions $\hat{\psi}_j(u)$ and eigenvalues $\hat{\lambda}_j$ are estimated, we incorporate these data-driven cross-sectional patterns into curves $r_t(u)$. The curves

$r_t(u)$ thus can be approximated well with finite $\mathcal{K}$ eigenfunctions,

$$r_t(u) \approx \mu(u) + \sum_{j=1}^{\mathcal{K}} \left\langle r_t(u), \hat{\psi}_j(u) \right\rangle \hat{\psi}_j(u) + \varepsilon_t(u), \tag{11}$$

where $\mu(u)$ is the functional mean, and the projection $\left\langle r_t(u), \hat{\psi}_j(u) \right\rangle$ is the time-varying functional loading for $\hat{\psi}_j(u)$.

We take the jth eigenfunction portfolio return $r_{p_j,t}$ as an example. It is not hard to see that $r_{p_j,t}$ is the inner production between the jth eigenfunction $\hat{\psi}_j(u)$ and the return $r_t(u)$,

$$r_{p_j,t} = \langle r_t(u), \hat{\psi}_j(u) \rangle. \tag{12}$$

Substituting Equation (11) into above equation, we then get

$$\begin{aligned} r_{p_j,t} &= \langle (\mu(u) + \sum_{j=1}^{\mathcal{K}} \left\langle r_t(u), \hat{\psi}_j(u) \right\rangle), \hat{\psi}_j(u) \rangle + \left\langle \varepsilon_t(u), \hat{\psi}_j(u) \right\rangle \\ &= \left\langle \mu(u), \hat{\psi}_j(u) \right\rangle + \langle \left\langle r_t(u), \hat{\psi}_j(u) \right\rangle, \hat{\psi}_j(u) \rangle + \left\langle \varepsilon_t(u), \hat{\psi}_j(u) \right\rangle. \end{aligned} \tag{13}$$

According to Theorem 3.1, the third term $\left\langle \epsilon_t(u), \hat{\psi}_j(u) \right\rangle$ is noise. It is clear that two sources generate the profits of jth eigenfunction portfolio, namely, a deterministic part and a stochastic part. The former is an inner product between $\hat{\psi}_j(u)$ and the functional mean $\mu(u)$, and the latter is a time-varying functional loading that projects $r_t(u)$ on the jth eigenfunction $\hat{\psi}_j(u)$. Since cross-sectional behavioral patterns in one market are likely to be formed in the long run, we can easily obtain profits by taking the deterministic part. But these behavioral patterns might vary in the short run, so that $r_{p_j,t}$ has to take a risk (variance) from the stochastic part. One potential way to improve the profitability of the eigenfunction portfolio is to predict the functional loadings; this will be difficult and more studies will be required.

7. Conclusion

This paper employed a functional principal component analysis approach to decompose cross-sectional Fama-French three-factor model residuals in the Chinese stock market. Based on our empirical results, two asset pricing factors were derived from the market data, which we interpreted as the momentum effect and the asymmetric disposition effect, respectively. According to behavioral finance theories, we explain these two data-driven factors through investors' overreaction, overconfidence, and the lead-lag effect. These two factors, combined with the Fama-French three factors explain almost all variation of cross-sectional returns in the Chinese stock market. Our findings empirically verified the existence of the momentum effect and the disposition effect in the Chinese stock market, and the orthogonality of empirical eigenfunctions suggests that the disposition effect could not explain the momentum effect. By incorporating these cross-sectional patterns into the portfolio selection, we found that the Sharpe ratio of the mean-variance optimized portfolio could be boosted to 1.27. Our functional data analysis approach can be generally applied to any market of interest, particularly to markets suffering from high volatilities. Future works can focus on the following issues: understanding the role of higher order eigenfunctions; investigating the relationship between data-driven factors and macroeconomic variables; and modeling market anomalies by using a functional multi-factor model.

Acknowledgements. We are grateful to Mr. Curtis Miller for his carefully reading our paper. We also thank the editor, Cheng-Few Lee, one anonymous referee, whose detailed and constructive comments helped to improve the quality of the paper.

References

[1] An L (2015) Asset pricing when traders sell extreme winners and losers. Rev. Financ. Stud., 29, 823-861.

[2] Asness CS, Moskowitz TJ, Pedersen LH (2013) Value and momentum everywhere. J. Finance, 68, 929-985.

[3] Barberis N, Shleifer A, Vishny R (1998) A model of investor sentiment. J. Financ. Econom., 49, 307-343.

[4] Barberis N, Xiong W (2009) What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. J. Finance, 64, 751-784.

[5] Ben-David I, Hirshleifer D (2012) Are investors really reluctant to realize their losses? trading responses to past returns and the disposition effect. Rev. Financ. Stud., 25, 2485-2532.

