# Risk Factor Timing and Mutual Fund Performance

Manuel Ammann[a], Sebastian Fischer[b], Florian Weigert[c]

October, 2018

**Abstract**

We investigate the relationship between a mutual fund's risk factor timing activity and its performance. Using a dynamic state space version of Carhart (1997)'s four factor model, we measure timing activity by the fund's variation in factor loadings to the market, size, book-to-market, and momentum factor. Our results indicate no evidence of positive timing skill. We find that a portfolio of funds with high risk factor timing activity underperforms a portfolio of funds with low timing activity by 147 basis points p.a. The results are important in the light of recent discussions about the predictability of asset pricing risk factors.

JEL Classification: G11, G14, G20, G23

**Keywords: Mutual Fund, Market Timing, Factor Timing, Kalman Filter**

We thank Yakov Amihud, the participants of the Annual Meeting of the German Finance Association 2017, the Cologne Colloquium on Financial Markets 2018, the Lancaster Frontiers of Factor Investing Conference 2018 and the participants of the Finance Seminars at the Universities of Constance and St. Gallen for their helpful comments. All errors are our own.

[a] Swiss Institute of Banking and Finance, University of St. Gallen, Email: manuel.ammann@unisg.ch

[b] Swiss Institute of Banking and Finance, University of St. Gallen, Email: sebastian.fischer@unisg.ch;
Address: s/bf-HSG, Unterer Graben 21, 9000 St.Gallen, Switzerland
Phone: +41 71 224 70 15

[c] Swiss Institute of Banking and Finance, University of St. Gallen, Email: florian.weigert@unisg.ch

## 1. Introduction

Among academics, there is a widely accepted consensus that mutual funds, on average, generate small positive abnormal gross returns, but fail to beat a risk-adjusted benchmark net of fees.[1] Therefore, the focus of the academic mutual fund literature has moved to the question which investment and fund characteristics lead to future abnormal returns and whether there are indicators that ex-ante identify top-performers. To achieve the goal of future benchmark-adjusted outperformance, a fund manager can generally pursue three different investment approaches. First, she can expose the fund to alternative risk factors, such as liquidity risk (Pástor and Stambough, 2003, and Dong et al., 2017), volatility risk (Ang et al., 2006), or tail risk (Kelly and Jiang, 2014, and Chabi-Yo et al., 2018) to earn the associated risk premium.[2] Second, she can deviate from the benchmark portfolio and engage in stock picking, i.e., tilt her portfolio towards stocks that are likely to outperform in the future (see Wermer, 2000, and Cremers and Petajisto, 2009). Third, the fund manager can attempt to time systematic risk factors, i.e., increase (decrease) her exposure to a risk factor when it is likely that it pays a high (low) premium in the future. Our paper is concerned with the latter investment approach and examines a fund manager's timing ability to different systematic risk factors in a comprehensive framework.[3]

To measure risk factor timing activity of mutual funds, we propose to apply the Carhart (1997) four-factor model with dynamic, i.e. time-varying, factor exposures that follow a mean-reverting process. We choose to apply the Carhart (1997) model because it is the current academic stand-

[1] Among others, Malkiel (1995), Gruber (1996), Carhart (1997), and Fama and French (2010) document that mutual funds underperform their respective benchmark net of fees. Using detailed portfolio holdings data, Grinblatt and Titman (1989, 1993) and Daniel et al. (1997) observe that gross of fees, mutual funds generate positive abnormal returns. Wermers (2000) combines both views and shows that mutual funds exhibit positive stock picking ability, which is – however – too low to cover expenses and transaction costs.

[2] Of course, this approach is only suitable if the mutual fund's benchmark does not account for these alternative risk factors.

[3] Our paper is concerned with studying a fund's timing ability to several risk factors at the same point in time. Other papers that examine timing ability to specific risk factors include Treynor and Mazuy (1966), Henriksson and Merton (1981), Ferson and Schadt (1996), and Kacperczyk and Seru (2007) in the case of market timing, Busse (1999), Giambona and Golec (2009), and Kim and In (2012) in the case of volatility timing, Bodnaruk et al. (2014) in the case of downside risk timing, and Bazgour et al. (2017) in the case of liquidity timing.

ard used to measure mutual fund performance.[4] Its systematic risk factors are the market return (MKT) factor, the size (SMB) factor, the book-to-market (HML) factor, and the momentum (UMD) factor. To estimate this model we use a Kalman filter and Kalman smoother technique. We apply the model to a period of 3 years of weekly return data in a rolling manner and measure factor timing activity by the volatility of the factor loadings during this estimation period with regard to the MKT factor, the SMB factor, the HML factor, and the UMD factor. To express a fund's overall level of factor timing we then compute an aggregated (overall) Timing Indicator by averaging and standardizing the individual market, size, value, and momentum timing measures.[5]

To better understand the concept of risk factor timing and the relevance of our aggregate Timing Indicator, we provide an example of two large and well-established equity mutual funds, the TIAA-CREF Growth & Income Fund and the Calamos Growth Fund in the time period from 2002 to 2016 in Figure 1.[6] Both funds follow a similar investment style and Morningstar classifies both as US Large Cap Growth Equity funds. However, comparing the two funds' factor loading volatilities reveals significant differences with regard to their risk factor timing. Whereas the

TIAA-CREF Growth & Income Fund's market beta measured within single calendar years between 2002 and 2016 varies between 0.95 to 1.09 ($\Delta$=0.14), Calamos Growth Fund's market beta fluctuates between 0.86 and 1.25 ($\Delta$=0.39) during the same time. The exposures to the SMB factor (-0.15 to 0.03 for the TIAA-CREF Growth & Income Fund vs. -0.12 to 0.67 for the Calamos

---

[4] See, e.g, Berk and van Binsbergen (2015) and Barber et al. (2016). Our results are stable for alternative factor models such as the Fama and French (1993) three-factor model and the Fama and French (2015) five-factor model.

[5] Ideally, we would like to restrict our sample to funds that intend to time risk factors as opposed to those who vary their factor exposure unintentionally. However, fund managers do not have to report their investment strategy in such a detailed way and even when they do, this description is likely to be misleading (see Sensoy, 2009, for the case of self-designated benchmark indices in the mutual fund industry). Given the popularity of the Carhart (1997) factor model, however, it is unlikely that fund managers are not aware of their exposure towards the four different risk factors. We therefore argue that even an unintended, but tolerated variation of factor exposures as a result of unsolicited trading should be considered factor timing.

[6] The TIAA-CREF Growth & Income Fund was incepted in 1997, the Calamos Growth Fund in 1990. At the end of 2016, USD 5.6 bn. were invested in the TIAA-CREF Growth & Income Fund, while the Calamos Growth Fund has total net assets of USD 1.8 bn.

Growth Fund), the HML factor (-0.19 to 0.05 vs. -0.80 to 0.42) and the UMD factor (-0.02 to 0.13 vs. -0.20 to 0.56) support the impression that the Calamos Growth Fund has more volatile risk factor exposures than the TIAA-CREF Growth & Income Fund. Figure 1 plots the two funds' Timing Indicators throughout our sample period with positive (negative) values indicating an above (below)-average overall timing activity.

[Insert Figure 1 around here]

This figure displays that the Calamos Growth Fund displays higher overall timing activity than the TIAA-CREF Growth & Income Fund over the sample period. In addition, we observe that differences in Timing Indicators can be traced back to differences in fund characteristics: In particular, the Calamos Growth Fund has a higher turnover ratio than the TIAA-CREF Growth & Income Fund (90% vs. 83% in 2016) and a less diversified portfolio (79 vs. 189 stock holdings as of the end of 2016). Finally, when comparing the performance of both funds, we observe that the TIAA-CREF Growth & Income Fund, which has little timing activity, outperformed the actively timed Calamos Growth Fund by 1.6% per year between 2002 and 2016.[7]

Based on this example, we ask whether performance differences between funds with high timing activity and funds with low timing activity are systematic. For this purpose, we investigate the relationship between factor timing activity and fund performance for a large sample of US equity mutual funds in the time period from the late 2000 up to 2016. Our results reveal the following main findings: First, we show that factor timing is a persistent fund characteristic, i.e., funds that are sorted into decile portfolios with the lowest (highest) factor timing in year *t* have a likelihood of 79% (75%) to remain in the lowest (highest) three deciles in year *t+3*. Second, and most importantly, risk factor timing is associated with future fund underperformance. A portfolio of the 20% funds with the highest Timing Indicator underperforms the 20% funds with the lowest Timing Indicator by risk-adjusted 147 basis points p.a. with statistical significance at the 1% level.

[7] The average yearly performance of the primary share classes was 7.2% for the TIAA-CREF Growth & Income Fund and 5.6% for the Calamos Growth Fund in our sample period.

Similarly, sorting funds on individual MKT-, HML-, or UMD-timing measures, results in underperformance of the most actively timed funds by 102, 82, and 120 basis points p.a. with statistical significance at least at the 5% level.[8] Third and finally, we provide evidence that the relationship between factor timing and future underperformance is not driven by the factor volatility of the fund's equity portfolio holdings nor by asset sales and purchases that result from investors' in- and outflows. Instead, we find that that risk factor timing is particularly prevalent among funds with long management tenure, high turnover, high total expense ratios, and high past fund inflows. These findings indicate that factor timing is driven by unsolicited trading of the fund manager and support the notion that (i) manager behavior is influenced by career concerns of young managers having no incentive to expose their portfolios to unsystematic risk (Chevalier and Ellison, 1999), (ii) risk factor timing is an actively enforced and expensive investment strategy, and (iii) risk factor timing is pursued by fund managers who were successful in the past, earn high inflows, and become overconfident in their trading decisions and risk factor forecasts (Puetz and Ruenzi, 2011).

The question whether mutual funds can successfully time risk factor exposures has so far mainly been studied in the context of market timing. Treynor and Mazuy (1966, TM) have been the first to adapt a non-linear market beta model to explore market timing abilities. The intuition behind their analyses is that a fund manager with such an ability will increase her exposure to the market when the market return is higher. Neither their research nor the majority of subsequent studies, e.g., by Henriksson and Merton (1981, HM), Ferson and Schadt (1996) or Kacperczyk and Seru (2007), find evidence for average market timing ability of mutual funds. To the contrary, more recent studies provide at least some evidence for successful market timing, such as in Mamaysky et al. (2008) and Jiang et al. (2007). Bollen and Busse (2001) find some market timing ability when applying the TM and HM measures to daily instead of monthly return data and controlling

[8] Funds with high SMB-timing underperform funds with low SMB-timing by 61 basis points p.a. The performance spread between high and low SMB-timing funds is statistically significantly indifferent from zero.

for spurious results. Moreover, Kacperczyk et al. (2014) show positive market timing ability of fund managers during recessions, while Elton et al. (2012) find positive market timing skill when using a one factor model (which disappears when applying a more complex factor model).

The literature on timing ability beyond the market factor is rather scarce. Investigating changes in fund holdings, Daniel et al. (1997) observe that mutual fund managers do not possess timing abilities with respect to stock characteristics and Benos et al. (2010), who extend the analysis of Bollen and Busse (2001) to a Carhart (1997) model, do not find factor timing abilities either. In contrast, Swinkels and Tjong-a-Tjoe (2007) detect positive risk factor timing skills within a very small US fund sample when applying the TM and HM measures to a four-factor model. Busse (1999), Giambona and Golec (2009), and Kim and In (2012) examine volatility timing of mutual funds, while Bodnaruk et al. (2014) document downside risk timing ability of some fund managers. Finally, Huang et al. (2011) document that funds that intensively shift their total risk exposure over time underperform funds with a stable risk level.[9] Like Benos et al. (2010) and Swinkels and Tjong-a-Tjoe (2007), our analysis focuses on the timing of risk factor exposures. In contrast to earlier research, we do not estimate timing ability from an ex-post perspective using the TM and HM models, but measure timing activity by the fund's volatility to different risk factors on a rolling basis.

Our research contributes to the mutual fund literature on market and risk factor timing threefold. First, our proposed measure of factor timing can directly assess a fund's timing activity whereas earlier models, e.g. TM and HM, only observe performance effects of timing activity. Thus, whereas existing models cannot distinguish funds with no timing activity from funds with excessive timing but no average return effect of this timing, our model allows us to directly observe a

[9] Empirical investigation with regard to timing abilities are not restricted to active mutual fund managers. Bhattacharya et al. (2017) find that individuals deteriorate their performance, among others, due to poor ETF timing.

fund's timing activity. This also enables us to observe high persistence in factor timing activity as an investment characteristic.[10]

Second, our model allows us to simultaneously estimate a fund's timing activity with respect to different risk factors. The vast majority of prior research on timing ability of mutual funds focuses on market timing only. Hence, our results of a negative return effect of overall timing activity goes beyond the most prominent findings of no positive market timing skill. We also contribute to the literature on fund activeness as timing is one element of activeness and is closely linked to – yet not covered by – earlier developed activeness measures such as the Amihud and Goyenko (2013) selectivity measure or the Huang et al. (2011) risk shifting measure.

Third, we contribute to the ongoing debate among academics and investment management practitioners, whether (and how) risk factors can be timed. Numerous papers suggest factor timing strategies, such as Barroso and Santa-Cara (2015) and Moreira and Muir (2017), who show that volatility predicts the momentum and other alternative risk premiums. Among others, Asness et al. (2000) and Arnott et al. (2016) advocate using risk factors' value spread as a signal to time factors. Yet, the question whether those results can be exploited out of sample remains unsolved. Asness (2016) articulates doubts about the performance of risk factor timing. We contribute to this discussion by documenting that professional and sophisticated investors, such as mutual fund managers, are unsuccessful in the timing of risk factors.

