Do mutual fund managers have risk factor timing skills?

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

December 2017

Abstract

We investigate whether mutual fund managers exhibit timing skills with regard to different asset pricing risk factors. Using a dynamic state space version of Carhart's (1997) four-factor model, we measure fund managers' timing activity by the variation of different risk factor loadings. Our results indicate no evidence of positive timing skill; instead, timing of market, size, or momentum risk factors is associated with future underperformance and undesirable risk characteristics. A one standard deviation increase of average market, size, or momentum timing leads to a decrease in average fund performance by 30, 39 or 24 basis points p.a., respectively. Moreover, abnormal return volatility of the most actively timing funds exceeds the abnormal return volatility of the least actively timing funds by 129%.

JEL Classification: G11, G14, G20, G23

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

We thank Yakov Amihud, the participants of the Finance Seminar at the University of Constance and the Finance Seminar at the University of St. Gallen for their helpful comments.

[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, Fax: +41 71 224 70 88

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

Do mutual fund managers have risk factor timing skills?

December 2017

Abstract

We investigate whether mutual fund managers exhibit timing skills with regard to different asset pricing risk factors. Using a dynamic state space version of Carhart's (1997) four-factor model, we measure fund managers' timing activity by the variation of different risk factor loadings. Our results indicate no evidence of positive timing skill; instead, timing of market, size, or momentum risk factors is associated with future underperformance and undesirable risk characteristics. A one standard deviation increase of average market, size, or momentum timing leads to a decrease in average fund performance by 30, 39 or 24 basis points p.a., respectively. Moreover, abnormal return volatility of the most actively timing funds exceeds the abnormal return volatility of the least actively timing funds by 129%.

JEL Classification: G11, G14, G20, G23

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

1 Introduction

Among academics, there is a widely accepted consensus that mutual funds on average fail to beat a risk-adjusted benchmark net of fees.[1] Therefore, the focus of fund literature has moved to the question which investment characteristics lead to future abnormal returns and whether there are indicators that ex ante identify top-performers. A benchmark-adjusted outperformance could be reached by either exposing the fund to additional risk factors (and therefore realizing alternative risk premiums), by selecting single stocks that will outperform (i.e., by stock selection) or by varying the systematic risk exposures and by that successfully timing risk factors (i.e., by risk factor timing). Our paper is concerned with the latter and focuses on the risk factor timing ability of mutual funds.

To measure risk factor timing activity of mutual funds we propose to apply a Carhart (1997) risk factor model with dynamic factor exposures that follow a mean-reverting process. To estimate this model we use a Kalman filter and Kalman smoother technique over a period of 156 weeks in a rolling manner. We then measure the factor timing activity by the degree of volatility within the mean-reverting process of factor loadings with regard to the market factor, the size (SMB) factor, the book-to-market (HML) factor, and the momentum (UMD) factor. To express a fund's overall level of factor timing compared to other funds we aggregate market, size, value and momentum timing into a single Timing Indicator.

The example of the Fidelity Magellan Fund, one of the best known US mutual funds, helps to better understand the idea of risk factor timing and the relevance of our timing measure. The fund is best known for its success and rapid growth under Peter Lynch's management during the 1980s

[1] Among others, Malkiel (1995), Gruber (1996), Carhart (1997), and Fama and French (2010) support this view. Studies of Grinblatt and Titman (1989, 1993) and Daniel et al. (1997) show that gross of fees mutual funds generate positive abnormal returns. Wermers (2000) combines both views as he finds mutual funds to exhibit positive stock picking ability but negative abnormal net returns.

and early 1990s. From 1996 to 2005 the fund was managed by Robert Stansky who was closely following the S&P500 index and was accused of ‘closet indexing’. Magellan’s strategy changed quickly when Harry Lange took office in 2005. He was known for buying blue chips as well as small cap firms – that’s why Kiplinger’s Personal Finance magazine called his management style a ‘go everywhere approach’. Comparing Stansky’s and Lange’s investment strategies characterizes different approaches to risk factor timing. Figure 1 plots the fund’s Timing Indicator during Stansky’s and Lange’s years in office where positive values indicate an above-average timing activity. As shown in the figure, our timing measures reflect the different risk factor timing activity within the strategies of Magellan’s fund managers.[2] While Robert Stansky is managing the fund, its overall timing activity as measured with our approach is below average and much lower than during Harry Lange’s management period. During Lange’s tenure the fund’s timing activity reaches an above-average-level. Finally, in fall 2011, when Jeffrey Feingold takes control of the fund, the timing activity goes down and again reaches a below-average-level, yet it is higher than under Stansky’s management.

We subsequently measure the risk factor timing activity of a large US mutual fund sample. Our results reveal the following main results: Firstly, we show that the absolute level of factor timing with regard to the SMB, HML, and UMD factor is substantial and much higher than the absolute level of factor timing with regard to the market. Secondly, risk factor timing is associated with future fund performance; a one standard deviation increase of market timing activity leads on average to a decrease of annualized abnormal returns by 30 basis points p.a. Increasing either size or momentum timing by one standard deviation leads to an abnormal return decrease of 39 or 24 basis points p.a. Finally, risk factor timing increases return volatility of funds. The standard

[2] Having a high timing activity does not necessarily mean that fund managers intentionally time risk factor exposures. Buying large cap stocks at one point in time and shifting to small cap stocks another time implied size timing even if this shift is not done in order to time the market but because some specific small cap stocks might suddenly seem more appealing.

deviation of abnormal returns of the most actively timing funds exceeds the abnormal return volatility of the least actively timing funds by 129%.

The question whether mutual funds can successfully time risk exposures is nothing new but has been widely studied with respect to market timing only. Treynor and Mazuy (1966, TM) have been the first to adapt a non-linear market beta model to explore market timing abilities. 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. There are few exceptions that provide evidence for market timing, e.g. Mamaysky et al. (2008) and Bollen and Busse (2001). Kacperczyk et al. (2014) find positive market timing ability during recessions. Elton et al. (2012) find positive timing skill when using a one factor model but when applying a more complex factor model the evidence disappears. The literature on factor timing beyond the market factor is rather scarce. Daniel et al. (1997) find that mutual fund managers do not possess timing abilities with respect to investment styles and Benos et al. (2010) do not find factor timing abilities either. A study by Swinkels and Tjong-a-Tjoe (2007) detects positive timing skills within a very small US fund sample. A closely related study by Huang et al. (2011) finds that funds which intensively shift their total risk exposure over time underperform funds with a stable risk level.

Our research contributes to the mutual fund literature on market and risk factor timing. Firstly, 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. While 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 fund’s timing activity. This also enables us to observe a high persistence in factor timing activity as an investment characteristic.

Secondly, 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. Our results of a negative return effect of timing activity goes even beyond the most prominent findings of no positive market timing skill. We also contribute to the literature on fund management activeness as timing is one element of activeness and is closely linked to already developed activeness measures such as the Amihud/Goyenko (2013) selectivity measure or the tracking error (which we will show later).

The reminder is structured as follows. Section 2 introduces our measure of factor timing activity. Section 3 describes the data. Section 4 contains the performance analysis and section 5 describes the characteristics of actively timing funds as well as the persistence of timing measures. Section 6 analyses the risk profile of factor timers. Sections 7 links our measure to other timing measures. Section 8 describes some robustness checks and section 9 concludes.