[6] Birru J (2015) Confusion of confusions: a test of the disposition effect and momentum. Rev. Financ. Stud., 28, 1849-1873.

[7] Blitz DC, Van Vliet P (2007) The volatility effect. J. Portfolio Manag., 34, 102-113.

[8] Blitz DC, Huij J, Martens M (2011) Residual momentum. J. Empir. Finance, 18, 506-521.

[9] Bondt WFM, Thaler R (1985) Does the stock market overreact? J. Finance, 40, 793-805.

[10] Cakici N, Chan K, Topyan K (2015) Cross-sectional stock return predictability in China. Eur. J. Financ., 1-25.

[11] Carhart MM (1997) On persistence in mutual fund performance. J. Finance, 52, 57-82.

[12] Chan LKC, Jegadeesh N, Lakonishok J (1995) Evaluating the performance of value versus glamour stocks the impact of selection bias. J. Financ. Econom., 38, 269-296.

[13] Chan SH, Kim KA, Rhee, SG (2005) Price limit performance: evidence from transactions data and the limit order book. J. Empir. Finance, 12, 269-290.

[14] Chen G, Kim KA, Nofsinger JR, Rui OM (2007) Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. J. Behavioral Decision Making, 20, 425-451.

[15] Cheema MA, Nartea GV (2014) Momentum returns and information uncertainty: Evidence from China. Pac.-Basin Finance J., 30, 173-188.

[16] Cooper MJ, Gutierrez RC, Hameed A (2004) Market states and momentum. J. Finance, 59, 1345-1365.

[17] Daniel K, Titman S (1997) Evidence on the characteristics of cross sectional variation in stock returns. J. Finance, 52, 1-33.

[18] Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: Empirical tests. J. Polit. Econ., 607-636.

[19] Fama EF, French KR (1992) The cross-section of expected stock returns. J. Finance, 47, 427-465.

[20] Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J. Financ. Econom., 33, 3-56.

[21] Fama EF, French KR (2012) Size, value, and momentum in international stock returns. J. Financ. Econom., 105, 457-472.

[22] Fama EF, French KR (2015) A five-factor asset pricing model. J. Financ. Econom., 116, 1-22.

[23] Frazzini A (2006) The Disposition Effect and Underreaction to News. J. Finance, 61, 2017-2046.

[24] Frazzini A, Pedersen LH (2014) Betting against beta. J. Financ. Econom., 111, 1-25.

[25] Gandhi P, Lustig H (2015) Size anomalies in us bank stock returns. J. Finance, 70, 733-768.

[26] Gibbons MR, Ross SA, Shanken J (1989) A test of the efficiency of a given portfolio. Econometrica: J. Econometric So., 2515-2547.

[27] Griffin JM, Ji X, Martin JS (2003) Momentum investing and business cycle risk: Evidence from pole to pole. J. Finance, 2515-2547.

[28] Grinblatt M, Han B (2005) Prospect theory, mental accounting, and momentum. J. Financ. Econom., 78, 311-339.

[29] Hall P, Hosseini-Nasab M (2006) On properties of functional princi- pal components analysis. J. Royal Stat. Society: Series B, 68, 109-126.

[30] Hörmann S, Kokoszka P (2010) Weakly dependent functional data. *The Annals of Statistics* 38, 1845-1884.

[31] Horváth L, Kokoszka P (2012) Inference for functional data with applications. Springer.

[32] Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: Implications for stock market efficiency. J. Finance, 48, 65-91.

[33] Jegadeesh N, Titman S (2001) Profitability of momentum strategies: An evaluation of alternative explanations. J. Finance, 56, 699-720.

[34] Kang J, Liu M, Ni SX (2002) Contrarian and momentum strategies in the china stock market: 1993-2000. Pac.-Basin Finance J., 10, 243-265.

[35] Kokoszka P, Miao H, Reimherr M, Taoufik B (2018) Dynamic Functional Regression with Application to the Cross-section of Returns. J. of Financ. Econometrics, 16, 461–485.

[36] Ni ZX, Wang DZ, Xue WJ (2015) Investor sentiment and its nonlinear effect on stock returns: New evidence from the Chinese stock market based on panel quantile regression model. Econ. Model., 50, 266–274.

[37] Lo AW, MacKinlay AC (1990) When are contrarian profits due to stock market overreaction? Rev. Financ. Stud., 3, 175-205.

[38] Ramsay JO, Silverman BW (2006) Functional Data Analysis. Wiley Online Library.

[39] Shefrin H, Statman M (1985) The disposition to sell winners too early and ride losers too long: Theory and evidence. J. Finance, 40, 777-790.

[40] Tan L, Chiang T, Mason JR, Nelling E (2008) Herding behavior in Chinese stock markets: An examination of A and B shares. Pac.-Basin Finance J., 16, 61-77.