The reminder of this paper is structured as follows. Section 2 describes the data and introduces our measure of factor timing activity. Section 3 links factor timing to mutual fund performance and Section 4 examines the drivers of factor timing. Section 5 concludes.

## 2. Data and the Factor Timing Indicator

[10] Jiang et al. (2007) note that the measures of TM and HM are subject to artificial timing biases and propose a holding-based measure. Our measure does not require any fund holding data, which might be difficult to access for most investors and which is generally only available on a low, quarterly frequency. The holding-based measure of Jiang et al. (2007) furthermore only captures the covariance of factor exposure and performance, whereas our approach measures factor-timing activity directly.

In this section we describe the data used in this study and discuss the methodology of the empirical analysis. We also provide summary statistics for the overall Timing Indicator and examine its persistence.

### *2.1. Data selection*

We investigate the relationship between risk factor timing and performance using a sample of actively managed US equity mutual funds. We select our fund universe from the CRSP survivorship-bias free mutual fund database and use daily net returns as well as quarterly updated fund characteristics in the empirical analysis. We start our data selection process with all mutual funds included in the CRSP survivorship-bias free mutual fund database during the 1998 - 2016 time period. This time window is determined by the availability of daily fund returns. We use Objective Codes from CRSP and Lipper as well as the Strategic Insights Objective Code to determine fund styles and assign each fund to either *Growth and Income, Growth, Income, Hedged, Mid Cap, Small Cap* or *Micro Cap*.[11] Funds that cannot be matched to one of these categories as well as funds with missing fund names are dropped from our sample. We exclude index funds, balanced funds, international funds, and sector funds according to the CRSP Index Fund Flag, CRSP Objective Code and by screening fund names for key terms such as “balanced” or “index”. We additionally exclude funds with less than 70% of equity holdings and funds with total net assets of less than 15 million USD. This leaves us with a total number of 3,816 funds in the sample.

We obtain quarterly data on fund age, management tenure, turnover ratio, total expense ratio and total assets under management as well as daily net returns for our sample funds and aggregate those data across all share classes of each fund. Fund age is the age of the oldest share class, total net assets are the sum of the total net assets of all share classes and turnover ratio, total expense ratio and daily returns are the weighted means of single share classes’ data, weighted by the

[11] We find that actively managed funds that mainly invest into large caps or whose name contains strings that indicate a large cap investment strategy are mostly classified as *Growth* or *Growth and Income*.

share classes' total net assets. We additionally calculate 12-months fund flows for each fund by $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1year,t)}))$, where $tna_t$ are the total net assets at time t and $ret_{(t-1year,t)}$ is the 1-year return (net of fees) during the past 12 months. We winsorize the data on age, tenure, turnover, expense ratio, flows, and total net assets at the 1%-level. For a sub-analysis in Section 4.1 of the paper, we also use equity portfolio holding implied returns. To calculate these returns, we obtain quarterly holding data from CRSP and use the securities' historical cusip number to link it to daily stock returns from CRSP.

For our empirical analysis, we aggregate daily returns into weekly as well as monthly data. Following Bollen and Busse (2001), we measure risk factor timing based on weekly returns. Our performance analysis is then based on monthly returns. Since we do not have monthly observations on fund characteristics, we assign the last available data point to each fund if it is not older than 12 months. We calculate weekly and monthly Fama and French (1993) as well as momentum risk factors from daily data, which we obtain from Kenneth R. French's website.[12] We also collect monthly data for the Fama and French (2015) five-factor model as well as a short and long term reversal factor from Kenneth R. French's website. In addition, we gather data on the Frazzini and Pedersen (2014) betting against beta factor from AQR, data for the Baker and Wurgler (2006) sentiment factor from Jeffrey Wurgler's website and data for the Pástor and Stambaugh (2003) liquidity factor from WRDS.[13]

## 2.2. *Factor Timing in a Dynamic Factor Model*

Traditional asset pricing factor models such as the capital asset pricing model (CAPM), the Fama and French (1993) three factor model, and the Carhart (1997) four factor model assume a linear relationship between an asset's excess return and the respective factor premia. The size of this

[12] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[13] Data for the betting against beta factor is retrieved from https://www.aqr.com/Insights/Datasets/Betting-Against-Beta-Equity-Factors-Monthly and data for market sentiment from http://people.stern.nyu.edu/jwurgler/.

relationship, represented by β, is traditionally assumed to be constant over time, which allows estimating values of β using an OLS regression framework. Even if this assumption of constant βs holds for single securities it might not be valid for managed portfolios such as mutual funds, as pointed out by Mamaysky et al. (2008), because any varying exposure due to timing attempts would not be reflected correctly. We model such timing attempts by applying the Carhart (1997) four factor model with time-varying risk factor exposures $\beta_t$, which is represented by the following state space model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{RMRF,i,t} * (r_{m,t} - r_{f,t}) + \beta_{SMB,i,t} * SMB_t + \beta_{HML,i,t} * HML_t + \beta_{UMD,i,t} * UMD_t + \varepsilon_{i,t},$$

$$\beta_{j,i,t} = \beta_{j,i,t-1} + \theta_{j,i}(\mu_{j,i} - \beta_{j,i,t-1}) + \eta_{j,i,t} \quad \text{for} \quad j \in \{RMRF, SMB, HML, UMD\},$$

where $r_{m,t}$ is the market return, $r_{f,t}$ the risk-free rate at time *t* and $SMB_t$, $HML_t$ and $UMD_t$ denote the Fama and French (1993) and Carhart (1997) risk factors at time *t*. The model differs from a classical Carhart (1997) model as it allows the factor loadings to change over time. In our main empirical specification, we assume the factor loadings to follow a mean-reverting process with four time-invariant mean factors $\mu$ (one with respect to each risk factor). The four time-invariant values of $\theta$ indicate the pace at which the loadings revert to its mean. Those values are unknown and estimated empirically together with the values of $\beta_t$. Forcing $\theta = 0$ leads to a model that assumes risk factor loadings to follow a random walk as introduced by Black et al. (1992). In our robustness check, we re-calculate our results enforcing this random walk. Our results remain qualitatively unchanged and remain statistically significant. The disturbance terms $\varepsilon_{i,t}$ and $\eta_{j,i,t}$ are normally distributed with zero mean and unknown standard deviations.

For each month we calculate a fund's factor timing activity. To do so, we apply the model to the past three years of weekly fund return data and use a Kalman filter and Kalman smoother technique to estimate the dynamics of all unknown parameters.[14] This yields a time series of 156 weekly values of $\beta_{RMRF}$, $\beta_{SMB}$, $\beta_{HML}$ and $\beta_{UMD}$ per fund in the three-year period. For each of the four βs, we compute the standard deviation across time, i.e. $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$. These standard deviations express the volatility of the fund's exposure to the respective risk factors during the three-year period. We therefore consider them as the direct measures of a fund managers' engagement in timing the four risk factors: Generally, a higher $\sigma(\beta)$ indicates a more actively timing fund with regard to a certain risk factor.[15]

To express a fund's overall level of factor timing with respect to all the risk factors we aggregate the four measures to one overall Timing Indicator. We determine this Timing Indicator as follows: At each point in time, we calculate the cross-sectional mean and standard deviation for each factor timing measure $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ and standardize all estimated values of $\sigma(\beta)$ by demeaning (using the cross-sectional mean) the estimates and dividing them by the respective cross-sectional standard deviation. Our Timing Indicator is then defined as the average of the four standardized values, i.e.,

$$Timing\ Indicator = \frac{1}{4}\left(\frac{\sigma(\beta_{RMRF})-\overline{\sigma(\beta_{RMRF})}}{SD(\sigma(\beta_{RMRF}))} + \frac{\sigma(\beta_{SMB})-\overline{\sigma(\beta_{SMB})}}{SD(\sigma(\beta_{SMB}))} + \frac{\sigma(\beta_{HML})-\overline{\sigma(\beta_{HML})}}{SD(\sigma(\beta_{HML}))} + \frac{\sigma(\beta_{UMD})-\overline{\sigma(\beta_{UMD})}}{SD(\sigma(\beta_{UMD}))}\right),$$

where $\overline{\sigma(\beta)}$ is the cross-sectional mean and $SD(\sigma(\beta))$ the cross sectional standard deviation of $\sigma(\beta)$. Subsequently, we will refer to a fund's risk factor timing measured over the past three years ending at time *t* as the fund's timing activity at time *t*. We will investigate the relationship between future fund performance and a fund's timing activity in Section 3.

### 2.3. *Summary Statistics and Factor Timing Persistence*

[14] We shortly describe the Kalman filter and the Kalman smoother technique in the Appendix. Within each three-year window we require funds to have at least 104 weekly return observations.

[15] To prevent outliers influencing our empirical tests, we censor observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ or $\sigma(\beta_{UMD})$ are among the highest 1% of all observations.

Daily fund returns – and hence, calculated weekly returns for our empirical tests – are available from CRSP by the end of 1998. We calculate our timing measures from past three years' net returns. If more than two but less than three years of data are available, we calculate the timing measures using the available data. Therefore, our final dataset reaches from the end of 2000 to 2016. It contains 300,519 observations and 3,816 distinct funds. Table 1 provides summary statistics for the main variables of the empirical analysis.

[Insert Table 1 around here]

Average and median fund sizes are 1,329 and 324 million USD, which indicate a skewed distribution of size across funds. On average, the age of a fund is 15.7 years and management has been in office for 7.5 years. The average turnover ratio is 75% per year, but there is a wide variance ranging from 3% to 342%. Total expenses range from 0.14% p.a. to 2.23% p.a. with a mean of 1.15%. The average yearly flow is positive (2.0% of past TNA) but its median is at -6.0% suggesting that there are high net inflows into few funds but smaller net outflows from the majority of funds. All four estimated parameters of $\sigma(\beta)$ show a pronounced heterogeneity in timing activity ranging from a very stable factor exposure ($\sigma(\beta)<0.0001$) to values as large as 4.2 times the average $\sigma(\beta)$.[16] The mean variation in factor loading ($\sigma(\beta)$) is highest for the HML risk factor, followed by the SMB-, the UMD-, and the market risk factor, which is in line with results of Engle (2016) who finds betas of industry portfolios to vary over time with the HML being the most volatile. As expected and by construction, the average Timing Indicator is close to 0, but there are some funds that are very active in risk factor timing (maximum Timing Indicator = 4.57) and some funds with very little timing activity (minimum Timing Indicator = -1.80). Panel B reports the timing estimates by fund style. Mid Cap, Small Cap, and Micro Cap funds tend to have a higher timing activity than Growth, Growth and Income, and Income funds. The row "other" summarizes very few observations of funds that were classified as large cap funds as

[16] The maximum values of the market, SMB, HML, and UMD timing measures are 0.42, 0.74, 1.00, 0.51.

well as funds that have been included in our sample but whose assigned styles change during the sample period.

[Insert Table 2 around here]

Table 2 reports the average cross-sectional correlations between the four timing measures (Panel A) as well as between the Timing Indicator and fund characteristics (Panel B). The correlation between single risk factor timing measures ranges from 0.20 to 0.33, thus indicating that the timing activity with regard to a single risk factor does not strongly imply timing activities with respect to other risk factors. Funds with high timing activity (measured by a high Timing Indicator) tend to be smaller, more expensive and show a higher turnover ratio. These results provide evidence that factor timing is not a randomly occurring observation but an intended active trading strategy. We investigate the relationship between factor timing and fund characteristics more thoroughly in Section 4.3 of the paper.

[Insert Figure 2 around here]

Figure 2 plots the time-series of equally-weighted average timing measures over all funds in our sample. Measures of timing with respect to the market, SMB and UMD risk factors appear to be relatively stable over time whereas HML timing slightly peeks during the pre-crises years and after 2013. Overall, factor timing seems to be an investment strategy that is prevalent in different market situations and periods of economic booms and recessions.

We also investigate the persistence of factor timing activity. If factor timing is related to mutual fund returns, long-term investors can only profit from this result if factor timing is a stable fund characteristic rather than a quickly changing investment trend. To study persistence, we sort funds into ten deciles by their Timing Indicator. We do so every month and leave those decile portfolios unchanged to observe the average value of the Timing Indicator during the 12 months prior and the 72 months after the formation period.

[Insert Figure 3 around here]

Figure 3 displays this time-series of average Timing Indicator values of funds sorted in decile portfolios. We find that the difference in Timing Indicators becomes smaller during the 12 months before and 36 months after the portfolio formation, but no two decile portfolios cross lines or converge to a common value. Even after 36 months, from where on the calculation window of the Timing Indicators does not overlap with the calculation window of the Timing Indicators at the formation period, the average Timing Indicators of the decile portfolios remain in an unchanged order. Thus, we conclude that the persistence of factor timing activity remains strong even in the long run (i.e., for a period up to 6 years in the future).

[Insert Table 3 around here]

The transition matrix in Table 3 underlines this conclusion numerically. It displays the likelihood that a fund sorted in decile portfolio *i* in year *t*, appears in decile portfolio *j* in year *t* and year *t+3*, respectively. Our results indicate that about 60% of all funds in the lowest (highest) timing decile remain in the lowest (highest) decile after one year and over 90% of all funds in the lowest (highest) timing decile remain in the lowest (highest) three deciles after one year. This might partially be by construction since factor timing has been estimated over a three-year time window. Panel B therefore displays transitions aver a period of three years. Results do not change qualitatively. After three years, 45% (41%) of the funds in the lowest (highest) timing activity decile still remain in this decile and 79% (75%) of all funds in the lowest (highest) timing decile remain within the lowest (highest) three deciles after three years.