2 Factor timing from a dynamic factor model

Traditional asset pricing factor models such as the single factor model, the Fama & 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 relationship, represented by β, is traditionally assumed to be constant over time which allows to estimate values of β using OLS regressions. Even if this assumption of constant βs holds for single securities it might not be valid for managed portfolios like mutual funds as pointed out by Mamaysky et al. (2008). Any varying exposure due to timing attempts would not be reflected correctly. We model such timing attempts and therefore apply the Carhart (1997) four factor model with time-varying risk factor exposures β_t which is represented by the following state space model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{mkt,i,t} * (r_{m,t} - r_{f,t}) + \beta_{smb,i,t} * SMB_t + \beta_{hml,i,t} * HML_t + \beta_{mom,i,t} * UMD_t + \varepsilon_{i,t},$$

$$\beta_{j,i,t} = \beta_{j,i,t-1} + \theta_{j,i}(\mu_{j,i} - \beta_{j,i,t-1}) + \eta_{j,i,t} \quad \text{for} \quad j \in \{mkt, smb, hml, mom\},$$

where $r_{m,t}$ is the market return and $r_{f,t}$ the risk-free rate at time t. SMB_t, HML_t and UMD_t represent the Fama & French and Carhart risk factors at time t. We assume the disturbance terms $\varepsilon_{i,t}$ and $\eta_{j,i,t}$ to be normally independently distributed and label their standard deviations σ_ε, and τ_{mkt}, τ_{smb}, τ_{hml}, τ_{mom}, respectively. The model differs from a classical Carhart (1997) model as it allows the factor loadings to change over time. The process of single factor loadings is assumed to be a mean-reverting process. The standard deviation of the white noise component of this process, e.g. τ_{mkt}, is an estimate of a fund's variation in risk factor exposure across time. We therefore consider it as a measure of a fund managers' involvement in timing the respective risk factor. The higher τ_{mkt} the more volatile (ceteris paribus) is the estimated process of $\widehat{\beta_{mkt,t}}$ (ceteris paribus) and the more does the fund's exposure to market risk vary from period to period. The same principle applies to the other risk factors. The model does not assume any dependencies or correlations between the dynamics of single risk factor exposures.

We use a Kalman filter and Kalman smoother technique to estimate the dynamics of this model for single mutual funds over a period of 156 weeks in a rolling manner.

Using return data from the [t-156;t-1] time window we estimate the model and get rolling estimates of β_t as well as for τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom}. These values of τ allow us to directly draw a conclusion about a fund's risk factor timing activity during the [t-156;t-1] time period: A higher τ describes a more actively timing fund. We also estimate values of θ and μ for each fund and each 156 week window. The four values of μ (one with respect to each risk factor) indicate the mean factor loadings the process reverts to. The four values of θ indicate the pace at which the loadings revert to its mean. 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 all our results enforcing this random walk. Our results remain unchanged and are statistically even more significant.

We also summarize a fund's timing activity with respect to different risk factors in an aggregate Timing Indicator. We calculate this Timing Indicator as follows: At each point in time we calculate the cross-sectional mean and standard deviation for each factor timing measure τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} and standardize all estimated values of σ by demeaning (with the cross-sectional mean) the estimates and dividing them by the respective cross-sectional standard deviation. Our Timing Indicator is the average of the four standardized values for τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom}, i.e.

$$Timing\ Indicator = \frac{1}{4}\left(\frac{\tau_{mkt}-\overline{\tau_{mkt}}}{SD(\tau_{mkt})} + \frac{\tau_{smb}-\overline{\tau_{smb}}}{SD(\tau_{smb})} + \frac{\tau_{hml}-\overline{\tau_{hml}}}{SD(\tau_{hml})} + \frac{\tau_{mom}-\overline{\tau_{mom}}}{SD(\tau_{mom})}\right),$$

where $\bar{\tau}$ is the cross-sectional mean and $SD(\tau)$the cross sectional standard deviation of τ.

3 Sample selection and data

We apply the dynamic factor model to actively managed mutual equity 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. We start out considering 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*. The Large Cap style category is primarily assigned to index funds and is therefore excluded from our further analysis.[3] Funds that cannot be matched to one of those 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 by 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. We get quarterly fund data on fund age, management tenure, turnover ratio, total expense ratio and total assets under management as well as daily net return data for the remainder sample. We aggregate those data across share classes of each fund. Fund age is the age of the oldest share class, total net assets is 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 class data, weighted by share classes' total net assets. We additionally calculate 12-months fund flows for each fund by $flow_t = (tna_t - tna_{t-1year})/(tna_{t-1year} * (1 + ret_{(t-1,t)}))$, where tna_t are the total net assets at time t and $ret_{(t-1,t)}$ is the 1-year return (net after fees) during the past 12 months.

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

We convert daily return data into weekly data as Bollen/Busse (2001) suggest to use more frequent than monthly return intervals. Since we do not have weekly observations on fund characteristics we assign the last available data to each fund if this is not older than 52 weeks. In case only less current fund characteristics are available we consider these data to be missing.

We apply the dynamic factor model over a period of 156 weeks to our sample and require funds to have at least 104 weekly return observations during the prior 156 weeks. Because daily return data are available from the end of 1998 our final dataset reaches from the end of 2000 to 2016. It contains 1,356,164 weekly observations of 3,900 funds. We winsorize the data on age, tenure, turnover, expense ratio, flows, and total net assets on the 1%-level and censor observations for which the estimated values of τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} are amongst the highest 1% of all observations.[4] This leaves us with 1,313,931 weekly observations. As daily risk factors, we use daily Fama and French (1993) as well as momentum risk factors from Kenneth R. French's website.

Table 1 provides summary statistics for the main variables of the empirical analysis. The average and median fund size is 1,333 and 325 million USD which indicates a skewed distribution of size across funds. On average, the age of the funds is 15.8 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 370%. Total expenses range from 0.14% p.a. to 2.26% p.a. with a mean of 1.14%. The average yearly flow has been positive (1.6%) but its median is at -6.1% suggesting that there are high net inflows into few funds but smaller net outflows from the majority of funds. All four estimated parameters of τ show a wide heterogeneity in timing activity ranging from very stable factor exposure variation (τ<0.001) to values as large as 10.8 times the average τ. The mean

[4] We drop the most extreme results because when applying the Kalman filter to more than 1.3 million time windows the maximum likelihood optimizing algorithm does not converge in some cases. In those very few cases we end up with extremely large and wrong numerical values. Cutting off the 1% most extreme results will drop those biased results. Our results remain qualitatively unchanged if we only drop the 0.5 % most extreme results.

variation in factor loading (τ) is highest for the HML 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 due to its construction the average Timing Indicator is 0 but there are some funds that are very engaged into timing (maximum Timing Indicator = 8.72). 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, which contain a large number of large cap funds. The row "other" summarizes very few observations of funds that were classified as large cap funds as well as funds that have been included in our sample but whose assigned styles change during the sample period.

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 timing measures is between 0.25 and 0.33. Funds with high timing activity (measure by a high Timing Indicator) tend to be smaller, more expensive and have a higher turnover ratio. This provides evidence that factor timing is not a randomly occurring observation but an intended and paid-for active trading strategy.

Figure 2 plots the median timing measures over time. Measures on SMB and HML timing appear to be very volatile with peaks in HML-timing between 2004 and 2007 as well as after 2014 and highest SMB-timing activity during 2009. The lower part of the graph shows the 50% confidence interval of the Timing Indicator. The 50% confidence interval reaches from about -0.45 to 0.25 thus indicating a positive skewness in timing which might be due to the fact that factor timing measures τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} are bounded by 0. The mean value of Timing Indicator is 0 by construction. The 50% confidence interval suggests a rather stable distribution of the Timing Indicator over time.