[41] Wong KA, Tan RSK, Liu W (2006) The cross-section of stock returns on the shanghai stock exchange.

Rev. Quant. Financ. Account., 26, 23-39.

[42] Wu Y (2011) Momentum trading, mean reversal and overreaction in Chinese stock market. Rev. Quant. Financ. Account., 37, 301-323.

Table 1: Results for F-F three-factor and Carhart four-factor models

Table 1 shows the OLS regression estimates on the past returns-sorted equal-weighted excess returns and asset pricing factors. Each column represents excess portfolio returns from the first decile (loser) to the last decile (winner). The returns on the winner-minus-loser portfolio are in column "*wml*". Panel A shows the estimation results for the three-factor Fama French model (Fama and French 1992). Panel B gives the estimation results of Carhart's four-factor model with an additional *wml* factor (Carhart, 1997). The final two columns show the F statistics and P values of the GRS test (Gibbons et al., 1989). *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

	Loser	2	3	4	5	6	7	8	9	Winner	*wml*	F(GRS)	p(GRS)
Panel A: Fama-French model													
α	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01^{*}	-0.02^{***}	-0.01^{**}		
rmrf	1.10^{***}	1.09^{***}	1.10^{***}	1.09^{***}	1.08^{***}	1.07^{***}	1.08^{***}	1.07^{***}	1.04^{***}	1.06^{***}	-0.02		
smb	0.97^{***}	0.92^{***}	0.88^{***}	0.91^{***}	0.89^{***}	0.84^{***}	0.90^{***}	0.80^{***}	0.85^{***}	1.71^{***}	0.37^{***}	4.82	0.0000
hml	-0.92^{***}	-1.31^{***}	-1.44^{***}	-1.57^{***}	-1.54^{***}	-1.69^{***}	-1.78^{***}	-1.86^{***}	-1.75^{***}	-0.79	0.45^{*}		
$AdjR^2$	0.83	0.85	0.85	0.86	0.86	0.86	0.85	0.86	0.82	0.79	0.15		
Panel B: Carhart model													
α	-0.01^{**}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01^{**}	0.00		
rmrf	1.08^{***}	1.08^{***}	1.09^{***}	1.09^{***}	1.08^{***}	1.07^{***}	1.08^{***}	1.07^{***}	1.04^{***}	1.08^{***}	0.00		
smb	1.24^{***}	1.13^{***}	1.05^{***}	1.05^{***}	0.99^{***}	0.89^{***}	0.90^{***}	0.73^{***}	0.64^{***}	1.24^{***}	0.00^{**}	3.65	0.0003
hml	-0.87^{***}	-1.28^{***}	-1.41^{***}	-1.54^{***}	-1.53^{***}	-1.68^{***}	-1.78^{***}	-1.87^{***}	-1.78^{***}	-0.87^{***}	0.00		
wml	-0.73^{***}	-0.58^{***}	-0.45^{***}	-0.39^{***}	-0.28^{**}	-0.15	-0.02	0.20	0.56^{***}	1.27^{***}	1.00^{***}		
$AdjR^2$	0.87	0.88	0.86	0.87	0.86	0.87	0.86	0.86	0.84	0.89	1.00		

Table 2: The percentage of variations explained by the first four eigenfunctions

The columns show the percentage of variability accounted for by the first four eigenfunctions of cross-sectional returns. The last column documents the total variations explained by the first four eigenfunctions. The first row shows the percentages computed from the real data, and the second row shows the percentage computed from the bootstrap with 5,000 replications.

	1st eigenfunction	2nd eigenfunction	3rd eigenfunction	4th eigenfunction	Total
Real data	64.62%	14.73%	5.16%	3.26%	87.77%
Bootstrap	74.72%	14.78%	5.22%	3.39%	88.11%

Table 3: Fama-MacBeth regression on $FPC1_t$ and $FPC2_t$

The estimated coefficients of the Fama-MacBeth regression on $FPC1_t$ and $FPC2_t$ with dependent variable past returns-sorted equally weighted portfolio returns are documented. First, we computed the factor risk exposures for each portfolio by regressing the portfolio returns on the risk factors. Second, we obtained coefficients for risk premiums by regressing the portfolio returns on the risk exposures month to month. To measure the strength of risk premiums instead of their directions, we used the absolute value of their coefficients. Bracketed values are Newey-West robust t-statistics.