We also provide summary statistics of the attrition rate, that is, the percentage of funds that leave our sample within the following one year or the following three years, respectively. Funds with a higher factor timing activity are more likely to drop from our sample within the next years. Only 6% (16%) of all the funds in the lowest timing activity decile leave our sample within the next year (three years), but the probability increases as factor timing activity increases and reaches 14% (26%) for the 10% of the funds with the highest measures of factor timing.

In summary, Section 2 displays summary statistics of the main variables in our study and shows that overall risk factor timing is a persistent characteristic of a mutual fund. Moreover, we find that funds with high timing activity are more likely to drop from our sample and that the correlation between individual factor timing measure is moderate. Hence, high timing activity of an individual factor does not necessarily imply high timing activity to another factor.

## 3. Factor Timing and Mutual Fund Performance

This section investigates the relationship between timing activity and future mutual fund performance. We examine univariate portfolio sorts in Section 3.1, multivariate Fama-MacBeth regressions in Section 3.2, and bivariate portfolio sorts in Section 3.3. We perform additional robustness checks and document the stability of our main results in Section 3.4.

### *3.1. Univariate portfolio sorts*

We are interested in the relationship between timing activity, measured by $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ as well as the overall *Timing Indicator*, and the future performance of mutual funds. We start by applying univariate portfolio sorts to investigate this relationship. Each month *t*, we sort all funds in our sample by factor timing with respect to either a specific risk factor (i.e., by either $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$, $\sigma(\beta_{UMD})$) or by the Timing Indicator and assign them to five quintile portfolios, each portfolio holding one fifth of all funds. As factor timing differs significantly between fund styles, we sort the funds within the same style, thus ensuring that the number of funds of a certain fund style is the same for all five quintile portfolios. We keep these portfolios unchanged for one month and calculate the quintile portfolio returns in month *t+1* as the equal-weighted mean of the funds' returns within this portfolio. We resort the portfolios every month by the most recent timing measure and therefore obtain a monthly return time series for each quintile portfolio.

[Insert Table 4 around here]

Table 4 reports the average abnormal, risk-adjusted returns of these portfolios with each column referring to a specific sorting criterion. As our asset pricing model for the risk-adjustment, we use the Carhart (1997) four-factor factor model. We specifically examine the differences in abnormal returns between funds with high and low timing activity, i.e., funds that are sorted in portfolio five and portfolio one according to each measure.

Our results reveal that the risk-adjusted spread between funds with high and low timing activity is negative and statistically significant (at least at the 5% significance level) for market, value, and momentum timing as well as for the overall Timing Indicator. Funds in the most actively timed portfolio underperform the funds in the least actively timed portfolio in terms of abnormal returns by 102 (market factor), 82 (value factor), 120 (momentum factor) and 147 (overall Timing Indicator) basis points p.a., respectively. Furthermore, the abnormal returns decrease monotonically in the market, value, momentum and the overall timing measures. The relationship between size timing and abnormal returns is also negative, yet statistically not significant.[17]

To rule out that these results are driven by other risk factors and/or the choice of the factor model, we repeat the portfolio sorts for the overall Timing Indicator and calculate each quintile's abnormal return for different alternative asset pricing models in Table 5. Again, we focus to interpret the results of the (5) – (1) difference portfolio between funds with high timing activity and low timing activity.

[Insert Table 5 around here]

To control for additional risk factors, we use the one-factor CAPM model, the Fama and French (1993) three-factor model, the Fama and French (2015) five-factor model and the Fama and French (1993) model plus a short term and a long term reversal factor provided by Kenneth French's homepage. We also apply the Carhart (1997) model including either the Frazzini and Pedersen (2014) betting against beta factor, the Baker and Wurgler (2006) sentiment factor or the

[17] Notably, no single quintile portfolio has a positive alpha. This is not surprising as we use net returns and funds are known to show, on average, significantly negative abnormal return after fees. The abnormal return is particularly low for funds in the highest timing quintile when sorted by any timing measure.

Pástor and Stambaugh (2003) liquidity factor in alternative specifications. We find that our results remain qualitatively unchanged and statistically significant for almost all alternative factor models (while getting even more significant for some of the additional models). Solely the Fama and French five-factor model reduces the return difference between low and high timing funds to 86 basis points. Although the level of statistically significance is slightly above the 10% level for the Fama and French five-factor model, the unexplained return difference remains economically significant. We thus conclude that the underperformance of risk factor timing by mutual funds is not explained by alternative asset pricing risk factors.

### *3.2. Fama-Macbeth regressions*

To check whether there is a negative impact of risk factor timing on performance when controlling for different fund characteristics at the same time, we proceed to investigate the relationship between factor timing and future fund returns using Fama-MacBeth regressions. We calculate a fund's abnormal return at month t, $\alpha_t$, as the difference between the actual fund performance during this month and the expected fund performance calculated from a Carhart (1997) model, that is $\alpha_t = r_{i,t} - \mathrm{E}[r_{i,t}]$, where

$$\mathrm{E}[r_{i,t}] = r_f + \beta_{mkt,i,t} * (r_{m,t} - r_{f,t}) + \beta_{smb,i,t} * SMB_t + \beta_{hml,i,t} * HML_t + \beta_{mom,i,t} * UMD_t,$$

and $\beta_t$ are estimated by an OLS regression over the previous three years of weekly return data.[18] We conduct Fama-MacBeth regressions with abnormal returns during the next month (or cumulated over the next six and twelve months) as the dependent variable and different fund characteristics as independent variables. As independent variables, we use a fund's ln(TNA), ln(fund

[18] We also obtain estimates of $\beta_t$ when applying the dynamic factor model during the same three-year period. Using those estimates of $\beta_t$ instead yields qualitatively unchanged results.

age), ln(manager tenure), expsenses, turnover, lagged alpha and past fund flows. Table 6 reports our results using Newey-West standard errors with a lag of 1 month and style dummies.

[Insert Table 6 around here]

Specification (1) reports that measures of market, size, value, and momentum timing have, on average, a negative effect on abnormal returns. This effect is statistically significant for market, size and momentum timing. In specification (2), we pool the individual measures to our overall Timing Indicator. The average effect of the Timing Indicator on future abnormal returns is negative and statistically significant at the 1% level. We also show that this result holds for the six-month and twelve-month abnormal returns in specifications (3) and (4). As a side note, we verify already established relationships between fund characteristics and performance in our multivariate regressions. In particular, we document a significantly negative relationship between fund size (expenses) and performance as well as a significantly positive relationship between past performance and performance.

We also analyze the economic impact of our results. The average cross-sectional standard deviation of market, size, value and momentum timing measures are 0.06, 0.12, 0.14, and 0.09. Thus, a one standard deviation increase of market, size, value, and momentum timing leads to a decrease of annualized abnormal returns by 35, 38, 15, and 22 basis points p.a. The economic impact of the overall timing measure is also substantial: Specification (2) reports that a one standard deviation increase of timing reduces abnormal future returns by 71 basis points p.a.

To demonstrate that our results are stable and do not depend on a specific economic environment, we split our sample in different subsets and repeat the Fama-MacBeth regressions as in Specification (2) of Table 6. We split the 192 sample months by business cycle into 166 months of expansion and 26 months of recession as defined by the NBER. We also split the sample into months with a positive and negative market risk premium and additionally consider a subsample that excludes the months of the financial crises, that is from November 2007 to February 2009.

We find that, throughout all subsets, risk factor timing as measured by the Timing Indicator is associated with lower future abnormal performance, as reported in Table 7.

[Insert Table 7 around here]

In addition to this result, we find that the coefficient estimate of the Timing Indicator is more negative during recessions, months with a negative market performance, and less negative for the sample excluding the financial crisis. The level of statistical significance varies across sub-periods, which is partially due to the decreased number of observations within the subsets.

*3.3. Bivariate portfolio sorts*

Factor Timing can be viewed as a form of fund manager activeness, which has already been linked to mutual fund performance in prior research. Amihud and Goyenko (2013) show that a low $R^2$ obtained from an OLS regression of fund returns on a Carhart (1997) model predicts future fund returns. They interpret this low $R^2$ as selectivity and claim that a higher selectivity might indicate a fund manager's conviction resulting from superior skill. Opposed to that, Huang et al. (2011) find that mutual funds, that change their risk levels significantly over time underperform mutual funds with a more stable risk level. The authors suggest that risk shifting might be either an indication of inferior manager ability or a result of agency issues.

To investigate whether the negative relationship between factor timing activity and mutual fund performance persists beyond those other measures ($R^2$ and risk shifting), we perform bivariate portfolio sorts based on the overall Timing Indicator and the measures of activeness. We calculate a fund's $R^2$ following Amihud and Goyenko (2013) from an OLS regression of net returns on a Carhart (1997) model but use three years of weekly return data to comply with the calculation of our timing measures. For those funds for which holding data are available from CRSP, we also calculate the holding based risk shifting measure following Huang et al. (2011).[19]

[19] Mutual fund holding data is available from CRSP starting in December 2001 for few funds and from June 2002 for a larger sample of funds. We need three years of data to calculate risk shifting and therefore our subsample for any analysis using holding data and risk shifting starts in July 2004 only. This reduces the overall number of fund-month observations to 233,251.

Whereas the risk shifting measure of Huang et al. (2011) is hardly correlated to our Timing Indicator (ρ=0.04, there is a considerable correlation between the Timing Indicator and the Amihud and Goyenko $R^2$ (ρ= -0.55).

We perform the bivariate portfolio sorts as follows: Each month we sort the funds by one of the two measures ($R^2$ and risk shifting) into five quintiles. As before, we define quintiles per style category to ensure an equal distribution of fund styles within each quintile. Within each of the resulting quintiles we sort funds by their Timing Indicator and form quintiles such that we end up with two sets of 5x5 quintile portfolios. We keep the portfolios unchanged over the next month and calculate the respective portfolio returns. For each quintile portfolio we calculate the abnormal return using thr Carhart (1997) model. Table 8 displays the results where the first sorting is done by either $R^2$ (Panel A) or the risk shifting measure (Panel B).

[Insert Table 8 around here]

Again, we are particularly interested in the difference between the portfolios of high and low timing activity formed within each $R^2$ or risk shifting portfolio, respectively. When looking at the results in Panel A, we observe that the quintiles with the highest timing indicator underperform the quintiles with the lowest timing indicator in all cases and the difference is statistically significant at the 1% level for almost all $R^2$-quintiles and significant at the 10% level for those with the lowest $R^2$. This result is important, because there is a negative relationship between $R^2$ and our timing indicator by construction. Funds with a high timing indicator have a volatile loading on risk factors and thus, a static risk factor model might not explain much of the return volatility and will have a low $R^2$ when estimated using an OLS regression. By showing that a high timing indicator is associated with lower abnormal returns even within $R^2$-quintiles, we provide strong evidence that our timing indicator measures a return pattern not captured by the $R^2$ measure.

Sorting on risk shifting and factor timing in Panel B supports our findings from above: Funds with a high timing indicator underperform funds with the lowest timing indicator in every risk-shifting quintile. The effect is statistically and economically particularly prevalent among funds with a low risk shifting measure. Hence, the return pattern due to timing activity is not subsumed by the effect of a fund manager's risk shifting (as measured by Huang et al., 2011).

### *3.4. Robustness tests*

We conduct a series of robustness tests to check that the negative relationship between factor timing activity and mutual fund performance remains strong when using value-weighted Fama-MacBeth regressions, using alternative performance measures, alternating the dynamics of our factor timing model or adding additional control variables. We adapt the Fama-MacBeth regressions presented in specification (2) of Table 6 and display the results of the stability checks in Table 9.

[Insert Table 9 around here]

In specification (1), we value-weight the funds during the first stage regressions of the Fama-MacBeth procedure. The results remain unchanged. Then, we regress alternative performance measures on the factor timing indicator and other fund characteristics. In Specification (2), we use the skill measure of Berk and van Binsbergen (2015), which measures the dollar value a fund manager generates, either presenting itself as a management fee or as over- or underperformance to the investor. Therefore, skill is defined as the product of fund size (total net assets) and the fund's gross excess return over the benchmark. We use the expected return from a Carhart (1997) factor model as a benchmark. We also apply a fund's Sharpe ratio and the manipulation-proof performance measure of Goetzmann et al. (2007) calculated from half a year of weekly returns as performance measures in specifications (3) – (5). For the latter, we set $\rho=2$ and $\rho=3$ to alternate the level of risk penalty. The relationship between a fund's factor timing activity – notably, measured during the period prior to the half year the performance measures were calculated for – remains negative and statistically significant at the 5% level.

Our dynamic factor timing model relies on an assumption about the underlying process of factor loadings and we assume a mean reverting process, that is for each fund i at time t:

$$\beta_{j,i,t} = \beta_{j,i,t-1} + \theta_{j,i}(\mu_{j,i} - \beta_{j,i,t-1}) + \eta_{j,i,t} \text{ for } j \in \{RMRF, SMB, HML, UMD\}.$$

As an additional robustness test, in Specification (6), we restrict this process to a random walk by setting $\theta_{j,i}$ to 0. This yields:

$$\beta_{j,i,t} = \beta_{j,i,t-1} + \eta_{j,i,t} \text{ for } j \in \{RMRF, SMB, HML, UMD\}.$$

We estimate the dynamics assuming this random walk, define factor timing by the variation of βs and calculate a corresponding version of the Timing Indicator as described in Section 2.2. The relationship between factor timing and mutual fund performance remains negative and economically and statistically significant at the 2%-level when using this alternative approach.