4 Performance effects of factor timing activity

We are interested in the relationship between timing activity and the performance of mutual funds and test whether timing activity, measured by τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} and the aggregated Timing Indicator, predicts future raw returns and future abnormal returns.

We use portfolio sorts to investigate the impact of factor timing on fund performance. Each week we sort all funds by factor timing with respect to either a specific risk factor (i.e. by either τ_{mkt}, τ_{smb}, τ_{hml}, or τ_{mom}) or the aggregated Timing Indicator and assign them to five quintile portfolios, each portfolio holding one fifth of all funds. Factor timing measures are determined by applying the dynamic factor model over the previous 156 weeks. As factor timing differs significantly between fund styles, we sort the funds within all funds of 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 those portfolios unchanged for one week and calculate the quintile portfolio returns during that week as the equal weighted mean of the funds' returns within this portfolio. We resort the portfolios every week by the most recent timing measures and this way get a return time series for each quintile portfolio. Panel A of table 3 reports the average returns of those return series. Each column refers to a specific sorting criteria. Regressing the return time series on a Carhart (1997) factor model gives abnormal returns. Panel B of table 3 reports those alphas. Results on raw returns are significant for momentum timing only. The most actively timing funds underperform the least actively timing funds (w.r.t. momentum timing) by 103 basis points in terms of absolute returns. The results on abnormal returns are significant for momentum timing as well as for the aggregated Timing Indicator. The most actively timing portfolio underperforms the least actively timing portfolio by 123 basis points and 140 basis points p.a., respectively. The impact of market, size, and value timing on either raw returns or abnormal returns is negative, yet statistically not significant. The abnormal return decreases monotonously as timing activity with respect to either risk factor increases (it decreases almost monotonously for size timing). Figure 3 visualizes this relationship using the aggregated Timing Indicator. The relationship is strongly monotone and even close to linear.

Notably, no single quintile portfolio has a positive alpha. This is not surprising as we use net returns and funds are known to on average have a significantly negative abnormal return.

In order to control for fund characteristics we also investigate the performance effects of timing using Fama-MacBeth regressions. We calculate a fund's abnormal return at time t, α_t, as the return difference between actual fund performance in week t and the expected fund performance calculated from a Carhart model, i.e. $\alpha_t = r_{i,t} - E[r_{i,t}]$, where

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

We estimate values of β_t when applying the dynamic factor model to the time interval (t-156,t-1). The Kalman filter calculates estimates of β_t and we therefore get abnormal returns for all the funds in our sample at each point in time. We drop observations for which we have less than 104 return data during the prior 156 weeks. We conduct Fama-MacBeth regressions using both, raw returns and abnormal returns. Table 4 reports our results with Newey-West standard errors. Measures of market timing, size timing, and momentum timing, i.e. τ_{mkt}, τ_{smb}, and τ_{mom} have on average a negative, yet statistically insignificant effect on raw returns. The average effect of the Timing Indicator on future raw returns is also negative, but statistically insignificant. Our results become clearer as we turn to risk-adjusted abnormal returns. Measures of market, size, and momentum timing have a negative effect on future abnormal fund returns and the effect is statistically significant. The effect remains significant if we use the aggregated Timing Indicator instead of single factor timing measures. There is no significant impact of value timing on either raw returns or abnormal returns.

The average cross-sectional standard deviation of market, size, and momentum timing measures are 0.04, 0.16, and 0.05. The economic interpretation of the Fama-MacBeth regression results is therefore as follows: A one standard deviation increase of market timing activity leads on average to a decrease of annualized raw return by 11 basis points. Increasing either size or momentum timing by one standard deviation leads to a raw return decrease of 31 or 21 basis points p.a. The average effect of a one standard deviation increase in market, size and momentum timing on abnormal returns is 30, 39, and 24 basis points p.a., respectively.

Next we focus on the mid- and long-term effect of timing activity. Each week we therefore sort funds into five quintile portfolios by the aggregated Timing Indicator and observe the portfolios' cumulative raw return and cumulative abnormal return over the following 1 week, 1 month, 6 month, 1 year or 3 year time horizon. Weekly abnormal returns are the difference between actual returns and the expected returns as estimated by a Carhart model. Panel A of table 5 reports the cumulative raw returns (net of fees) and Panel B the cumulative annualized abnormal returns over the respective time periods. We use Newey-West standard errors with a lag length of 156 weeks as 3-year cumulative returns will be autocorrelated by construction.

Results show that the negative effect of factor timing on both, raw returns and abnormal returns gets economically smaller over time but its statistical power increases and results become significant even for raw returns for 1-month time horizons and above. Funds with high factor timing activity underperform those with low timing activity by 94 basis points p.a. over the next week, 40 basis points p.a. over the next year and 17 basis points p.a. over the next 3 years in terms of raw return. The underperformance even increases to a statistically significant 135 basis points p.a. during the next week, 98 basis points p.a. during the next year and 76 basis points p.a. during a 3-year period when controlling for risk, i.e. measuring abnormal returns instead of raw returns. The abnormal returns increase monotonously when turning from most actively timing to least actively timing quintile portfolios for all time horizons.

5 Characteristics of factor timers and persistence of timing activity

Fund characteristics

Factor timing seems to be very unfavorable for investors. We study which fund characteristics determine a fund's factor timing activity. Since timing measures were estimated using 3-year time windows during our 09/1998-12/2016 sample period, we split our sample into 6 non-overlapping subperiods, namely 1999-2001, 2002-2004, 2005-2007, 2008-2010, 2011-2013, and 2014-2016. We regress the fund characteristics at the beginning of these periods on the timing measures during those periods to observe the relationship between ex-ante fund characteristics and timing activity. Table 6 reports the results. Management tenure, total expense ratio and turnover ratio have strong positive effects on factor timing. The positive relationship between manager tenure and factor timing with respect to each factor is in line with predictions of Chevalier and Ellison (1999) who suggest that younger managers have an incentive to not expose their portfolios to unsystematic risk. The impact of management tenure on timing activity is decreasing as both, the use of ln(manager tenure) and the negative coefficient on a the squared term suggest (unreported results). The positive regression coefficients of expenses as well as turnover help us understand the characteristics behind factor timing. As factor timing increases with turnover ratios, factor timing describes an intended and actively enforced investment strategy rather than a coincidental return series characteristic. The positive relationship between timing activity and expense ratio suggests that timing risk factors is associated with costs either due to higher operational expenses (e.g. trading costs) or due to higher management fees because factor timing requires skill and research effort. Funds with a better past performance (a larger lagged alpha) tend to be actively timing momentum. Other characteristics either don't have a statistically significant impact on factor timing or their impact is ambiguous across risk factors. We also include style dummies into our regressions and find that factor timing is higher for growth funds as well as mid, small and especially micro-cap funds.

Persistence of timing activity

Investors can only avoid factor timing funds and not suffer from their bad performance if factor timing is a persistent fund characteristic rather than a quickly changing investment trend. We therefore study the persistence in timing. Each week we sort funds into 10 deciles by the aggregated Timing Indicator. The transition matrix in table 7 displays changes in deciles over a period of 1 year and 3 years as well as the attrition rate, i.e. the percentage of funds that leave our sample within those next 1 year and 3 years respectively. We exclude observations from the last year (in the one year case) and the last three years (in the 3-year case) of our sample period to not falsely increase the attrition rate. Results are unequivocally: Factor timing is a very persistent fund characteristic. Almost three fourth of all funds in the lowest timing decile remain in the lowest decile after one year and 70% of all funds in the highest timing decile remain in this decile after one year. This might partially be by construction since factor timing has been estimated over a 3-year time window. Panel B therefore displays transitions aver a period of 3 year. Results do not change qualitatively. After three years 53% (47%) of the funds in the lowest (highest) timing activity decile still remained in this decile. Figure 4 underlines this finding graphically. It displays the time series of average values of the aggregated Timing Indicators of the ten decile portfolios sorted by the Timing Indicator, starting 52 weeks prior and ending 312 weeks after portfolio formation. The difference in Timing Indicators gets smaller during the 52 weeks before and 156 weeks after the portfolio formation but the order does not change and no two decile portfolios converge to a common value. After 156 weeks the average Timing Indicators remain almost unchanged so one can conclude the persistence in factor timing activity remains strong even in the long run (i.e. 6 years).