	Jan.2007-Jun.2017				Jan.2007-Jun.2017				Jan.2007-Jun.2017			
	Entire	Jan	Feb-Nov	Dec	Entire	Jan	Feb-Nov	Dec	Entire	Jan	Feb-Nov	Dec
$FPC1_t$	0.13	0.18	0.12	0.15					0.13	0.17	0.12	0.14
	(3.43)	(4.56)	(3.23)	(4.45)					(3.99)	(5.53)	(3.75)	(5.09)
$FPC2_t$					0.07	0.08	0.07	0.07	0.07	0.08	0.07	0.07
					(1.92)	(1.87)	(1.93)	(1.92)	(2.47)	(2.93)	(2.41)	(2.62)
$AdjR^2$	0.17	0.28	0.16	0.20	0.08	0.09	0.08	0.09	0.25	0.37	0.24	0.26

Table 4: Results for multi-factor models with FPCs

The OLS estimates of the coefficients in regressing past returns-sorted excess portfolio returns on Fama-French risk factors, and functional momentum factors $FPC1_t$ and $FPC2_t$ are presented. Columns represent excess portfolio returns from the first decile (loser) to last decile (winner), and returns on the winner-minus-loser portfolio (wml). Panel A displays the three Fama-French risk factors and $FPC1_t$. Panel B replaces $FPC1_t$ with $FPC2_t$. Panel C considers both $FPC1_t$ and $FPC2_t$, forming a five-factor model. The final two columns present F statistics and P values of the GRS tests (Gibbons et al. 1989). *, ** and *** denote statistical significance at 10%, 5% and 1% level.

	Loser	2	3	4	5	6	7	8	9	Winner	*wml*	F(GRS)	p(GRS)
Panel A: F-F with $FPC1_t$													
α	−0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	−0.01*	0.00		
rmrf	1.07***	1.07***	1.09***	1.08***	1.07***	1.07***	1.08***	1.08***	1.06***	1.10***	0.02*		
smb	1.15***	1.07***	0.99***	1.01***	0.96***	0.87***	0.88***	0.74***	0.68***	1.39***	0.12***	2.99	0.0022
hml	−0.97***	−1.36***	−1.48***	−1.60***	−1.56***	−1.69***	−1.78***	−1.84***	−1.70***	−0.70*	0.14**		
$FPC1_t$	−0.12***	−0.10***	−0.08***	−0.07***	−0.05*	-0.02	0.01	0.04*	0.11***	0.22***	0.17***		
$AdjR^2$	0.87	0.87	0.86	0.86	0.85	0.85	0.85	0.86	0.85	0.89	0.95		
Panel B: F-F with $FPC2_t$													
α	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	−0.01***	−0.02***	−0.01***		
rmrf	1.09***	1.09***	1.10***	1.09***	1.08***	1.07***	1.08***	1.07***	1.04***	1.06***	-0.02		
smb	0.82***	0.99***	0.97***	1.09***	1.10***	1.09***	1.17***	1.09***	1.18***	1.55***	0.36***	3.44	0.0006
hml	−1.18***	−1.17***	−1.29***	−1.24***	−1.17***	−1.22***	−1.29***	−1.36***	−1.16*	−1.09*	0.05		
$FPC2_t$	0.08**	-0.04	-0.05	−0.10***	−0.12***	−0.14***	−0.16***	−0.16***	−0.18***	0.09	0.00		
$AdjR^2$	0.84	0.85	0.85	0.87	0.87	0.88	0.87	0.88	0.84	0.79	0.15		
Panel C: F-F with FPC1 and FPC2													
α	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
rmrf	1.07***	1.07***	1.08***	1.08***	1.07***	1.07***	1.08***	1.08***	1.06***	1.10***	0.02*		
smb	1.01***	1.15***	1.09***	1.19***	1.17***	1.13***	1.17***	1.02***	1.00***	1.21***	0.10***		
hml	−1.22***	−1.21***	−1.31***	−1.26***	−1.18***	−1.23***	−1.28***	−1.34***	−1.13***	−1.01***	0.10*	2.16	0.0243
$FPC1_t$	−0.12***	−0.10***	−0.08***	−0.07**	−0.05*	-0.02	0.00	0.04	0.11***	0.22***	0.17***		
$FPC2_t$	0.08**	-0.05	-0.05	−0.10***	−0.12***	−0.14***	−0.16***	−0.16***	−0.18***	0.09*	0.01		
$AdjR^2$	0.87	0.88	0.87	0.88	0.87	0.88	0.87	0.88	0.88	0.89	0.95		

Table 5: Relative strength of momentum and contrarian portfolios

Table 5 shows the relative strength of the average monthly returns from momentum and contrarian portfolio strategies. Formation periods are denoted by K, which equals to 3, 6, 9 and 12 months for short-run past returns and 60 months for long-run past returns. Holding periods vary from 3 to 24 months in each column. We apply the method proposed in Jegadeesh and Titman (1993) to construct equally weighted portfolios with overlapping holding periods. Bracketed values are Newey-West robust t-statistics.