Another methodological alternative only considers the idiosyncratic variation of betas. That is, we estimate the dynamics of βs using a mean-reverting process. Instead of measuring the variation of βs over time we use the standard deviation of ηs as a measure of factor timing. For each of the four risk factors, the error term η is normally distributed and we take the standard deviation of these distributions as timing measures with respect to the four risk factors. As before, we calculate the Timing Indicator as the average of the cross-sectional standardized measures and use Fama-MacBeth regressions to determine the relationship between factor timing and fund performance in Specification (7). Again, we find a negative and statistically significant relationship. We additionally test the robustness of our results by adding the measures of fund activeness discussed in Section 3.3 as additional control variables in Specification (8). As a result, the relationship between factor timing and fund performance slightly weakens economically when including Amihud and Goyenko (2013)'s $R^2$ and Huang et al. (2011)'s risk shifting measure as independent variables, but it remains statistically significant at the 10%-level.

Finally, we perform a placebo test to examine the relationship between timing activity and fund performance for a sample of index funds. For these funds, any variation of risk factor exposure should be coincidental and we should not expect any relationship between timing activity and future performance. We exactly follow the data selection procedure from Section 2.1 but instead of dropping index funds, we solely keep index funds in our sample. We identify those funds by the index fund flag from CRSP and additionally hand-pick funds whose names include one of the terms “Index”, “S&P”, “Wilshire”, “Dow” or “Russell”. This leaves us with 631 index funds and 33,515 fund-month observations. Specification (9) shows the results of the Fama-MacBeth regressions on this index fund sample. As expected, the relationship between the Timing Indicator and future abnormal fund performance is close to zero. This result adds further credibility to our main result of a negative relation between risk factor timing and future fund performance.

To summarize, in this section we document that the relationship between risk factor timing and future risk-adjusted performance is negative in univariate portfolio sorts, multivariate regressions, and bivariate portfolio sorts when explicitly controlling for related measures. We confirm this result in a large battery of robustness checks and show that our results are not sensitive to several choices we make in our empirical analysis.

## 4. Drivers of factor timing

In this section, we analyze the drivers of mutual funds’ timing activity. We look at equity-induced factor timing in Section 4.1, and investigate whether factor timing is related to fund flows and thus induced by funds’ asset fire sales and purchases in Section 4.2. We finally relate factor timing to correlated fund characteristics in Section 4.3

### *4.1. Factor timing induced by time-varying factor exposures of equity holdings*

There might be two potential sources of a factor timing activity measured by our approach. On the one hand, a fund’s trading activity might cause the variation of βs if the fund management shifts holdings accordingly, for example between large cap and small cap stocks to vary its expose to the SMB factor. Section 4.2 will further break down this channel into forced and unforced trading. On the other hand, even a buy-and-hold strategy might have volatile risk factor exposures if the holdings' factor exposures vary over time. Prior research finds evidence consistent with the latter explanation: Armstrong et al. (2013) show that stock with high risk factor loading uncertainty with respect to the MKT factor, the SMB factor, the HML factor, and the UMD factor earn low future returns.[20] This pattern is likely to be present also on the fund level.

We aim to disentangle factor timing induced by changes in a fund's asset allocation from factor timing caused by the volatility of the holdings' factor exposure. Therefore, we calculate an additional set of timing measures directly imputed from mutual fund equity portfolio holdings. Most funds report holdings at the end of each quarter and we then calculate weekly returns during a quarter q as the weighted average stock returns during this week, weighted by the fund's portfolio weights as of the end of quarter q-1. [21] The underlying assumption of constant portfolio weights between reporting dates is, among others, in line with the implicit assumptions of the holding-based market timing approach of Jiang et al. (2007). This yields a return series, where short-term investment decisions and timing of trades of the fund manager remain unconsidered. As in Section 2.2, we apply the Kalman filter and smoother to estimate our dynamic version of the Carhart model for this holding-based return series instead of actual fund returns. As before, we compute the volatility of factor loadings with regard to MKT, SMB, HML, and UMD factor over a period of 156 weeks and form an overall Timing Indicator by averaging the standardized values of these factor timing measures . We then investigate whether this overall Timing Indicator calculated from fund holdings is also related to future abnormal returns of the fund using Fama-MacBeth regressions in Table 10.

[Insert Table 10 around here]

Specification (1) repeats the baseline regression setup of specification (2) in Table 6. In specification (2), we report the results of the relationship between factor timing based on equity hold-

[20] Opposed to this view, Lenz (2017) finds that stocks with more volatile market betas earn systematically higher returns than stocks with persistent market risk exposures.

[21] We do not consider a fund whenever the most recent holdings were reported more than one year ago and are missing in the upcoming quarters.

ings data and fund performance. In line with the results of Armstrong et al. (2013), we find that the association between the holdings-based Timing Indicator and future abnormal returns is significantly negative. However, we also observe that the coefficient estimate of the Timing Indicator decreases by more than 30% in comparison to the Timing Indicator based on actual net returns. If we use both Timing Indicators as explanatory variables in regression (3), we document that solely the coefficient of the Timing Indicator calculated from actual net returns remains statistically significant and is 3.4 times as large as the coefficient on the holding-based Timing Indicator.

Altogether, these results indicate that the holding-based Timing Indicator relates negatively to future abnormal returns; however, it cannot explain the negative association between the Timing Indicator calculated from actual net returns and future performance. Hence, we conclude that factor timing induced by fund managers' active trading decisions (as opposed to the volatility of fund holdings' risk factor exposures) is the main driver of a fund's Timing Indicator. Section 4.2 investigates whether this result might be driven by forced trading due to inflows and outflows rather than strategic or tactical asset allocation decisions.

### *4.2. Forced versus unforced trading*

Section 4.1 shows that the negative relationship between factor loading volatility and future fund performance is due to factor loading volatility that is induced by fund managers' trading. This section aims to distinguish between unsolicited trading and forced trading, i.e., trading that is required according to a fund's investor flows. If investors withdraw large amounts from a fund (or invest new money into the fund), the fund management will be forced to sell (or buy) assets and the risk factor exposure might vary as a result of this forced trading.[22] If the negative relationship between factor timing and fund performance is stronger and only present among funds

[22] Coval and Stafford (2007) discuss the phenomena of asset fire sales and purchases for mutual funds. They show that, among others, funds experiencing large outflows tend to decrease existing positions, which creates price pressure in the underlying securities held by the fund.

that experience large inflows or outflows, it might not be due to the disability of fund managers to time risk factors but due to investors' flows and resulting asset sales and purchases.

We measure factor timing over a three-year period and investigate the impact of contemporaneous flows, observed over the identical period. Hence, we compute a fund's three-year flow as the sum of yearly flows as described in Section 2.1. To detect the impact of fund inflows and outflows on the timing-performance relationship we construct three subsamples. One subsample consists of all fund-month observatons for which the three-year flow lies below the 30% quantile of three-year flows during the same time period. A second subsample consists of all observations with a three-year flow above the 70% quantile. All remainder funds, those with a medium three-year flow between the 30% and 70% quantile constitute a third subsample. We chose the 30% and 70% thresholds because we observe that values of three-year flows sharply increase after the 70% quantile. Within each subsample we repeat the Fama-MacBeth regression. Table 11 displays the results. If high outflows or inflows were driving our results, we would expect the relationship between factor timing and future fund performance to be particularly large for funds with low or high three-year-flows. The empirical results do not support this idea. In fact, the coefficient of the Timing Indicator is lowest for funds in the medium-flow sample, that is funds with moderate flows.

[Insert Table 11 around here]

Summing up flows over the previous three years might disguise cases where funds had to react to large inflows in one year and to large outflows during another year, i.e., these flows could level out each other. We therefore calculate an absolute flow measure as the sum of the absolute flow values during the three years.[23] We repeat the subsample analysis using this absolute flow. We consider subsamples with the 30% highest, 30% lowest and 40% median absolute flows. Additionally we use subsamples with the 15% and 5% highest absolute flows. If large inflows or

[23] As an example, a fund with yearly flows of +50%, -50% and +20% would have an absolute flow measure of 120%.

outflows drove our results, we would expect the timing-return-relationship to be stronger for funds with a higher absolute flow measure and would disappear for all other funds. However, the results in Panel B of Table 11 do not support this expectation. We thus conclude that the negative relationship between risk factor timing and future fund performance cannot be explained by fire sales and is mainly due to unsolicited trading decisions.

### *4.3. Fund determinants*

To understand which funds are pursuing active factor timing, we study the relationship between fund characteristics and the individual timing measures with regard to MKT, SMB, HML, and UMD as well as the overall Timing Indicator. Since timing measures are estimated using 3-year time windows during our 09/1998-12/2016 sample period, we split our sample into six non-overlapping sub-periods, namely 1999-2001, 2002-2004, 2005-2007, 2008-2010, 2011-2013, and 2014-2016. We regress the timing measures during those periods on the fund characteristics at the beginning of these periods to observe the relationship between ex-ante fund characteristics and timing activity. Table 12 reports the results of the multivariate regressions.

[Insert Table 12 around here]

Specifications (1) – (4) show the results with the individual timing measures as dependent variables, while specification (5) adapts the overall timing indicator as the dependent variables. We focus to interpret the results of regression (5) which documents a significant relationship between three sets of fund characteristics and the overall Timing Indicator.

First, we observe that risk factor timing is most common among funds that are old and that are managed by fund managers with long manager tenure. The result is in line with predictions of Chevalier and Ellison (1999) who suggest that manager's behavior is influenced by career concerns and that younger managers have an incentive to not expose their portfolios to unsystematic risk and hold more conventional portfolios. Second, risk factor timing is positively related to a fund's expenses and portfolio turnover. This finding is in line with Huang et al. (2011) as well as Amihud and Goyenko (2013), who document a positive relationship between a fund's expense ratio and turnover as well as their measures of fund activity. These relationships also confirm our results from Section 4.1 that our Timing Indicator captures an intended actively implemented investment strategy rather than a coincidental return series characteristic. Furthermore, the positive relationship between a fund's expense ratio and timing activity is either due to additional trading costs for the fund manager's trading strategy (e.g., due to high trading costs or research efforts) or it might indicate investors' willingness to pay for factor timing activity.[24] Finally, we observe that risk factor timing is pursued by fund managers who were successful in the past and have earned high inflows into their funds.[25] We argue that receiving new inflows can trigger higher exposure to active factor timing strategies due to (i) the availability of cash for new investment strategies, and (ii) changes in the mindset of (successful) managers who become overconfident and spend their money in costly active trading strategies (see Puetz and Ruenzi, 2011).[26]

To summarize, our results reveal that a part of the negative relationship between risk factor timing and future performance is due to factor-loading uncertainty of funds' stock holdings (see Armstrong et al., 2013). However, this effect only partly explains the negative association between our main Timing Indicator and future fund performance. We show that this Timing Indicator is strongly correlated to certain fund characteristics, such as fund manager tenure, fund's expenses and portfolio turnover, and past flows.

## 5. Conclusion

[24] Amihud/Goyenko (2013) have made this argument in the context of selectivity.

[25] This result does not contradict earlier findings. Flows do not change the timing-performance relationship but lead to a higher timing activity.

[26] We also include style dummies in our regressions and find that factor timing is higher for growth funds as well as mid, small and especially micro-cap funds.

Mutual fund managers vary their exposure to risk factors over time. To measure this investment pattern, we propose a new measure of factor timing activity that is based on a dynamic version of the Carhart (1997) four-factor factor model. Based on this measure, we investigate whether a variation in factor exposure is linked to fund performance within a sample of US mutual funds during the time period from the late 2000 up to 2016.

In this paper we find that factor timing is a persistent fund characteristic and associated with future underperformance. A portfolio of the 20% funds with the highest Timing Indicator underperforms the 20% funds with the lowest Timing Indicator by risk-adjusted 147 basis points p.a. with statistical significance at the 1% level. Similarly, sorting funds on individual MKT-, HML-, or UMD-timing measures, results in underperformance of the most actively timed funds by 102, 82, and 120 basis points p.a. with statistical significance at least at the 5% level. We also show that the underperformance is not explained by different risk factors, fund characteristics or similar activeness measures, such as the $R^2$–selectivity measure by Amihud and Goyenko (2013) or the Huang et al. (2011) risk shifting measure.

Our results also provide evidence that the relationship between factor timing and performance is mainly driven by fund managers' active trading decisions and less by the variation of single stocks' factor exposures. Moreover, it is not driven by asset sales and purchases in response to investment flows into or out of the fund. We show that risk factor timing is particularly prevalent among funds with long management tenure, high turnover and total expense ratio, and high past fund inflows. Our results do not support the hypothesis that deviations in risk factor exposures are a signal of skill and we recommend that investors should resist the temptation to invest in funds that intentionally or coincidentally vary their exposure to risk factors over time.

## Appendix: Kalman Filter

Kalman filtering was introduced to engineering in 1960[27]. The algorithm derives estimates of unobservable state variables from a time-series of observable variables that contains statistical noise. In our case, the unobservable state variables are the risk factor loadings, which are estimated from a return time-series. The Kalman filter requires a mathematical model that describes the dynamics of the unobservable state variables. In our main specification, we assume the factor loadings to follow a mean-reverting process.