The transition matrices (table 7) reveal another result: Funds with a higher factor timing activity are more likely to drop from our sample within the next years. Only 6% (15%) of all the funds in the lowest timing activity decile leave our sample within the next year (3 years), but the probability increases almost monotonously as factor timing activity increases and reaches 15% (27%) for the

10% of the funds with highest measures of factor timing. This result is not very surprising if one considers the prior discussed relationship between factor timing and returns as well as size: Fund that time risk factors more actively tend to are smaller and have negative future returns that will cause additional outflows. Those funds are less likely to survive.

6 On the riskiness of factor timing

If a fund's factor timing activity has a negative impact on future fund return, does a more desirable risk profile possibly compensate for this negative performance? So far we have shown that factor timing has a negative influence on *average risk adjusted returns*. We will now turn to the distribution of those abnormal returns. Each week we calculate each fund's abnormal return, defined by the difference between the fund's actual return and the expected return according to a Carhart model with the betas estimated over the previous 3 years by an OLS regression. Each week we additionally assign each fund to one of five timing quintiles by either its market timing, size timing, value timing, momentum timing activity or by the aggregated Timing Indicator. Once again, those measures are estimated using the dynamic factor timing model over the previous 3 years. Figure 5 plots the cumulative distribution function of the panel data of abnormal returns by quintile. The graphs indicate that funds with a higher timing activity with respect to either risk factor have a wider distribution of abnormal returns. Most importantly, they have a larger number of large negative abnormal returns but also seem to have more relatively large positive returns. Unreported results on the standard deviation of abnormal returns confirm this impression. The standard deviation of this distribution of abnormal returns monotonously increases from 0.43% for least actively timing funds to 0.54%, 0.63%, 0.78% for the more actively timing funds to 0.99% for funds within the highest actively timing quintile when sorted on the aggregated Timing Indicator.

Further evidence of the higher risk resulting from factor timing is gained from a regression of risk indicators on factor timing measures.

We consider four different risk indicators, namely the market beta from a one-factor model, the weekly return volatility, and two versions of the standard deviation of the residuals from a Carhart factor model regression (representing idiosyncratic risk). Those risk indicators are regressed on the market, size, value and momentum factor timing measures calculated from the dynamic factor model over a 3 year period. The risk indicators are calculated over a period of 3 years. The Return volatility $\sigma(r_{weekly})$ is the standard deviation of weekly raw returns. The market beta, β_{CAPM}, is calculated from an OLS regression of returns on a one-factor model. The two versions of residual volatility, $\sigma(\varepsilon)_{OLS}$ and $\sigma(\varepsilon)_{DFM}$, are calculated from OLS regression and the dynamic factor model, respectively. Table 8 summarizes the results using Fama-MacBeth regressions with Newey-West standard errors. The time periods factor timing and risk indicators are obtained from either overlap (panel A) or the factor timing measures are calculated during the 3 years preceding the 3 years of risk indicator calculation (panel B). The results support the impressions we gained from the cumulative distribution functions: Timing activity with respect to either risk factor is linked to an overall higher return volatility. It is especially linked to a higher systematic market risk. Those results hold when measuring timing either contemporaneously or by lagged values. Timing activity also tends to increase the idiosyncratic risk, although results are not statistically significant for all risk factor timing measures when measuring idiosyncratic risk from the dynamic factor model. Overall, we conclude that funds with a higher factor timing activity expose their investors to additional systematic (with respect to market risk) and idiosyncratic risk.

7 Timing Activity, Tracking Error and R^2

We have measured timing activity as the standard deviation of factor loadings in a dynamic factor model. Timing activity might be closely linked to an overall fund activeness. Tracking Error (TE), i.e. the standard deviation of the return difference between a fund and its self-declared benchmark, is a traditional measure of fund activeness. A study of Amihud & Goyenko (2013) introduces the R^2 obtained from a multi-factor OLS-regression as a measure of fund activeness and finds it to be a predictor of future returns. We will therefore link our findings to those measures of fund activeness.

Tracking Error (TE)

TE is the standard deviation of a fund's excess return over its benchmark. We use the TE data available from Morningstar Direct, which is calculated with respect to a fund's self-declared benchmark index. Morningstar Direct allows us customize the calculation method and we download TE values calculated from 3 years of weekly return data.

A fund's TE should be positively correlated to fund activeness since a fund that closely follows a benchmark index and therefore has a low TE is expected to have low timing activity. We get first evidence on this negative relationship between factor timing and TE by computing the correlations of TE and our factor timing measures in our panel data, where TE is calculated over the same 3 year horizon the timing measures are estimated on. The correlation is between 0.12 for value timing and 0.27 for size timing.

We split our sample into non-overlapping 3-year subperiods, namely 1999-2001 to 2014-2016. This splitting is done to handle the problem of autocorrelation of TE and the timing measures that was induced by construction since we calculated those values with rolling regression windows. For each subperiod we estimate the Timing Indicators from the dynamic factor model and we calculate values of TE. We then regress TE on the obtained factor timing measures. Table 9 reports the results. As expected, the factor timing measures with respect to either risk factor have a significant negative influence on TE and they even explain 44% of the cross-sectional variation of TE-values.

Since factor timing is so closely linked to TE we will study how the return predictive effect of factor timing measures change if we also control for TE. We therefore repeat the Fama-MacBeth regressions of one week abnormal returns on factor timing measures but also include the TE calculated from the previous 3 years into the set of independent variables. Results are reported in table 10. Adding TE to the regression model does not change the return effect of factor timing. We furthermore find that TE has a negative, yet statistically not significant, impact on future abnormal fund returns when controlling for our factor timing measures. Any return predictive effect of TE (as found, e.g. by Wermers (2003)) is captured in our timing measures.

Amihud & Goyenko (2013)

Amihud and Goyenko (2013) state that a low R^2 obtained from an OLS regression of fund returns on a Carhart 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. This approach is linked to our factor timing measures as a high activity in factor timing would lead to low R^2 when estimating OLS regressions. Following this argumentation our results would contradict the findings of Amihud and Goyenko (2013).

We compute the correlations of R^2 and the factor timing measures in the panel data, where the R^2 is obtained from OLS regressions on a Carhart model. R^2 and timing measures are estimated over the same 3 year time horizon. The correlation is between -0.42 for market timing and -0.31 for momentum timing.

As before, table 9 reports the results of a regression of R^2 on measures of factor timing activity. We split our sample into non-overlapping 3-year subperiods to deal with autocorrelation. Values of factor timing and R^2 are calculated over those 3-year periods. Factor timing shows to have a significant negative influence on R^2 and explain 37% of the cross-sectional variation of R^2.