	Momentum					Contrarian				
K	3	6	9	12	24	3	6	9	12	24
3 months	-0.0030	-0.0028	-0.0025	-0.0025	-0.0024	0.0028	0.0025	0.0023	0.0023	0.0023
	(-1.52)	(-1.90)	(-2.36)	(-2.85)	(-4.48)	(1.42)	(1.78)	(2.19)	(2.63)	(4.24)
6 months	-0.0042	-0.0039	-0.0036	-0.0034	-0.0032	0.0042	0.0038	0.0035	0.0034	0.0031
	(-2.30)	(-2.85)	(-3.14)	(-3.43)	(-4.23)	(2.25)	(2.78)	(3.05)	(3.34)	(4.15)
9 months	-0.0036	-0.0032	-0.0030	-0.0028	-0.0023	0.0034	0.0030	0.0028	0.0026	0.0021
	(-2.00)	(-2.55)	(-2.74)	(-2.85)	(-2.77)	(1.92)	(2.39)	(2.55)	(2.66)	(2.55)
12 months	-0.0042	-0.0039	-0.0036	-0.0034	-0.0027	0.0038	0.0034	0.0032	0.0029	0.0022
	(-2.19)	(-2.78)	(-3.09)	(-3.25)	(-3.09)	(2.00)	(2.45)	(2.68)	(2.81)	(2.58)
60 months	-0.0025	-0.0023	-0.0022	-0.0021	-0.0018	0.0018	0.0013	0.0011	0.0008	0.0002
	(-1.64)	(-2.21)	(-2.48)	(-2.63)	(-2.89)	(1.14)	(1.21)	(1.15)	(0.99)	(0.47)

Table 6: Serial and cross-Serial correlations

Serial and cross-serial correlations in the past returns-sorted portfolio returns are displayed. Panel A describes the serial correlations up to lag 4. Panel B describes the cross-serial correlation with lag 1. The correlation between lagged loser returns and contemporary winner returns is 0.06, and the correlation between lagged winner returns and contemporary loser returns is 0.02. The sample includes stocks without missing data from January 2007 to June 2017 (1307 stocks).

Panel A: Serial Correlations

	Loser	G2	G3	G4	G5	G6	G7	G8	G9	Winner
Lag1	-0.08	-0.04	-0.09	-0.05	0.01	0.01	-0.10	-0.02	-0.07	-0.04
Lag2	-0.07	-0.06	-0.07	-0.18	-0.15	0.04	-0.13	-0.02	-0.10	-0.08
Lag3	-0.12	-0.05	-0.15	-0.13	0.01	-0.06	0.04	-0.08	-0.16	-0.09
Lag4	-0.01	-0.12	0.03	-0.06	-0.12	-0.16	-0.06	-0.15	0.04	-0.04

Panel B: Cross-serial Correlations

	Loser	G2	G3	G4	G5	G6	G7	G8	G9	Winner
Loser(Lag1)	-0.11	-0.08	0.18	-0.08	-0.15	-0.10	-0.03	0.05	0.08	0.06
G2(Lag1)	0.11	0.03	-0.04	0.00	0.10	-0.05	0.08	-0.12	0.06	-0.06
G3(Lag1)	0.01	0.23	-0.03	0.04	-0.15	0.11	-0.04	-0.22	0.03	-0.09
G4(Lag1)	-0.15	0.11	0.10	-0.02	-0.04	0.02	-0.06	-0.01	-0.19	-0.09
G5(Lag1)	0.20	0.00	0.12	-0.07	-0.03	-0.14	-0.04	0.07	0.08	0.04
G6(Lag1)	-0.08	0.01	0.02	-0.06	-0.15	0.17	-0.01	-0.06	0.03	-0.14
G7(Lag1)	0.12	0.12	-0.01	0.01	-0.03	0.12	0.00	0.03	-0.11	-0.03
G8(Lag1)	0.21	-0.11	-0.01	0.04	-0.16	-0.05	0.03	0.12	0.05	0.16
G9(Lag1)	-0.08	0.07	-0.02	0.17	-0.10	0.09	-0.15	-0.16	-0.08	-0.18
Winner(Lag1)	0.02	-0.02	0.05	0.01	0.05	-0.13	-0.03	-0.12	-0.01	-0.06

Table 7: Turnover ratios in pair-wise groups

Turnover ratios for different groups are shown. 'Extreme' implies the loser and winner groups; 'Mediocre' implies groups 5 and 6. G2/9, G3/8 and G4/7 imply the combinations of groups 2 and 9, 3 and 8, and 4 and 7, respectively.

	G2/9	G3/8	G4/7	Mediocre
Extreme	1.3200	1.3859	1.4007	1.4124
	Extreme	G3/8	G4/7	Mediocre
G2/9	0.7780	1.0446	1.0523	1.0615
	Extreme	G2/9	G4/7	Mediocre
G3/8	0.7558	0.9676	1.0093	1.0164
	Extreme	G2/9	G3/8	Mediocre
G4/7	0.7535	0.9626	0.9970	1.0106
	Extreme	G2/9	G3/8	G4/7
Mediocre	0.7495	0.9573	0.9898	0.9963

Table 8: Turnover ratios in individual groups

Turnover ratios for individual groups are given. Each sub-panel shows turnover ratios between the subject of its row and each remaining group.