The optimization follows a recursive two-step process. At each time t, the Kalman filter uses information up to time t to estimate the current state variables (i.e., factor loadings) as well as their uncertainties. It then uses the observed noisy measurement (i.e., the fund return) to update the estimate using a weighted average forecast. The algorithm gives more weight to estimates with lower uncertainty. In addition to the Kalman *filter* technique, we also apply a Kalman *smoother* in our estimations. The Kalman smoother additionally contains a backward procedure that utilizes observations that occur after time t to estimate state variables at time t. The Kalman smoother is more suitable to estimate the factor loading dynamics from an ex-post perspective. Rachev et al. (2007) provide an introduction to the Kalman filter and its application in finance. Racicot and Théore (2009) provide an overview over the historical use of Kalman filters in finance, which started in the 1980s. Black, Fraser, and Power (1992) have been the first to measure time-varying factor exposures via the Kalman filter and similar approaches have later been used e.g. by Wells (1994), Brunnermeier and Nagel (2004), Jostova and Philipov (2005), Swinkels and Van Der Sluis (2006), Mamaysky, et al. (2007) and Mamaysky, et al. (2008). An important difference among earlier studies and our paper is the assumed process of factor loadings. Whereas some papers assume a random walk, we follow Wells (1994), Jostova and Philipov (2005) who assume a mean-reverting process. This is also in line with the findings of

---

[27] See Kalman (1960).

Blake et al. (1999) who document a mean reversion in funds' portfolio weights within a sample of U.K. pension funds.

We execute the Kalman filter using adapted functions from the Jouni Helske's KFAS package (Helske, 2016) in the software environment R.

## References


Amihud Y., Goyenko R., 2013. Mutual fund $R^2$ as predictor of performance. Review of Financial Studies, 26, 667–694.

Ang A., Hodrick R. J., Xing Y., Zhang X., 2013. The cross-section of volatility and expected Returns. The Journal of Finance, 61, 259–299.

Armstrong C. S., Banerjee S., Corona, C., 2013. Factor-loading uncertainty and expected returns. The Review of Financial Studies, 26, 158–207.

Arnott, R. D., Beck, N., Kalesnik, V., West, J., 2016. How can 'smart beta' go horribly wrong? Working Paper.

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

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

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

Barber B. M., Huang X., Odean T., 2016. Which factors matter to investors? Evidence from mutual fund flows. The Review of Financial Studies, 29, 2600–2642.

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. Journal of Financial Economics, 116, 111–120.

Bazgour T., Bodson L., Sougné D., 2017. What style liquidity timing skills do mutual fund managers possess? Financial Review, 52, 597–626.

Benos, E., Jochec, M., Nyekel, V., 2010. Can mutual funds time risk factors? The Quarterly Review of Economics and Finance, 50, 509–514.

Bhattacharya, U., Loos, B., Meyer, S., & Hackethal, A., 2017. Abusing ETFs. Review of Finance, 21, 1217–1250.

Black, A., Fraser, P., Power, D., 1992. UK unit trust performance 1980–1989: a passive time-varying approach. Journal of Banking & Finance, 16, 1015–1033.

Blake D., Lehmann B., Timmermann A., 1999. Asset allocation dynamics and pension fund performance. Journal of Business, 72, 429–461.

Bodnaruk A., Chokaev B., & Simonov, A., 2014. Downside risk timing by mutual funds. Working Paper.

Bollen N. P., Busse J. A., 2001. On the timing ability of mutual fund managers. The Journal of Finance, 56, 1075–1094.

Brunnermeier K., Nagel S., 2004. Hedge Funds and the Technology Bubble. The Journal of Finance, 59, 2013–2040.

Busse J. A., 1999. Volatility timing in mutual funds: Evidence from daily returns. Review of Financial Studies, 12, 1009–1041.

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

Chabi-Yo F., Ruenzi S., Weigert F., 2018. Crash sensitivity and the cross-section of expected stock returns. Journal of Financial & Quantitative Analysis, 53, 1059–1100.

Chevalier J., Ellison G., 1999. Career concerns of mutual fund managers. Quarterly Journal of Economics, 114, 389–432.

Coval, J., Stafford E., 2007: Asset fire sales (and purchases) in equity markets. Journal of Financial Economics, 86, 479-512.

Cremers K. J. M., Petajisto A., 2009. How active is your fund manager? A new measure that predicts performance. Review of Financial Studies, 22, 3329–3365.

Daniel K., Grinblatt M., Titman S., Wermers R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. The Journal of Finance, 52, 1035–1058.

Dong X., Feng S., Sadka R., 2017. Liquidity Risk and Mutual Fund Performance. Management Science. forthcoming.

Elton E. J., Gruber J. M., Blake C. R., 2012. An examination of mutual fund timing ability using monthly holdings data. Review of Finance, 16, 619–645.

Engle R. F., 2016. Dynamic conditional beta. Journal of Financial Econometrics, 14, 643–667.

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

Fama E. F., French K. R., 2010. Luck versus skill in the cross-section of mutual fund returns. The Journal of Finance, 65, 1915–1947.

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

Ferson W E., Mo, H., 2016. Performance measurement with selectivity, market, and volatility timing. Journal of Financial Economics, 121, 93–110.

Ferson W. E., Schadt, R. W., 1996. Measuring fund strategy and performance in changing economic conditions. The Journal of Finance, 51, 425–461.

Frazzini A., Pedersen L. H., 2014. Betting against beta. Journal of Financial Economics, 111, 1–25.

Giambona, E., Golec, J., 2009. Mutual fund volatility and management fees. Journal of Banking and Finance, 33, 589–599.

Grinblatt M., Titman S., 1989. Mutual fund performance: An analysis of quarterly portfolio holdings. Journal of Business, 62, 393–416.

Grinblatt M., Titman S., 1993. Performance measurement without benchmarks: An examination of mutual fund returns. Journal of Business, 66, 47–68.

Gruber M. J., 1996. Another puzzle: The growth in actively managed mutual funds. The Journal of Finance, 51, 783–810.

Helske J., 2016. KFAS: Exponential Family State Space Models in R. Working Paper.

Henriksson R. D., Merton R. C., 1981. On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. Journal of Business, 54, 513–533.

Huang J., Sialm, C., Zhang, H., 2011. Risk shifting and mutual fund performance. Review of Financial Studies, 24, 2575–2616.

Jiang G. J., Yao T., Yu, T., 2007. Do mutual funds time the market? Evidence from portfolio holdings. Journal of Financial Economics, 86, 724–758.

Jostova G., Philipov A., 2005. Bayesian analysis of stochastic betas. Journal of Financial and Quantitative Analysis, 40, 747–778.

Kacperczyk M., Nieuwerburgh S. V., Veldkamp L., 2014. Time-varying fund manager skill. The Journal of Finance, 69, 1455–1484.

Kacperczyk M., Seru A., 2007. Fund manager use of public information: New evidence on managerial skills. The Journal of Finance, 62, 485–528.

Kalman R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering, 35–45.

Kelly B., Jiang H., 2014. Tail risk and asset prices. Review of Financial Studies, 27, 2841–2871.

Kim S., In F., 2012. False discoveries in volatility timing of mutual funds. Journal of Banking and Finance, 36, 2083–2094.

Lenz M., 2017. Factor loading persistence and expected returns. Working paper.

Malkiel B. G., 1995. Returns from investing in equity mutual funds 1971 to 1991. The Journal of Finance, 50, 77–91.

Mamaysky H., Spiegel M., Zhang H., 2007. Improved forecasting of mutual fund alphas and betas. Review of Finance, 11, 359–400.

Mamaysky H., Spiegel M., Zhang H., 2008. Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies, 21, 233–264.

Moreira A., Muir T., 2017. Volatility-managed portfolios. The Journal of Finance, 72, 1611–1644.

Pástor L., Stambaugh R. F., 2003. Liquidity risk and expected stock returns. Journal of Political economy, 111, 642–685.

Puetz A., Ruenzi S., 2011. Overconfidence among professional investors: Evidence from mutual fund managers. Journal of Business Finance & Accounting, 38, 684–712.

Rachev S. T., Mittnik S., Fabozzi S. M., 2015. Financial Econometrics: From basics to advanced modeling techniques. Wiley.

Racicot F. É., Théore, R., 2009. Modeling hedge fund returns using the Kalman filter: An errors-in-variables perspective. Working Paper.

Sensoy B.A., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, Journal of Financial Economics, Journal of Financial Economics, 92, 25-39.

Swinkels L., Tjong-A-Tjoe L., 2007. Can mutual funds time investment styles? Journal of Asset Management, 8, 123–132.

Swinkels L., Van Der Sluis P. J., 2006. Return-based style analysis with time-varying exposures. European Journal of Finance, 12, 529–552.

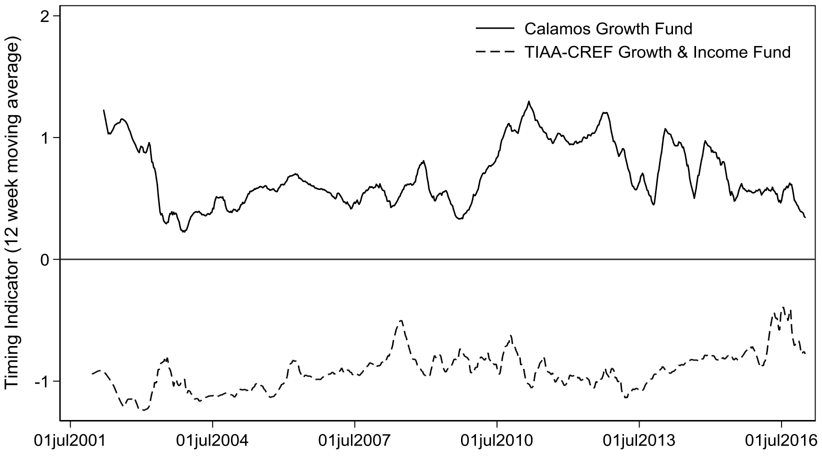
Treynor J., Mazuy, K., 1966. Can mutual funds outguess the market? Harvard Business Review 44, 131–136.

Wells C., 1994. Variable betas on the Stockholm exchange 1971-1989. Applied Financial Economics, 4, 75–92.

Wermers R., 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. The Journal of Finance, 55, 1655–95.

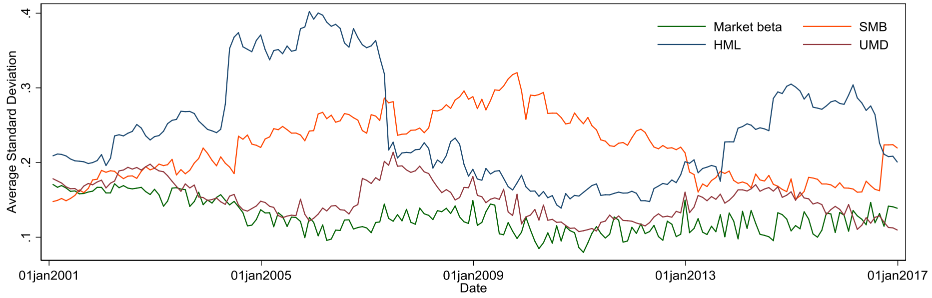
# Figures and Tables

## Figure 1: Timing Indicator of Calamos Growth Fund and TIAA-CREF Growth & Income Fund



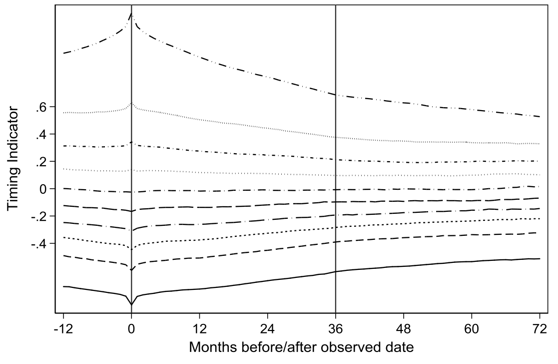

This figure plots the 2002-2016 time series of the Timing Indicator for the Calamos Growth Fund and the TIAA-CREF Growth & Income Fund. We calculate the Timing Indicator from the timing measures obtained from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is calculated from the past three years of weekly return data. A Timing Indicator >0 indicates an above average timing activity.

# Figure 2: Timing measures over time




This figure shows the evolution of cross-sectional mean timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ over time. We calculate a fund's timing measures as the standard deviation of its factor loadings during three years of weekly net returns. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process.

## Figure 3: Timing Indicator persistence




This figure shows the evolution of the mean Timing Indicator (see Section 2.2 for the calculation of the Timing Indicator) of decile portfolios over time. Each month funds are sorted into ten deciles by the current value of the Timing Indicator, which is calculated over the past three years of weekly net returns. The average values of the Timing Indicator of those deciles are displayed over time, starting 12 months prior to and ending 72 months after the formation period.