As in the case of TE we add R^2 as a control variable to the Fama-MacBeth regression. Results are reported in table 10. Including R^2 in the regression model does not change the return effect of factor timing. As suggested by prior research, R^2 has a negative impact on future abnormal fund returns, even when controlling for our factor timing measures. We therefore conclude that our factor timing measure contains return predictive information beyond what is captured by the R^2 measure proposed by Amihud & Goyenko (2013).

Including both, TE and R^2 into the regression model does not change the return effects of either measure.

8 Robustness tests

We conduct additional robustness checks to validate our findings. We use a Fama & French 3-factor model instead of the Carhart model for estimating factor timing activity. The factor timing measures with respect to market risk, the size and the value factor picked up some of the momentum timing activity and increased slightly but they were highly correlated to the timing measured we obtained from our dynamic four factor model. The results on the return predictive power remained qualitatively unchanged.

We also conducted our analysis based on gross returns (before fees). We added the ex post reported total expense ratio to the weekly returns prior to estimating timing measures and recalculated all our analysis. The Results remain qualitatively unchanged, thus suggesting that higher fees are not driving our results.

We also used different dynamics within the dynamic factor model. Following Black et al. (1992) we assume the factor loadings to follow a random walk instead of a mean-reverting process. Our main results are qualitatively comparable and statistically even stronger than with the mean-reverting model. We yet decided to use the mean-reverting process throughout our paper as we consider it to be the economically more plausible model.

We also conduct portfolio sorts using value weighted portfolios. Results remain unchanged.

To mitigate concerns that the return effect of factor timing might be caused by fund characteristics other than timing we apply a matching procedure. Each week and for each of the 20% funds with the highest market timing activity we match a fund with low market timing by either a propensity score matching with the propensity score of being amongst the most actively timing funds estimated from a logit regression of fund characteristics. In an alternative approach we match fund by the mahalanobis distance on size, past alpha and size, value and momentum timing. For both matched samples we repeat the Fama-MacBeth regression. The results confirms our findings on market timing and return predictability. We repeat this matching procedure for size, value and momentum timing and once again our results remain valid.

Instead of including style dummies we run the return analysis for the subsamples of funds with the same style. We observe the negative effect of factor timing within each style category.

We also ran the return analysis with beta estimates calculated from a static OLS regression instead of obtaining beta estimated from our dynamic factor model. This doesn't change our results. In the same way we analyzed the distribution of abnormal returns as an indicator for the higher risk of actively timing funds with beta estimated from dynamic model instead of a static OLS regression. Our results remain robust with respect to this alternation.

9 Conclusion

Mutual fund managers vary their exposure to risk factors over time. Can mutual fund managers successfully time risk factors and generate positive abnormal returns from it? We propose a new measure of a fund's factor timing activity and asses the timing ability within a sample of US mutual funds during the 2000-2017 time period. Factor timing activity appears to be a very persistent fund characteristic. We find that factor timing with respect to either the market, size, or the momentum risk factor is associated with negative future abnormal returns. Furthermore, funds with a higher risk factor timing activity expose investors to additional systematic and idiosyncratic risk.

We conclude that investors should resist the temptation to invest into fund that intentionally or coincidentally vary their exposure to risk factors over time. Deviations in risk factor exposures should rather be interpreted as a signal of weak conviction and a lacking investment strategy than a signal of skill.

References

Amihud Y., Goyenko R. (2013). Mutual fund R^2 as predictor of performance. Review of Financial Studies 2, 667–94.

Benos, E., Jochec, M., & Nyekel, V. (2010). Can mutual funds time risk factors?. The Quarterly Review of Economics and Finance, 50(4), 509-514.

Black, A., Fraser, P., & Power, D. (1992). UK unit trust performance 1980–1989: a passive time-varying approach. Journal of Banking & Finance, 16(5), 1015-1033.

Bollen, N. P., Busse, J. A. (2001). On the timing ability of mutual fund managers. The Journal of Finance, 56(3), 1075-1094.

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

Cremers K. J. M., Petajisto A. (2009). How active is your fund manager? A new measure that predicts performance. Review of Financial Studies 22, 3329–65.

Daniel K., Grinblatt M., Titman S., Wermers R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–58.

Doshi, H., Elkamhi, R., & Simutin, M. (2015). Managerial activeness and mutual fund performance. Review of Asset Pricing Studies, 5(2), 156-184.

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(4), 643-667.

Fama, E. F., French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.

Fama E. F., French K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. Journal of Finance 65, 1915–47.

Ferson, W. E., Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. The Journal of Finance, 51(2), 425-461.

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. Journal of Finance 51, 783–810.

Henriksson, R. D., Merton, R. C. (1981). On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. Journal of business, 513-533.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. Review of Financial Studies, 24(8), 2575-2616.

Jensen M. C. (1968). The performance of mutual funds in the period 1945-1964. Journal of Finance 23, 389–416.

Kacperczyk, M., Nieuwerburgh, S. V., Veldkamp, L. (2014). Time-varying fund manager skill. The Journal of Finance, 69(4), 1455-1484.

Kacperczyk, M., Seru, A. (2007). Fund manager use of public information: New evidence on managerial skills. The Journal of Finance, 62(2), 485-528.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. The Journal of Finance, 60(4), 1983-2011.

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

Mamaysky, H., Spiegel, M., Zhang, H. (2008). Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies, 21(1), 233-264.

Petajisto, A. (2013). Active share and mutual fund performance. Financial Analysts Journal, 69(4), 73-93.

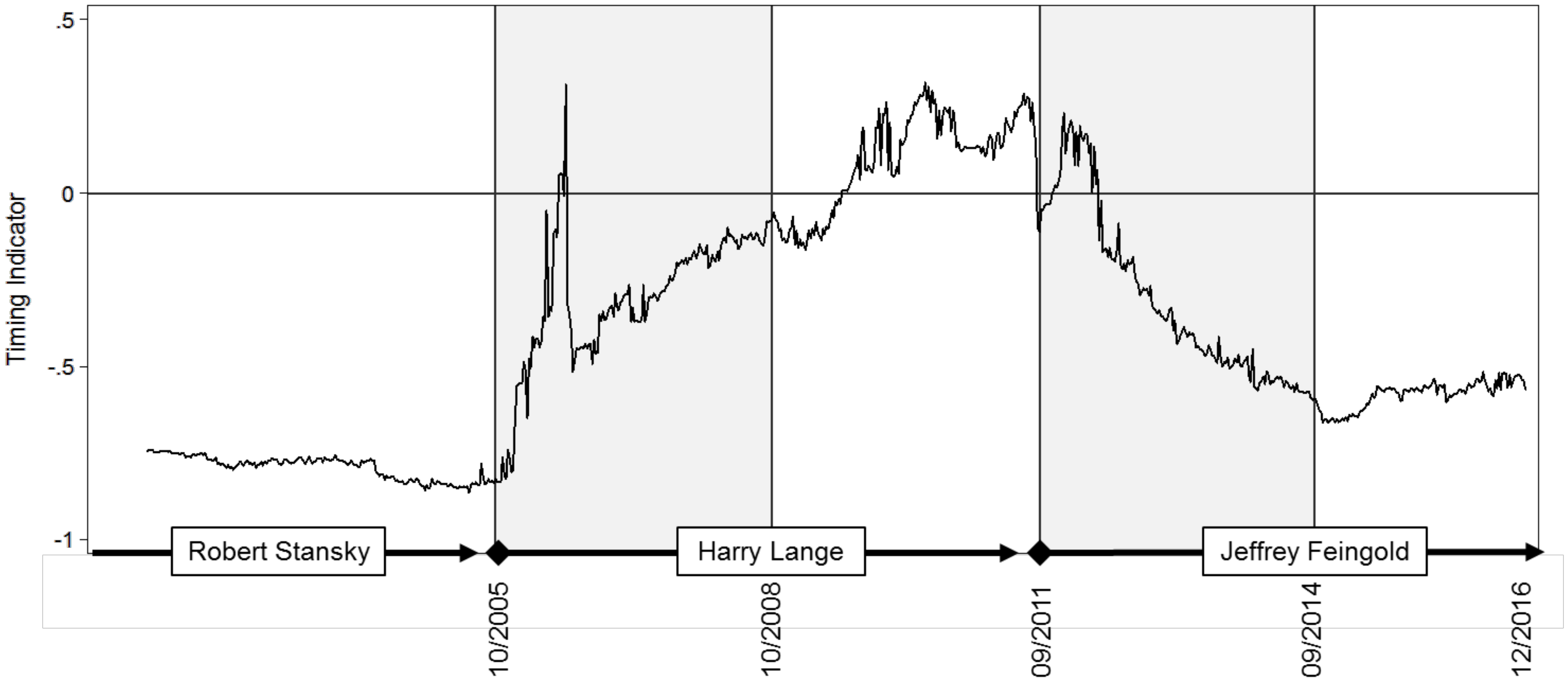
Swinkels, L., & Tjong-A-Tjoe, L. (2007). Can mutual funds time investment styles?. Journal of Asset Management, 8(2), 123-132.