	G2	G3	G4	G5	G6	G7	G8	G9	Winner
Loser	**1.0862**	**1.0744**	**1.0615**	**1.0402**	**1.0209**	**0.9896**	**0.9682**	**0.9089**	**0.6385**
	Loser	G3	G4	G5	G6	G7	G8	G9	Winner
G2	0.9451	0.9881	0.9709	0.9520	0.9351	0.9079	0.8887	0.8362	0.5879
	Loser	G2	G4	G5	G6	G7	G8	G9	Winner
G3	0.9691	1.0242	0.9858	0.9654	0.9465	0.9196	0.8997	0.8481	0.5983
	Loser	G2	G3	G5	G6	G7	G8	G9	Winner
G4	0.9977	1.0494	1.0278	0.9846	0.9640	0.9347	0.9144	0.8601	0.6064
	Loser	G2	G3	G4	G6	G7	G8	G9	Winner
G5	1.0211	1.0739	1.0506	1.0274	0.9831	0.9530	0.9307	0.8746	0.6157
	Loser	G2	G3	G4	G5	G7	G8	G9	Winner
G6	1.0519	1.1063	1.0800	1.0547	1.0309	0.9748	0.9512	0.8941	0.6317
	Loser	G2	G3	G4	G5	G6	G8	G9	Winner
G7	1.0879	1.1471	1.1205	1.0915	1.0668	1.0408	0.9795	0.9162	0.6468
	Loser	G2	G3	G4	G5	G6	G7	G9	Winner
G8	1.1226	1.1838	1.1555	1.1253	1.0980	1.0700	1.0321	0.9368	0.6583
	Loser	G2	G3	G4	G5	G6	G7	G8	Winner
G9	1.2245	1.2972	1.2691	1.2326	1.2015	1.1723	1.1247	1.0910	0.7001
	Loser	G2	G3	G4	G5	G6	G7	G8	G9
Winner	**1.8426**	**1.9600**	**1.9267**	**1.8759**	**1.8245**	**1.7850**	**1.7123**	**1.6543**	**1.5080**

Table 9: Statistics of portfolios

Average monthly returns, Sharpe ratios and the maximum draw-down of five portfolios are displayed. The first two columns are the benchmarks market index and contrarian-lmw portfolios. Considering that the momentum-wml portfolio is unprofitable and the contrarian-lmw is profitable, the portfolio weight for $EP1$ is $-\hat{\psi}_1$ instead of $\hat{\psi}_1$, and the weight for EP2 is $\hat{\psi}_2$. The last three columns represent portfolios $EP1$, $EP2$ and their mean-variance optimization. Assuming no transaction costs, we hold all portfolios from January 2007 to June 2017.

	Benchmark Portfolios		Eigenfunction Portfolios		
	Market index	Contrarian	$EP1$	$EP2$	Optimized portfolio
average return	0.0168	0.0182	0.0236	0.0248	0.0264
Sharpe ratio	0.0550	0.4757	0.4936	0.3272	1.2736
maximum draw-down	-1.2368	-0.0773	-0.1245	-0.2141	-0.0208

Figure 1: The 13 B-spline basis functions

The figure plots 13 basis functions used on 100 grouped cross-sectional returns. Nine dashed lines are interior breakpoints or knots, splitting 100 groups into ten intervals and requiring 13 basis functions for smoothing.

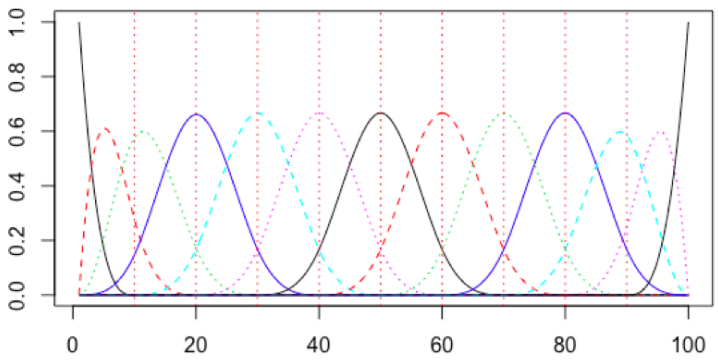

Figure 2: Functional Residual Curves and Decomposed Eigenfunctions

Functional objects and the corresponding first four estimated eigenfunctions are displayed. The upper sub-figure shows 126 functional curves, which are cross-sectional past returns-sorted residuals smoothed by a 13 cubic B-spline smoother. The lower sub-figure shows the first four estimated eigenfunctions computed from 126 functional objects through FPCA. The first four account for 87.77% of the total variation.