## Table 1: Descriptive statistics and timing activity by fund style

**Panel A: Fund characteristics**

| | # Obs. | Mean | 1% | 25% | 50% | 75% | 99% |
|---|---|---|---|---|---|---|---|
| Number of funds | 3,816 | | | | | | |
| Fund-Week-observations | 300,519 | | | | | | |
| Total assets (in mn. USD) | 300,519 | 1,329 | 19 | 107 | 324 | 1,049 | 21,268 |
| Fund age (years) | 300,361 | 15.73 | 2.57 | 7.65 | 12.68 | 19.04 | 72.42 |
| Manager tenure (years) | 241,442 | 7.51 | 0.42 | 3.66 | 6.33 | 10.09 | 25.76 |
| Turnover ratio | 265,402 | 0.75 | 0.03 | 0.30 | 0.58 | 0.99 | 3.42 |
| Total expense ratio (in %) | 266,142 | 1.15 | 0.14 | 0.92 | 1.14 | 1.37 | 2.23 |
| Relative fund flow | 300,311 | 0.02 | -0.59 | -0.15 | -0.06 | 0.08 | 1.90 |
| $\sigma(\beta_{RMRF})$ | 300,519 | 0.1220 | 0.0260 | 0.0828 | 0.1105 | 0.1495 | 0.3343 |
| $\sigma(\beta_{SMB})$ | 300,519 | 0.2179 | 0.0250 | 0.1242 | 0.1936 | 0.2847 | 0.6311 |
| $\sigma(\beta_{HML})$ | 300,519 | 0.2401 | 0.0200 | 0.1295 | 0.2035 | 0.3099 | 0.7831 |
| $\sigma(\beta_{UMD})$ | 300,519 | 0.1465 | 0.0138 | 0.0814 | 0.1278 | 0.1925 | 0.4273 |
| Timing Indicator | 301,908 | 0.0065 | -1.1529 | -0.4874 | -0.0937 | 0.3895 | 2.0284 |

**Panel B: Mean values of factor Timing Indicators by fund style**

| Fund Style | # Funds / # Obs. | $\sigma(\beta_{RMRF})$ | $\sigma(\beta_{SMB})$ | $\sigma(\beta_{HML})$ | $\sigma(\beta_{UMD})$ | Timing Indicator |
|---|---|---|---|---|---|---|
| Growth and Income | 773 / 57,877 | 0.105 | 0.173 | 0.194 | 0.118 | -0.343 |
| Growth | 1,636 / 127,455 | 0.120 | 0.212 | 0.235 | 0.144 | -0.029 |
| Hedged | 49 / 2,369 | 0.132 | 0.223 | 0.236 | 0.127 | 0.071 |
| Income | 202 / 14,016 | 0.108 | 0.175 | 0.209 | 0.130 | -0.247 |
| Mid Cap | 430 / 36,644 | 0.141 | 0.263 | 0.275 | 0.184 | 0.365 |
| Small Cap | 675 / 58,711 | 0.133 | 0.249 | 0.277 | 0.159 | 0.217 |
| Micro Cap | 45 / 4,275 | 0.155 | 0.315 | 0.321 | .0190 | 0.629 |
| Other | 6 / 172 | 0.105 | 0.173 | 0.194 | 0.118 | -0.177 |

Panel A of this table provides a descriptive overview over the sample size and fund characteristics. Size, age, management tenure, turnover ratio and total expense ratio are obtained from the CRSP survivorship bias free database and relative fund flows are calculated over the past year using $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1year,t)}))$. Fund styles are mainly determined by a fund's CRSP objective code. Funds are aggregated on a portfolio level and size is the sum of all share classes' total assets, fund age is the age of the oldest share class and all other characteristics as well as returns are calculated as the size-weighted mean of all share classes. The timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are the standard deviation of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. Panel B reports the average timing measures and Timing Indicators by fund style.

## Table 2: Cross-sectional correlations between timing measures and fund characteristics

**Panel A: Average cross sectional correlations between timing measures**

| | $\sigma(\beta_{RMRF})$ | $\sigma(\beta_{SMB})$ | $\sigma(\beta_{HML})$ | $\sigma(\beta_{UMD})$ | Timing Indicator |
|---|---|---|---|---|---|
| $\sigma(\beta_{RMRF})$ | 1.00 | | | | |
| $\sigma(\beta_{SMB})$ | 0.26 | 1.00 | | | |
| $\sigma(\beta_{HML})$ | 0.28 | 0.20 | 1.00 | | |
| $\sigma(\beta_{UMD})$ | 0.33 | 0.31 | 0.29 | 1.00 | |
| Timing Indicator | 0.69 | 0.65 | 0.65 | 0.71 | 1.00 |

**Panel B: Average cross sectional correlations between fund characteristics and the Timing Indicator**

| | Timing Indicator | Total exp. ratio | Turnover ratio | Relative fund flow | ln(total assets) | ln(fund age) | ln(tenure) |
|---|---|---|---|---|---|---|---|
| Timing Indicator | 1.00 | | | | | | |
| Total exp. ratio | 0.35 | 1.00 | | | | | |
| Turnover ratio | 0.22 | 0.22 | 1.00 | | | | |
| Relative fund flow | 0.00 | -0.05 | -0.05 | 1.00 | | | |
| ln(total assets) | -0.11 | -0.32 | -0.16 | 0.07 | 1.00 | | |
| ln(fund age) | -0.01 | -0.07 | -0.06 | -0.19 | 0.37 | 1.00 | |
| ln(tenure) | 0.04 | -0.03 | -0.15 | -0.02 | 0.05 | 0.18 | 1.00 |

Panel A of this table reports the average cross-sectional correlations between timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ and the Timing Indicator. The timing measures are the standard deviation of a fund's weekly factor exposures over the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Panel B reports the correlations between fund characteristics and the Timing Indicator. Fund size, age, management tenure, turnover ratio and total expense ratio are obtained from the CRSP survivorship bias free database and relative fund flows are calculated over the past year using $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1year,t)}))$. Funds are aggregated on a portfolio level and size is the sum of all share classes' total assets, fund age is the age of the oldest share class and all other characteristics as well as returns are calculated as the size-weighted mean of all share classes. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample.

## Table 3: Factor timing transition matrix

**Panel A: 1-year transition matrix and attrition rate**

| Current Decile | Mean initial / final Timing Indicator | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Attrition Rate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -0.86 / -0.73 | 60.47 | 22.31 | 8.82 | 4.15 | 2.14 | 0.96 | 0.58 | 0.29 | 0.18 | 0.09 | 6.47 |
| 2 | -0.60 /-0.51 | 21.02 | 32.09 | 22.07 | 12.27 | 6.40 | 3.20 | 1.53 | 0.87 | 0.47 | 0.08 | 6.92 |
| 3 | -0.45 / -0.37 | 8.62 | 21.01 | 25.10 | 19.59 | 12.35 | 6.72 | 3.62 | 1.99 | 0.75 | 0.24 | 6.65 |
| 4 | -0.31 / -0.26 | 4.52 | 11.79 | 18.65 | 21.98 | 18.61 | 11.68 | 7.10 | 3.59 | 1.62 | 0.46 | 7.21 |
| 5 | -0.16 / -0.14 | 2.27 | 6.21 | 11.63 | 17.55 | 21.31 | 17.72 | 12.18 | 6.98 | 3.23 | 0.92 | 6.70 |
| 6 | -0.02 / -0.18 | 1.45 | 3.49 | 6.78 | 11.47 | 17.31 | 20.66 | 18.45 | 12.63 | 6.05 | 1.73 | 7.33 |
| 7 | 0.14 / 0.12 | 0.86 | 2.05 | 3.94 | 6.85 | 11.77 | 17.40 | 22.59 | 19.55 | 11.07 | 3.92 | 6.73 |
| 8 | 0.34 / 0.27 | 0.72 | 1.19 | 2.11 | 3.73 | 7.05 | 11.90 | 18.95 | 24.77 | 21.05 | 8.54 | 7.51 |
| 9 | 0.63 / 0.50 | 0.41 | 0.72 | 1.17 | 2.01 | 3.42 | 6.76 | 11.55 | 20.24 | 31.38 | 22.34 | 8.34 |
| 10 | 1.28 / 0.95 | 0.30 | 0.39 | 0.52 | 0.67 | 1.36 | 2.09 | 4.33 | 8.91 | 23.32 | 58.12 | 14.24 |

**Panel B: 3-year transition matrix and attrition rate**

| Current Decile | Mean initial / final Timing Indicator | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Attrition Rate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -0.86 / -0.61 | 45.34 | 21.04 | 12.64 | 7.70 | 5.20 | 3.37 | 2.03 | 1.26 | 0.82 | 0.59 | 16.10 |
| 2 | -0.60 /-0.39 | 19.82 | 22.22 | 17.89 | 13.17 | 9.74 | 6.38 | 4.42 | 3.17 | 2.20 | 0.99 | 18.29 |
| 3 | -0.45 / -0.28 | 11.59 | 17.72 | 17.16 | 15.24 | 12.53 | 9.19 | 6.84 | 4.85 | 3.51 | 1.37 | 18.70 |
| 4 | -0.31 / -0.19 | 7.79 | 13.06 | 15.30 | 14.79 | 14.11 | 11.03 | 9.45 | 7.66 | 4.63 | 2.17 | 18.69 |
| 5 | -0.16 / -0.10 | 4.91 | 9.28 | 12.05 | 13.91 | 14.42 | 13.42 | 12.35 | 9.95 | 6.52 | 3.19 | 18.70 |
| 6 | -0.02 / -0.01 | 3.43 | 6.81 | 10.04 | 11.49 | 13.76 | 13.52 | 14.35 | 12.48 | 8.93 | 5.18 | 19.26 |
| 7 | 0.14 / 0.10 | 2.45 | 5.09 | 7.30 | 9.18 | 11.15 | 13.78 | 15.47 | 14.69 | 13.22 | 7.67 | 18.69 |
| 8 | 0.34 / 0.21 | 1.77 | 3.50 | 4.64 | 6.94 | 9.55 | 12.08 | 15.57 | 16.97 | 17.24 | 11.73 | 19.37 |
| 9 | 0.63 / 0.38 | 1.16 | 2.44 | 3.24 | 4.88 | 7.30 | 9.90 | 12.80 | 17.12 | 21.03 | 20.12 | 20.43 |
| 10 | 1.28 / 0.69 | 0.67 | 1.18 | 1.59 | 2.51 | 4.22 | 6.13 | 8.60 | 12.84 | 21.61 | 40.66 | 26.22 |

This table displays a transition matrix of mutual funds between deciles sorted on the Timing Indicator over a period of one year (Panel A) and three years (Panel B). A fund's Timing Indicator is defined as the mean of its cross-sectionally standardized timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$. Timing measures are the weekly standard deviation of a fund's factor loadings obtained from a dynamic version of Carhart's (1997) four-factor model over the previous three years. Each week we sort funds by their Timing Indicator. The first column reports the average Timing Indicator of funds within each decile upon its formation as well as one or three years later. The last column reports the percentage of funds within each decile that drop out of our sample within the next year or the next three years, respectively. For all other funds the table reports the transitions between the original decile and the decile funds would have been sorted into if the sorting was done one year or three years later.

**Table 4: Abnormal returns of quintile portfolios sorted by timing measures**

| | (1) $\sigma(\beta_{RMRF})$ | (2) $\sigma(\beta_{SMB})$ | (3) $\sigma(\beta_{HML})$ | (4) $\sigma(\beta_{UMD})$ | (5) Timing Indicator |
|---|---|---|---|---|---|
| Low Timing | -0.96%*** | -1.17%*** | -1.01%** | -0.91%*** | -0.80%** |
| | (-2.91) | (-4.07) | (-2.38) | (-2.82) | (-2.56) |
| (2) | -1.28%*** | -1.39%*** | -1.10%*** | -1.12%*** | -0.97%*** |
| | (-3.54) | (-4.12) | (-3.01) | (-3.17) | (-2.92) |
| (3) | -1.33%*** | -1.25%*** | -1.43%*** | -1.19%*** | -1.34%*** |
| | (-3.25) | (-3.24) | (-3.77) | (-3.07) | (-3.66) |
| (4) | -1.44%*** | -1.40%*** | -1.59%*** | -1.62%** | -1.58%*** |
| | (-3.12) | (-2.91) | (-4.06) | (-3.48) | (-3.17) |
| High Timing | -1.98%*** | -1.78%*** | -1.82%*** | -2.11%*** | -2.27%*** |
| | (-3.47) | (-2.88) | (-3.38) | (-3.56) | (-3.50) |
| High-Low Timing | -1.02%** | -0.61% | -0.82%** | -1.20%*** | -1.47%*** |
| | (-2.24) | (-1.33) | (-2.43) | (-2.68) | (-2.76) |

This table reports the abnormal returns of fund portfolios sorted on factor timing activity. Each month we sort funds into five quintiles by either a single factor timing measure $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ or by the Timing Indicator. The timing measures are the standard deviation of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. The sorting is done within each style category, where fund styles are mainly determined by a fund's CRSP objective code. We keep the portfolios constant for one month and calculate the equal weighted portfolio return from funds' net return. Each column represents the sorting by a distinct timing measure. We report reports Carhart (1997) alphas for each quintile portfolio (Rows 1-5) as well as the difference between the most and the least active portfolios (High-Low). We regress the return time series on a Carhart (1997) factor model and report the annualized alphas. T-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