Treynor J., Mazuy, K. (1966). Can Mutual Funds Outguess the Market? Haward Business Review 44, July-August 1966, 131-136.

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

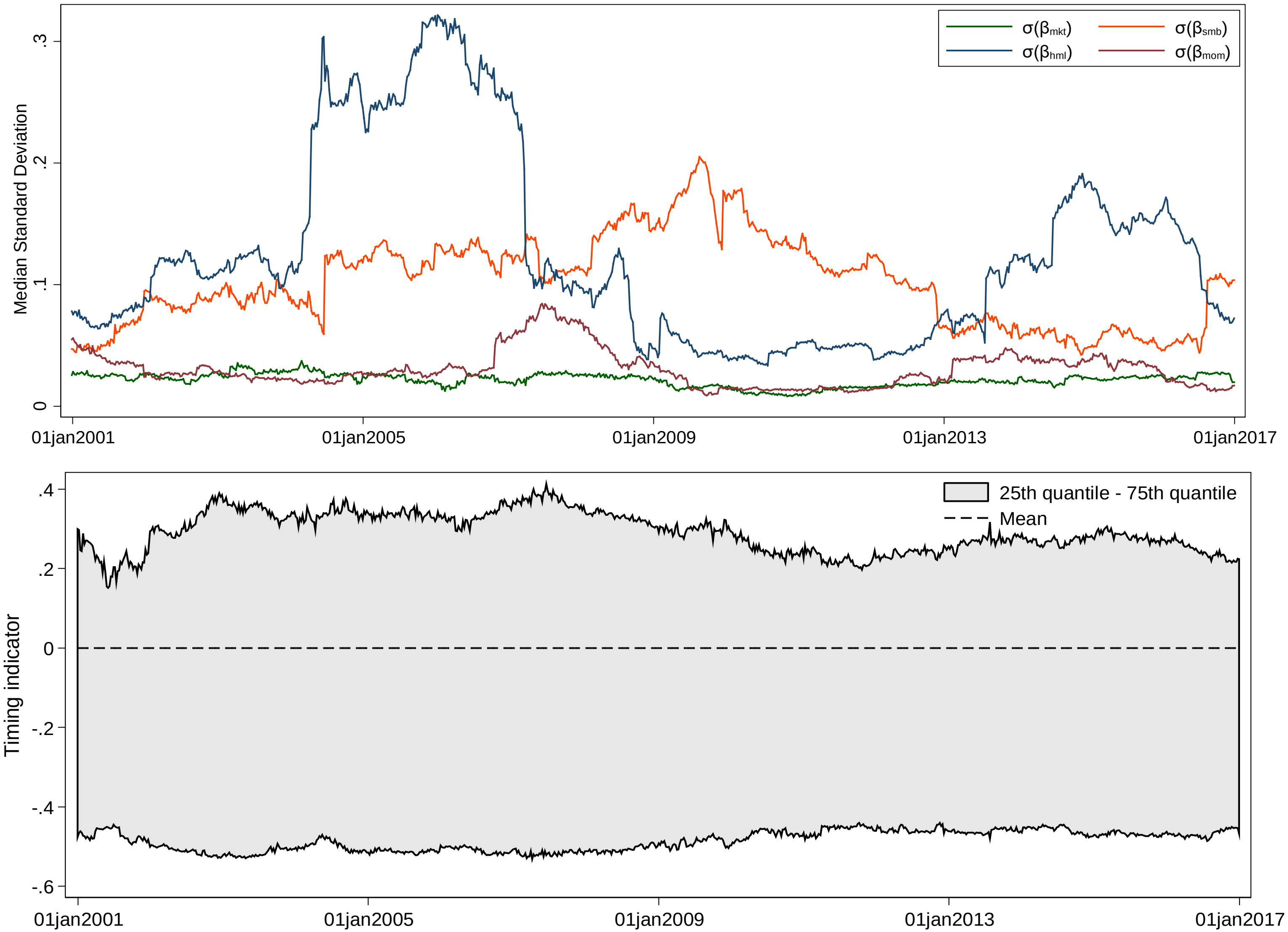
Wermers R. (2003). Are Mutual Fund Shareholders Compensated for Active Management “Bets”? Working Paper.

Figure 1: Timing Indicator of the Fidelity Magellan fund, 2002-2016



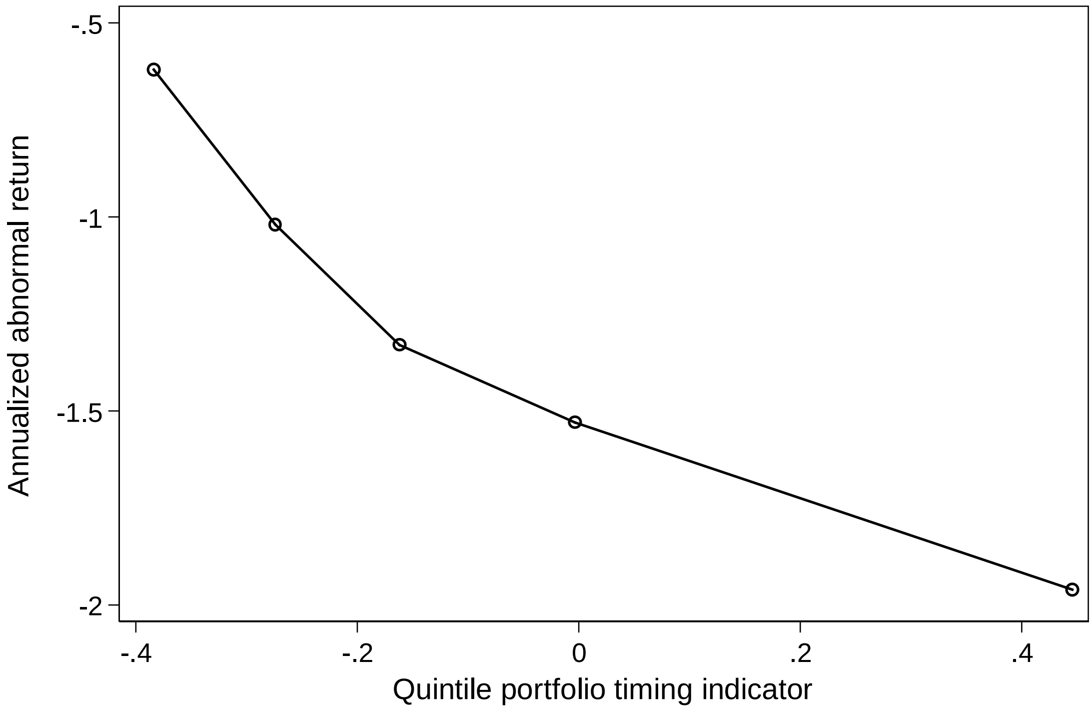
The graph plots the 2002-2016 time series of the aggregated Timing Indicator for Fidelity's Magellan fund. We calculate the Timing Indicator from the timing measures obtained from a dynamic version of Carhart's (1997) 4-factor model as introduced in section 2. A Timing Indicator >0 indicates an above average timing activity with respect to the market, size, value and momentum risk factors. The graph highlights management changes in October 2005 (Lange follows Stansky) and September 2011 (Feingold follows Lange). The timing measures are estimated from 3 years of historical data. The grey shaded areas indicate time periods for which the timing activity is estimated using fund returns from more than one management periods.