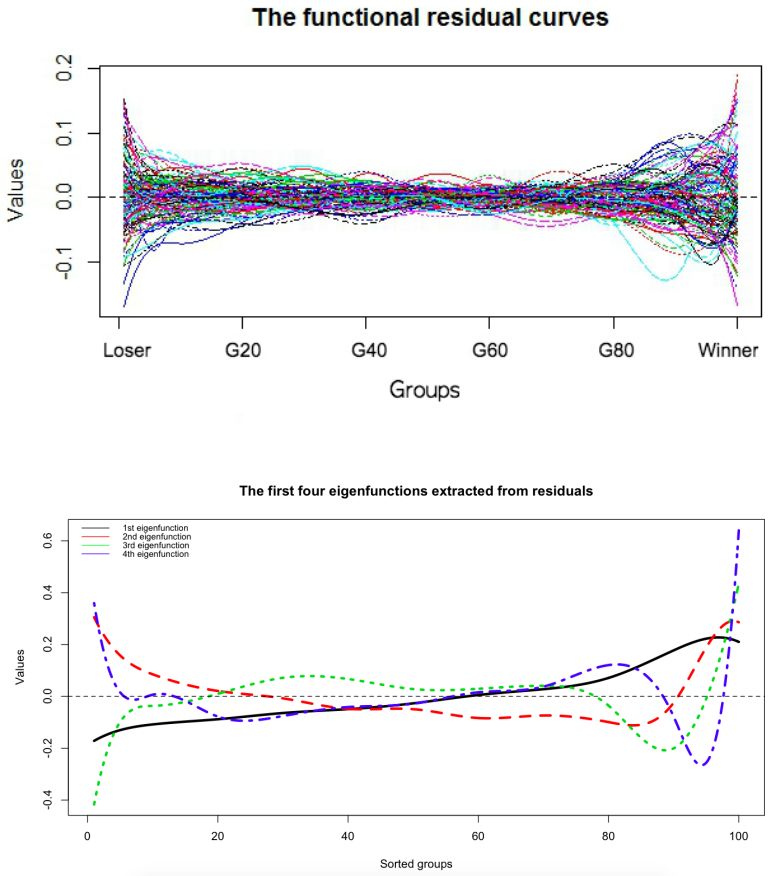

Figure 3: Confidence envelopes of the first four eigenfunctions

Figure 3 displays 95% confidence envelopes of the eigenfunctions. By bootstrapping functional cross-sectional residual curves 5,000 times, we computed 95 % confidence envelopes for the first four functional eigenfunctions.