**Table 5: Abnormal returns of quintile portfolios sorted by the Timing Indicator under different factor models**

| | (1) Carhart (1997) | (2) 1-Factor | (3) Fama/French 3 Factors (1993) | (4) Fama/French 5 Factors (2015) | (5) FF3 + Rever-sal | (6) Carhart + BaB | (7) Carhart + Sentiment | (8) Pástor-Stambaugh |
|---|---|---|---|---|---|---|---|---|
| Low Timing | -0.80%** (-2.56) | -0.14% (-0.28) | -0.76%** (-2.37) | -1.28%*** (-4.40) | -0.81%** (-2.55) | -1.20%*** (-4.01) | -0.62%* (-1.89) | -0.94%*** (-3.01) |
| (2) | -0.97%*** (-2.92) | -0.37%* (-0.74) | -0.94%*** (-2.76) | -1.31%*** (-4.06) | -0.98%*** (-2.91) | -1.34%*** (-4.10) | -0.84%** (-2.36) | -1.20%*** (-3.71) |
| (3) | -1.34%*** (-3.66) | -0.76% (-1.43) | -1.30%*** (-3.46) | -1.45%*** (-3.92) | -1.34%*** (-3.67) | -1.70%*** (-4.68) | -1.21%*** (-3.12) | -1.57%*** (-4.44) |
| (4) | -1.58%*** (-3.17) | -0.96% (-1.42) | -1.52%** (-2.93) | -1.71%** (-3.26) | -1.53%*** (-2.97) | -1.98%*** (-3.94) | -1.41%*** (-2.66) | -1.83%*** (-3.70) |
| High Timing | -2.27%*** (-3.50) | -1.61%* (-1.83) | -2.20%*** (-3.27) | -2.14%*** (-3.15) | -2.20%*** (-3.34) | -2.82%*** (-4.32) | -2.12%*** (-3.09) | -2.50%*** (-3.75) |
| High-Low Timing | -1.47%*** (-2.76) | -1.47%** (-2.08) | -1.43%*** (-2.65) | -0.86% (-1.60) | -1.39%*** (-2.67) | -1.61%*** (-2.93) | -1.50%*** (-2.78) | -1.56%*** (-2.77) |

This table reports the abnormal returns of fund portfolios sorted by the Timing Indicator. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures, which are defined as the standard deviation of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in section 2.2. We assume risk factor exposures to follow a mean-reverting process. The sorting is done within each style category, where fund styles are mainly determined by a fund's CRSP objective code. We keep the portfolios constant for one month and calculate the equal weighted portfolio return from funds' net return. We regress each quintile portfolio's return time series on different factor models. Each column refers to one factor model, namely the one-factor model including only the market factor, the Fama/French (1993) three-factor model, the Carhart (1997) four-factor model, the Fama/French (2015) five factor model, a Fama/French (1993) three-factor model extended by a short and long term reversal factor as well as a Carhart (1997) model extended by the Frazzini/Pedersen (2014) betting against beta factor, the Baker/Wurgler (2006) sentiment factor or the Pástor/Stambaugh (2003) liquidity factor. We report the annualized alphas for each quintile portfolio (Rows 1-5) as well as the difference between the most and the least active portfoli-os (High-Low). T-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 6: Fama-MacBeth regressions

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Explanatory variables | annualized $alpha_{j,t}$ | annualized $alpha_{j,t}$ | 6-months CAR | 12-months CAR |
| $\sigma(\beta_{RMRF})$ | -0.060*** (-3.33) | | | |
| $\sigma(\beta_{SMB})$ | -0.032** (-2.15) | | | |
| $\sigma(\beta_{HML})$ | -0.011 (-1.15) | | | |
| $\sigma(\beta_{UMD})$ | -0.025* (-1.78) | | | |
| Timing Indicator | | -1.036*** (-2.78) | -0.763*** (-3.99) | -0.616*** (-5.21) |
| ln(tna) | -0.158** (-2.02) | -0.160** (-2.03) | -0.157*** (-3.50) | -0.135*** (-4.96) |
| ln(fund age) | 0.144 (1.03) | 0.163 (1.18) | 0.150** (2.46) | 0.090* (1.77) |
| ln(manager tenure) | 0.013 (0.17) | 0.006 (0.08) | -0.024 (-0.50) | -0.017 (-0.56) |
| Expenses | -0.777*** (-5.72) | -0.783*** (-5.58) | -0.743*** (-12.63) | -0.731*** (-18.50) |
| Turnover | -0.207 (-0.96) | -0.211 (-0.96) | -0.349** (-2.56) | -0.348*** (-3.16) |
| Lagged Alpha | 0.262*** (5.86) | 0.259*** (5.70) | 0.189*** (8.11) | 0.160*** (8.70) |
| Fund Flows | 0.181 (0.68) | 0.201 (0.75) | 0.105 (0.84) | -0.101 (-1.12) |
| Style Dummies | YES | YES | YES | YES |
| Average $R^2$ | 0.14 | 0.13 | 0.14 | 0.15 |

This table reports the results of Fama-MacBeth regressions of annualized abnormal fund returns on timing measures and controls. Each month, expected returns are calculated from a Carhart (1997) model where the factor loadings are estimated over the past three years of weekly return data from an OLS regression. Abnormal returns are the differences between actual monthly returns and the expected returns. Fama-MacBeth regressions are applied on the panel data of monthly abnormal returns. The first two columns report results where the dependent variable is the next month's abnormal return, the last two columns report results where the cumulated abnormal return over the next six or 12 months is regressed on fund characteristics. Funds are aggregated on a portfolio level and fund characteristics are calculated as described in Section 2.1. The Timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are the standard deviations of a fund's weekly factor exposures over the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. T-statistics are reported in parentheses. We use Newey-West standard errors (lag = 1 month) for the regressions with abnormal returns as the dependent variable. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 7: Fama-MacBeth-regressions within sub-periods

| | (1) NBER Expansion | (2) NBER Recession | (3) Positive MRP | (4) Negative MRP | (5) Without crises (11/2007 – 02/2009) |
|---|---|---|---|---|---|
| Timing Indicator | -0.736** | -2.953 | -0.644 | -1.635** | -0.880** |
| | (-2.33) | (-1.62) | (-1.56) | (-2.52) | (-2.55) |
| ln(tna) | -0.094* | -0.587 | 0.085 | -0.535*** | -0.136* |
| | (-1.70) | (-1.21) | (1.07) | (-3.70) | (-1.84) |
| ln(fund age) | 0.014 | 1.115 | 0.114 | 0.237 | 0.074 |
| | (0.13) | (1.50) | (0.78) | (1.06) | (0.57) |
| ln(manager tenure) | 0.025 | -0.112 | 0.120 | -0.167 | -0.003 |
| | (0.35) | (-0.36) | (1.30) | (-1.16) | (-0.04) |
| Expenses | -0.729*** | -1.128** | -0.605*** | -1.055*** | -0.793*** |
| | (-5.10) | (-2.19) | (-3.16) | (-4.23) | (-5.43) |
| Turnover | -0.197 | -0.305 | 0.243 | -0.904** | -0.185 |
| | (-1.00) | (-0.29) | (1.11) | (-2.24) | (-0.87) |
| Lagged Alpha | 0.280*** | 0.126 | 0.302*** | 0.194** | 0.265*** |
| | (5.75) | (1.00) | (5.14) | (2.58) | (5.64) |
| Fund Flows | 0.026 | 1.321 | 0.654* | -0.491 | 0.283 |
| | (0.10) | (1.09) | (1.85) | (-1.24) | (1.00) |
| Style Dummies | YES | YES | YES | YES | YES |
| | 166 | 26 | 116 | 76 | 176 |
| Average $R^2$ | 0.12 | 0.16 | 0.13 | 0.13 | 0.13 |

This table reports the results of Fama-MacBeth regressions of annualized one-month abnormal fund returns on the Timing Indicator and controls for several subperiods. Each month, expected returns are calculated from a Carhart (1997) model where the factor loadings are estimated over the past three years of weekly return data from an OLS regression. Abnormal returns are the differences between actual monthly returns and the expected returns. Fama-MacBeth regressions are applied on the panel data of monthly abnormal returns. Funds are aggregated on a portfolio level and fund characteristics are calculated as described in Section 2.1. The Timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are the standard deviations of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. The regressions are conducted for several sub-periods, that is, expansion and recession months as defined by the NBER, months with a positive and negative market risk premium, as well as during all months besides the 11/07 - 02/09 financial crisis. T-statistics are reported in parentheses. We use Newey-West standard errors (lag = 1 month) for the regressions with abnormal returns as the dependent variable. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

**Table 8: Portfolio Double Sorts, Activeness and Factor Timing**

| | Low Timing | (2) | (3) | (4) | High Timing | High-Minus-Low |
|---|---|---|---|---|---|---|
| | | | Panel A: Sorting on $R^2$ and Timing Indicator | | | |
| Low $R^2$ | 0.14%<br>(0.22) | -0.23%<br>(-0.32) | -0.69%<br>(-0.82) | -0.78%<br>(-0.93) | -1.30%<br>(-1.21) | -1.44%*<br>(-1.68) |
| (2) | -0.49%<br>(-0.83) | -0.47%<br>(-0.79) | -1.29%**<br>(-2.02) | -0.64%<br>(-1.01) | -2.52%***<br>(-3.78) | -2.04%***<br>(-2.60) |
| (3) | -0.71%<br>(-1.56) | -1.07%**<br>(-2.29) | -1.15%**<br>(-2.34) | -2.16%***<br>(-4.29) | -2.40%***<br>(-3.90) | -1.69%***<br>(-2.64) |
| (4) | -0.71%*<br>(-1.77) | -1.64%***<br>(-4.24) | -1.76%***<br>(-4.30) | -2.07%***<br>(-4.49) | -2.63%***<br>(-5.39) | -1.92%***<br>(-3.56) |
| High $R^2$ | -0.71%***<br>(-2.93) | -1.52%***<br>(-5.21) | -2.10%***<br>(-5.69) | -1.84%***<br>(-4.63) | -2.55%***<br>(-5.60) | -1.84%***<br>(-4.08) |
| | | | | | Average Coefficient | -1.79%*** |
| | | | Panel B: Sorting on Risk Shifting and Timing Indicator | | | |
| Low Risk Shifting | -1.52%***<br>(-4.52) | -1.56%***<br>(-3.71) | -1.65%***<br>(-3.43) | -2.54%***<br>(-4.43) | -3.31%***<br>(-4.68) | -1.79%***<br>(-3.01) |
| (2) | -1.52%***<br>(-5.31) | -1.35%***<br>(-3.99) | -1.08%***<br>(-2.90) | -1.23%***<br>(-2.64) | -1.98%***<br>(3.33) | -0.46%<br>(-0.82) |
| (3) | -0.79%***<br>(-2.99) | -0.97%***<br>(-3.05) | -1.07%***<br>(-2.90) | -1.21%***<br>(-2.62) | -1.79%***<br>(-3.10) | -1.00%*<br>(-1.87) |
| (4) | -0.80%***<br>(-2.71) | -0.79%**<br>(-2.44) | -1.09%***<br>(-2.75) | -0.49%<br>(-1.00) | -1.02%*<br>(-1.67) | -0.21%<br>(-0.42) |
| High Risk Shifting | -0.83%**<br>(-2.48) | -0.42%<br>(-0.95) | -1.23%**<br>(-2.55) | -1.25%**<br>(-2.20) | -1.34%*<br>(-1.68) | -0.50%<br>(-0.80) |
| | | | | | Average Coefficient | -0.79% |

This table reports the results of bivariate portfolio sorts. Funds are first sorted into five quintiles by either the Amihud/Goyenko (2013, Panel A) $R^2$ measure or the fund's Huang et al. (2011, Panel B) risk shifting measure. Within each quintile funds are sorted into five quintiles by their Timing Indicator. $R^2$ is calculated from a Carhart (1997) model over three years of return data and we follow Huang et al. (2011) to calculate the risk shifting measure. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures, which are defined as the standard deviation of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The sorting is done within each style category, where fund styles are mainly determined by a fund's CRSP objective code. We keep the portfolios constant for one month and calculate the equal weighted portfolio returns. We regress each quintile portfolio's return time series on the Carhart (1997) four-factor model. Within each panel, we report the annualized alphas for each 5x5 portfolio as well as the difference between the most and the least active portfolios (High-Low) within each of 5 quintile portfolios from the first sorting step. T-statistics are reported in parentheses. We also report average coefficients and t-statistics from the High-Low returns. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 9: Robustness checks