Figure 2: Timing measures and aggregated Timing Indicator over time



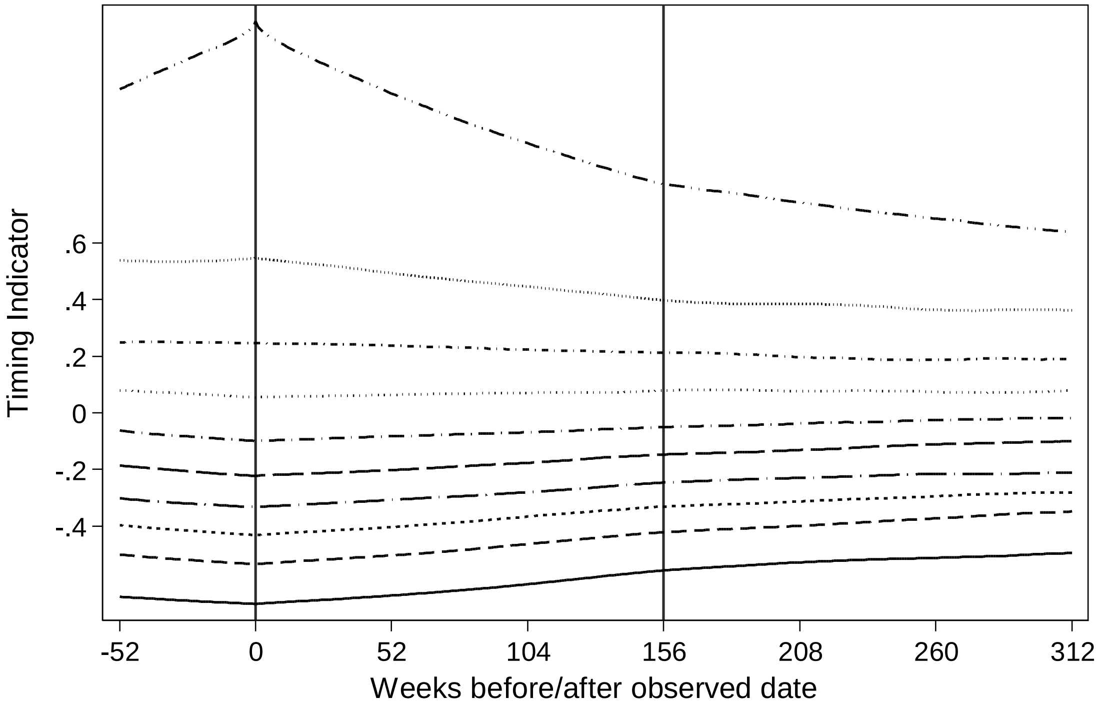
The upper graph shows the evolution of cross-sectional median timing measures τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} over time. The lower part shows the mean Timing Indicator as well as the 25% and 75% quantile. See section 2 for the calculation of the timing measures and the Timing Indicator.

Figure 3: Average quintile portfolio timing and abnormal returns



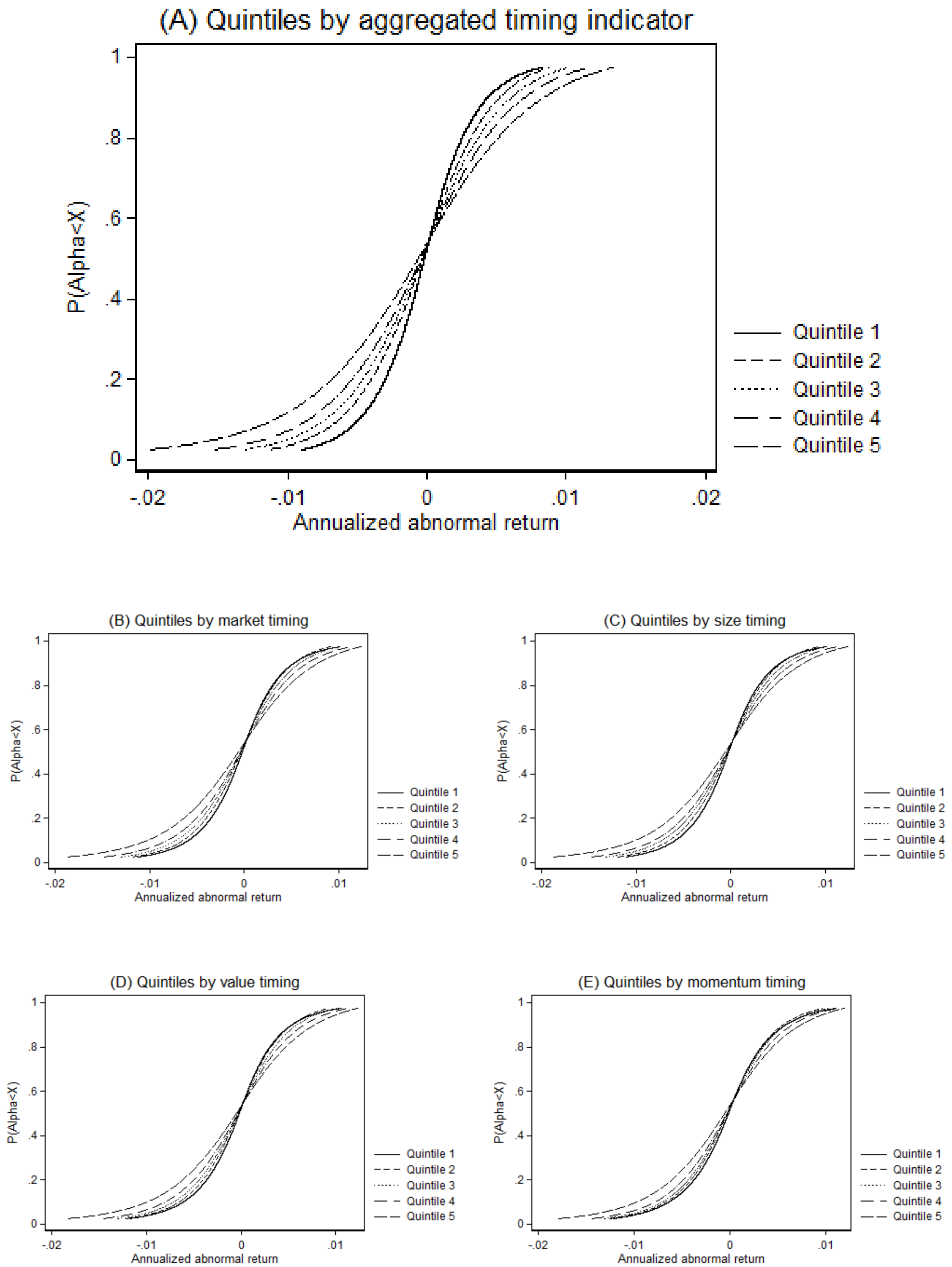
The graphs shows the time series average of the Timing Indicator (see section 2 for the calculation of the Timing Indicator) for quintile portfolios sorted by the respective timing measure (x-axis) and the abnormal return of the quintile portfolio (y-axis). The portfolios are resorted every week and portfolio returns are calculated as the equal weighted mean of all the funds within a portfolio. A portfolio's abnormal return is calculated from regressing the portfolio return series on a Carhart (1997) risk factor model and annualizing the alpha.