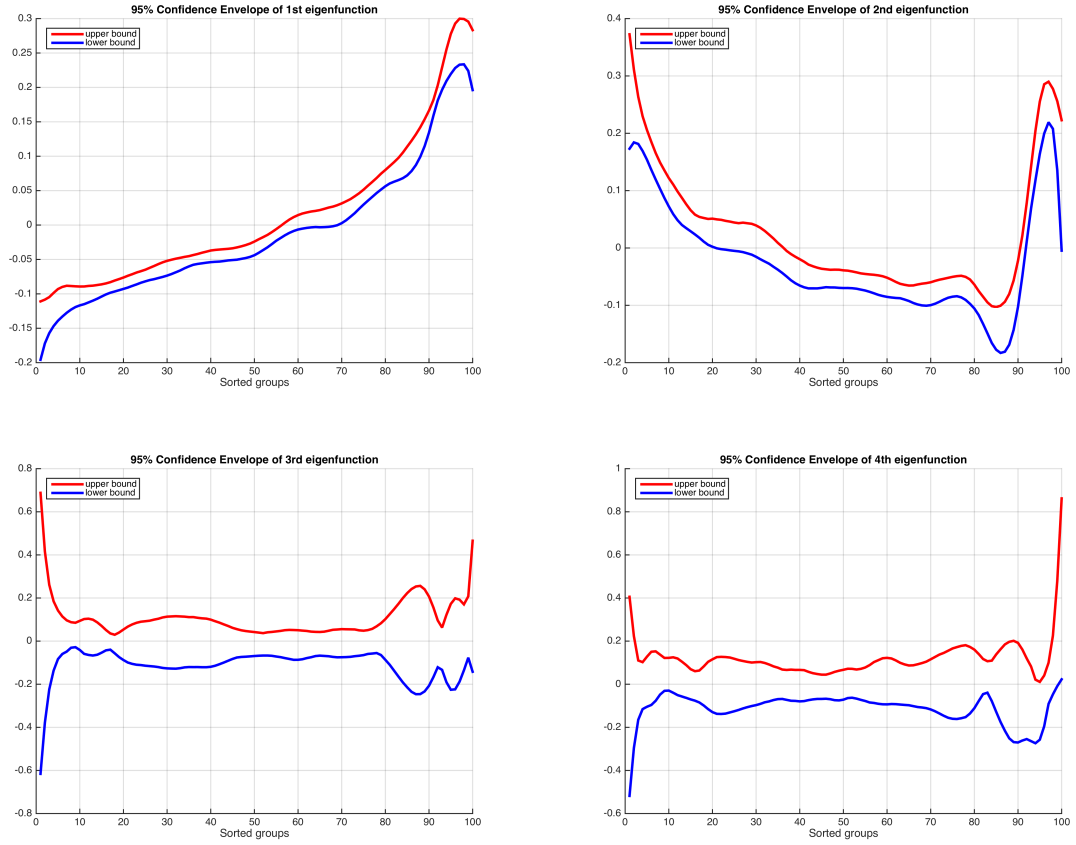

Figure 4: Information coefficients

The figure plots information coefficients between past returns-sorted cross-sectional returns and the first two eigenfunctions. There are 126 information coefficients from January 2007 to June 2017. To assess the strength of the correlation, we set the domain of the coefficients to be between 0 and 1 by taking their absolute values.

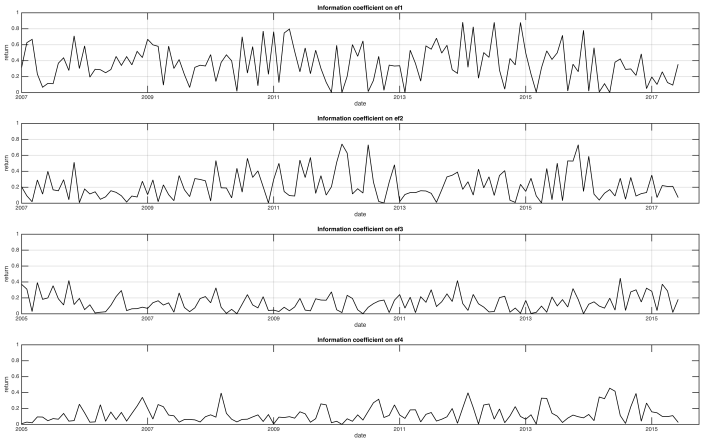

Figure 5: Asymmetric turnover surface

The turnovers of cross-sectional shares sorted by firm-specific 12-month momentum are displayed. The x-axis measures groups from loser to winner, and splitting into deciles. The y-axis denotes periods between January 2007 and June 2017. The z-axis shows the value of turnovers.

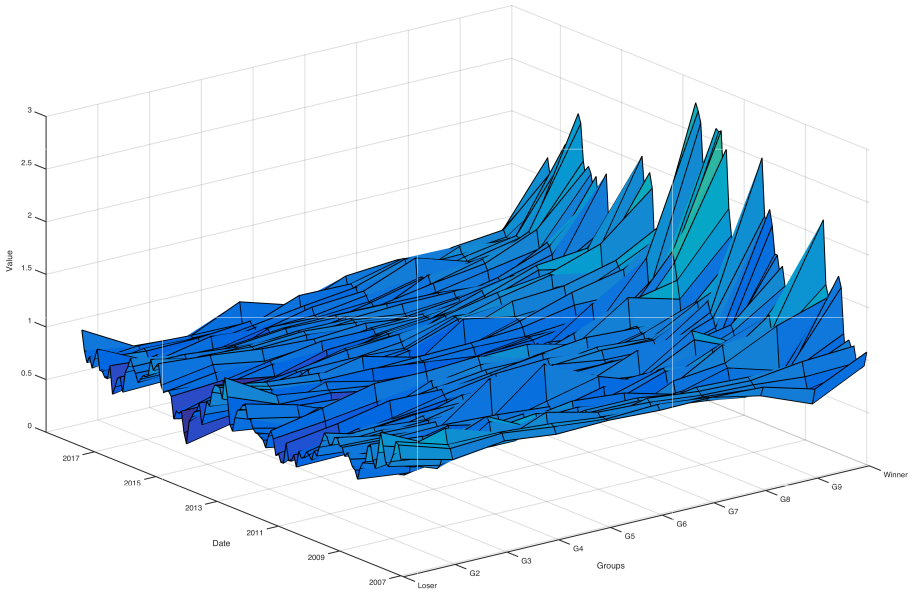

Figure 6: Portfolio performance

Cumulative portfolio returns and monthly average returns for different portfolios are shown. The upper sub-figure shows the performance of the contrarian, $EP1$, and $EP2$ portfolios and their optimization. The lower sub-figure displays the monthly average return performance. Terms "-1st ef" and "2nd ef" indicate $EP1$ and $EP2$, respectively.