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | Size-weighted | Alternative Performance Measures | | | | Other Models | | Additional Control Var. | Index Funds |
| Explanatory variables | annualized alphaj,t | skill | Sharpe ratio (26 weeks) | MPPM (rho=2) | MPPM (rho=3) | Random Walk | Idiosynchratic Beta-Volatility | annualized $alpha_{j,t}$ | annualized $alpha_{j,t}$ |
| Timing Indicator | -1.029** | -977.482** | -0.031*** | -0.828** | -1.103*** | -0.974** | -0.842** | -0.670* | -0.045 |
| | (-2.38) | (-2.04) | (-2.69) | (-2.40) | (-3.02) | (-2.43) | (-2.54) | (-1.87) | (-0.12) |
| ln(tna) | -0.132 | -13.973** | -0.004 | -0.230*** | -0.253*** | -0.151* | -0.165 | -0.119* | -0.040 |
| | (-1.08) | (-2.19) | (-1.42) | (-3.34) | (-3.56) | (-1.68) | (-2.06) | (-1.74) | (-0.88) |
| ln(fund age) | 0.081 | 0.634 | 0.009 | 0.309*** | 0.336*** | 0.086 | 0.079 | 0.119 | 0.083 |
| | (0.49) | (0.35) | (1.55) | (3.42) | (3.65) | (0.72) | (0.59) | (1.10) | (0.60) |
| ln(manager tenure) | -0.053 | -1.017 | -0.010*** | -0.068 | -0.063 | 0.033 | 0.022 | -0.032 | 0.120 |
| | (-0.33) | (-0.26) | (-3.14) | (-0.89) | (-0.81) | (0.43) | (0.29) | (-0.46) | (1.32) |
| Expenses | -1.109*** | 190.744 | -4.497*** | -0.785*** | -0.798*** | -0.761*** | -0.753*** | -1.015*** | -1.024*** |
| | (-3.58) | (0.58) | (-7.15) | (-9.83) | (-9.63) | (-7.31) | (-5.46) | (-6.45) | (-3.79) |
| Turnover | -0.351 | -134.092 | -0.791 | -0.090 | -0.136 | 0.208 | -0.216 | -0.100 | -0.110 |
| | (-1.30) | (-0.51) | (-1.16) | (-0.49) | (-0.74) | (-0.80) | (-0.96) | (-0.46) | (-0.65) |
| Lagged Alpha | 0.248*** | 233.256*** | 1.689*** | 0.213*** | 0.231*** | 0.246*** | 0.261*** | 0.249*** | 0.283*** |
| | (4.36) | (3.42) | (10.78) | (7.01) | (7.46) | (5.51) | (5.75) | (5.21) | (3.47) |
| Fund Flows | 0.024 | 4.698 | -0.011 | -0.247 | -0.221 | 0.210 | 0.198 | 0.005 | -0.178 |
| | (0.07) | (1.32) | (-1.28) | (-1.65) | (-1.46) | (0.95) | (0.75) | (0.03) | (-0.74) |
| $R^2$ | | | | | | | | -0.027 | |
| | | | | | | | | (-0.93) | |
| Risk Shifting | | | | | | | | 0.402** | |
| | | | | | | | | (2.36) | |
| Style Dummies | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Average $R^2$ | 0.20 | 0.08 | 0.21 | 0.27 | 0.27 | 0.13 | 0.13 | 0.14 | 0.42 |

This table reports the results of Fama-MacBeth regressions of alternative performance measures on the Timing Indicator and control variables as well as different robustness checks. Column (1) reports results of Fama-MacBeth regressions with the first-step regressions being a size-weighted regression. In columns (2) – (5) the dependent variable consists of an alternative performance measure: the skill measure of Berk/van Binsbergen (2015), the sharpe ratio computed over the past 26 weeks, and the 26-week manipulation-proof performance measure of Goetzmann et al. (2007) with ρ=2 and ρ=3. Columns (6) and (7) report results for alternative definitions of factor timing. For column (6) we assume risk factor exposures to follow a random walk instead of a mean-reverting process, and for column (7) we measure the standard deviation of the idiosyncratic component of factor loading dynamics. In Column (8), the Amihud/Goyenko (2013) $R^2$ measure and the Huang et al. (2011) risk shifting measure are added as control variables. Column (9) displays the results for a sample of index funds. Expected returns are calculated from a OLS regression of a Carhart (1997) model. Abnormal returns are the differences between actual monthly returns and the expected returns. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. T-statistics are reported in parentheses. We use Newey-West standard errors (lag=1 month) for the regressions with abnormal returns as the dependent variable. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 10: Equity Portfolio Holdings

| Explanatory variables | (1) | (2) | (3) |
|---|---|---|---|
| Timing Indictor (based on returns) | -1.036*** (-2.78) | | -0.640* (-1.90) |
| Timing Indicator (based on holdings) | | -0.725* (-2.11) | -0.186 (-0.99) |
| ln(tna) | -0.160** (-2.03) | -0.075 (-0.97) | -0.073 (-0.97) |
| ln(fund age) | 0.163 (1.18) | 0.110 (0.92) | 0.103 (0.85) |
| ln(manager tenure) | 0.006 (0.08) | -0.018 (-0.28) | -0.016 (-0.25) |
| Expenses | -0.783*** (-5.58) | -0.900*** (-4.25) | -0.869*** (-4.22) |
| Turnover | -0.211 (-0.96) | -0.180 (-0.66) | -0.166 (-0.63) |
| Lagged Alpha | 0.259*** (5.70) | 0.271*** (5.07) | 0.272*** (5.18) |
| Fund Flows | 0.201 (0.75) | 0.093 (0.41) | 0.090 (0.40) |
| Style Dummies | YES | YES | YES |
| Average $R^2$ | 0.13 | 0.12 | 0.13 |

This table reports the results of Fama-MacBeth regressions of annualized one-month abnormal fund returns on timing measures and controls. Expected returns are calculated from a OLS regression of a Carhart (1997) model. Abnormal returns are the differences between actual monthly returns and the expected returns. Fama-MacBeth regressions are applied on the panel data of monthly abnormal returns. Funds are aggregated on a portfolio level and fund characteristics are calculated as describes in section 2.1. The Timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are the standard deviations of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Besides the Timing Indicator calculated from funds' net returns we calculate a second Timing Indicator from funds' equity portfolio holdings. Portfolio holding are reported on a quarterly basis and we assume that between those reporting dates a fund held constant portfolio weights. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ for either the net return based or the holding based approach are amongst the highest 1% are dropped from the sample. T-statistics are reported in parentheses. We use Newey-West standard errors (lag = 1 month) for the regressions with abnormal returns. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 11: Fama-MacBeth Regression by Flow

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | 30% lowes past flows | Medium flows | 30% highest past flows | 30% lowest past absolute flow | Medium absolute flows | 30% highest past absolute flow | 15% highest past absolute flow | 5% highest past absolute flow |
| Timing Indicator | -0.668* | -1.117** | -0.967*** | -0.984** | -1.036** | -0.943*** | -1.099*** | -1.276 |
| | (-1.89) | (-2.55) | (-2.62) | (-2.32) | (-2.44) | (-2.69) | (-2.78) | (-1.28) |
| ln(tna) | -0.069 | -0.134* | -0.191* | -0.140 | -0.181* | -0.218** | -0.188** | -0.419 |
| | (-0.72) | (-1.79) | (-1.77) | (-1.50) | (-1.91) | (-1.98) | (-2.12) | (-0.72) |
| ln(fund age) | -0.266 | 0.139 | -0.077 | 0.028 | 0.206 | 0.034 | 0.248* | 0.207 |
| | (-1.22) | (1.00) | (-0.41) | (0.14) | (1.05) | (0.17) | (1.68) | (0.09) |
| ln(manager tenure) | 0.106 | -0.136 | -0.135 | 0.159** | -0.099 | -0.169 | -0.063 | 2.232 |
| | (1.01) | (-0.89) | (-0.61) | (1.97) | (-0.74) | (-0.77) | (-0.65) | (1.19) |
| Expenses | -0.929** | -0.960*** | -0.524** | -0.980*** | -0.642*** | -0.969*** | -0.741*** | -1.602 |
| | (-2.40) | (-4.35) | (-2.03) | (-2.99) | (-2.83) | (-3.47) | (-4.91) | (-1.42) |
| Turnover | -0.529 | -0.255 | -0.115 | -0.512 | -0.253 | -0.052 | -0.154 | -0.453 |
| | (-1.43) | (-1.00) | (-0.52) | (-1.44) | (-0.97) | (-0.24) | (-0.70) | (-0.59) |
| Lagged Alpha | 0.250*** | 0.312*** | 0.239*** | 0.317*** | 0.244*** | 0.248*** | 0.256*** | 0.472*** |
| | (4.72) | (5.78) | (3.93) | (5.14) | (4.73) | (4.59) | (5.59) | (4.78) |
| Fund Flows | 1.809 | 0.405 | 0.356 | -0.351 | 0.512 | 0.223 | 0.207 | 0.370 |
| | (1.82) | (0.70) | (1.43) | (-0.35) | (0.96) | (0.85) | (0.77) | (0.14) |
| Style Dummies | YES | YES | YES | YES | YES | YES | YES | YES |
| Average $R^2$ | 0.16 | 0.17 | 0.17 | 0.17 | 0.16 | 0.16 | 0.13 | 0.39 |

This table reports the results of Fama-MacBeth regressions of annualized abnormal fund returns on the Timing Indicator measures and controls within fund subsamples. Funds are sorted into subsamples by either the past 3-year flow (columns 1-3) or the past 3-year absolute flow (columns 4-8). Yearly flows are calculated as $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1year,t)}))$. The 3-year flow is the sum of the year flows during the most recent three years. The 3-year absolute flow is calculated as the sum of the absolute values of yearly flows during the previous three years. Within each subsample we apply the Fama-MacBeth regression as follows. Each month, expected returns are calculated from a Carhart (1997) model where the factor loadings are estimated over the past three years of weekly return data from an OLS regression. Abnormal returns are the differences between actual monthly returns and the expected returns. We regress the abnormal returns on the Timing Indicator and further control variables. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. T-statistics are reported in parentheses. We use Newey-West standard errors (lag = 1 month) for the regressions with abnormal returns as the dependent variable. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

## Table 12: Determinants of factor timing activity

| | (1) $\sigma(\beta_{RMRF})$ | (2) $\sigma(\beta_{SMB})$ | (3) $\sigma(\beta_{HML})$ | (4) $\sigma(\beta_{UMD})$ | (5) Timing Indicator |
|---|---|---|---|---|---|
| ln(tna) | -2.89e-4<br>(-0.56) | -1.47e-4<br>(-0.15) | -3.02e-3<br>(-1.27) | 2.12e-3<br>(1.38) | 1.06e-3<br>(0.10) |
| ln(fund age) | 2.18e-3<br>(1.47) | 6.27e-4<br>(0.18) | 6.67e-3*<br>(1.86) | 2.14e-3<br>(1.11) | 3.43e-2<br>(1.55) |
| ln(manager tenure) | 2.55e-3**<br>(2.45) | 8.77e-3<br>(4.57) | 7.28e-3***<br>(2.86) | 6.44e-3***<br>(7.46) | 6.37e-2***<br>(6.51) |
| Expenses (in %) | 2.78e-2***<br>(6.78) | 6.61e-2***<br>(18.35) | 5.99e-2***<br>(8.25) | 3.99e-2***<br>(8.46) | 5.14e+1***<br>(9.42) |
| Turnover ratio | 5.19e-3***<br>(3.09) | 1.39e-2***<br>(5.82) | 1.23e-2***<br>(3.05) | 1.56e-2***<br>(11.71) | 1.21e-1***<br>(8.01) |
| Past Alpha | 7.11e-3<br>(0.17) | -3.31e-2<br>(-0.69) | 1.10e-2<br>(0.40) | 4.93e-3<br>(0.18) | 5.71e-2<br>(0.18) |
| Fund Flows | 1.98e-6***<br>(3.26) | 2.41e-3<br>(1.19) | 1.50e-2**<br>(2.52) | 2.53e-3**<br>(2.13) | 4.75e-2***<br>(4.86) |
| *Style dummy variables* | | | | | |
| *Growth and Income* | – | – | – | – | – |
| *Growth* | 0.013***<br>(4.75) | 0.030***<br>(4.64) | 0.033***<br>(6.33) | 0.017***<br>(4.24) | 0.239***<br>(9.05) |
| *Hedged* | 0.032**<br>(2.49) | 0.052**<br>(2.24) | 0.061*<br>(1.71) | 0.004<br>(0.44) | 0.388***<br>(3.48) |
| *Income* | 0.003*<br>(1.85) | -0.007<br>(-0.51) | 0.007<br>(-0.83) | 0.002<br>(0.32) | 0.012<br>(0.21) |
| *Micro* | 0.036***<br>(6.35) | 0.080***<br>(4.36) | 0.079***<br>(6.24) | 0.063***<br>(5.02) | 0.701***<br>(7.87) |
| *Mid* | 0.029***<br>(3.25) | 0.066***<br>(8.31) | 0.065***<br>(3.53) | 0.053***<br>(6.56) | 0.560***<br>(6.85) |
| *Small* | 0.021***<br>(2.97) | 0.057***<br>(7.55) | 0.062***<br>(2.73) | 0.034***<br>(5.11) | 0.432***<br>(6.53) |
| $R^2$ | 0.24 | 0.17 | 0.23 | 0.26 | 0.24 |

This table reports the results of multivariate regressions of future factor timing activity on fund characteristics. We split our sample into non-overlapping 3-year subperiods, that is 1999-2001, 2002-2004, etc. up to 2014-2016. We regress the timing measures estimated from the dynamic factor model during those periods on the fund characteristics measured at the beginning of these periods. Fund size, age, management tenure, turnover ratio and total expense ratio are obtained from the CRSP survivorship bias free database and relative fund flows are calculated over the past year using $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1year,t)}))$. Fund styles are mainly determined by a fund's CRSP objective code. Funds are aggregated on a portfolio level and size is the sum of all share classes' total assets, fund age is the age of the oldest share class and all other characteristics as well as returns are calculated as the size-weighted mean of all share classes. The Timing measures $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are the standard deviations of a fund's weekly factor exposures during the past three years. The factor loadings are estimated from a dynamic version of Carhart's (1997) four-factor model as introduced in Section 2.2. We assume risk factor exposures to follow a mean-reverting process. The Timing Indicator is the mean of the four cross-sectionally standardized timing measures. Data on age, tenure, turnover, expense ratio, flows, and total net assets are winsorized at the 1%-level and observations for which the estimated values of $\sigma(\beta_{RMRF})$, $\sigma(\beta_{SMB})$, $\sigma(\beta_{HML})$ and $\sigma(\beta_{UMD})$ are amongst the highest 1% are dropped from the sample. Standard errors are double clustered on fund level and time period. T-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.