Figure 4: Persistence of the aggregated Timing Indicator



This graph displays the evolution of the aggregated Timing Indicator (see section 2 for the calculation of the Timing Indicator) over time. At a fixed time t funds are sorted into ten deciles by the actual value of the aggregated Timing Indicator. The average values of the aggregated Timing Indicators of those unchanged portfolios are observed over time, starting 52 weeks prior and ending 312 weeks after time t. We do so for each t during our sample period and display average values for the 10 decile portfolios during the 52 weeks prior to and the 312 weeks after the portfolio formation.

Figure 5: Cumulative distribution function of abnormal returns



Each week all funds are assigned to one of five timing quintiles by either the market timing (B), size timing (C), value timing (D), momentum timing activity (E) or by the aggregated Timing Indicator (A), each estimated from the dynamic factor timing model over the previous 3 years. Fund's weekly abnormal returns are calculated from a Carhart (1997) model where the betas are estimated from an OLS regression over the previous 3 years. The graphs (A)-(E) plot the respective cumulative distribution functions of the abnormal return panel data.

Table 1: Descriptive statistics and timing activity by fund style

Panel A: Fund characteristics

	# Obs.	Mean	Min	25%	50%	75%	Max
Number of funds	3,900						
Fund-Week-observations	1,313,931						
Total assets (in mn. USD)	1,313,931	1,333	18	107	325	1,052	21,268
Fund age (years)	1,313,229	15.76	2.57	7.67	12.72	19.06	72.42
Manager tenure (years)	1,053,750	7.50	0.42	3.68	6.33	10.11	25.85
Turnover ratio	1,158,594	0.75	0.03	0.30	0.58	0.99	3.70
Total expense ratio (in %)	1,161,652	1.15	0.14	0.92	1.13	1.37	2.26
Relative fund flow	1,313,034	0.017	-0.601	-0.151	-0.061	0.078	1.899
τ_{mkt}	1,313,931	0.033	< 0.001	0.009	0.020	0.042	0.310
τ_{smb}	1,313,931	0.148	< 0.001	0.040	0.096	0.193	1.174
τ_{hml}	1,313,931	0.185	< 0.001	0.039	0.100	0.227	1.989
τ_{mom}	1,313,931	0.046	< 0.001	0.010	0.026	0.060	0.390
Timing Indicator	1,313,931	0	-0.982	-0.480	-0.176	0.285	8.722

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

Fund Style	# Funds / # Obs.	τ_{mkt}	τ_{smb}	τ_{hml}	τ_{mom}	Timing Indicator
Growth and Income	795 / 254,387	0.021	0.096	0.119	0.030	-0.325
Growth	1,667 / 551,621	0.032	0.143	0.177	0.044	-0.028
Hedged	70 / 10,724	0.027	0.158	0.174	0.043	0.010
Income	188 / 61,515	0.021	0.099	0.127	0.032	-0.292
Mid Cap	447 / 160,438	0.045	0.197	0.230	0.060	0.217
Small Cap	677 / 256,370	0.040	0.180	0.242	0.057	0.056
Micro Cap	46 / 18,112	0.059	0.282	0.331	0.089	0.826
Other	10 / 764	0.036	0.146	0.133	0.029	0.182

Panel A provides a descriptive overview over the sample size and fund characteristics. Relative fund flows are calculated over the past 52 weeks. Timing measures τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} are calculated applying the dynamic factor model over a rolling 156 weeks window. Timing Indicator is the mean of the four standardized timing measures. Panel B reports the average timing measures and Timing Indicators by fund style.
Data on age, tenure, turnover, expense ratio, flows, and total net assets are wisorized on the 1%-level and observations for which the estimated values of τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} are amongst the highest 1% were dropped from the sample.

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

Panel A: Average cross sectional correlations between transfer factor loading standard deviations

	τ_{mkt}	τ_{smb}	τ_{hml}	τ_{mom}	Timing Indicator
τ_{mkt}	1.00				
τ_{smb}	0.33	1.00			
τ_{hml}	0.33	0.29	1.00		
τ_{mom}	0.25	0.30	0.26	1.00	
Timing Indicator	0.70	0.70	0.69	0.66	1.00

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

	Timing Indicator	Total exp. ratio	Turnover ratio	Relative fund flow	Total assets	ln(fund age)	ln(tenure)
Timing Indicator	1.00						
Total exp. ratio	0.34	1.00					
Turnover ratio	0.19	0.22	1.00				
Relative fund flow	0.00	-0.05	-0.05	1.00			
Total assets	-0.12	-0.31	-0.16	0.07	1.00		
ln(fund age)	-0.03	-0.06	-0.06	-0.19	0.36	1.00	
ln(tenure)	0.05	-0.03	-0.15	-0.01	0.05	0.18	1.00

Panel A reports the average cross-sectional correlations between timing measures τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} and the Timing Indicator. Timing measures are calculated applying the dynamic factor model over a rolling 156 weeks window. Timing Indicator is the mean of the four standardized timing measures. Panel B reports the correlations between fund characteristics and the Timing Indicator.

Data on age, tenure, turnover, expense ratio, flows, and total net assets are wisorized on the 1%-level and observations for which the estimated values of τ_{mkt}, τ_{smb}, τ_{hml}, and τ_{mom} are amongst the highest 1% were dropped from the sample.

Table 3: Abnormal returns of quintile portfolios sorted by timing measures

	Timing Indicator	τ_{mkt}	τ_{smb}	τ_{hml}	τ_{mom}
	Panel A: Raw returns of quintile portfolios				
Low Timing	5.94 (1.30)	5.45 (1.21)	5.57 (1.24)	5.57 (1.25)	5.84 (1.30)
(2)	5.47 (1.20)	5.37 (1.19)	5.70 (1.28)	5.46 (1.23)	5.51 (1.22)
(3)	5.21 (1.14)	5.35 (1.19)	5.13 (1.13)	5.37 (1.19)	5.35 (1.18)
(4)	5.22 (1.15)	5.40 (1.18)	5.31 (1.16)	5.16 (1.12)	5.37 (1.18)
High Timing	5.00 (1.10)	4.23 (1.12)	5.13 (1.09)	5.31 (1.13)	4.81 (1.13)
High-Low Timing	-0.94 (-1.26)	-1.03 (-0.21)	-0.44 (-0.56)	-0.26 (-0.21)	-1.03* (-1.79)
	Panel B: Abnormal returns of quintile portfolios				
Low Timing	-0.65** (-2.28)	-1.10*** (-3.79)	-1.03*** (-3.62)	-1.04*** (-2.82)	-0.85*** (-2.39)
(2)	-1.02*** (-3.50)	-1.14*** (-3.63)	-0.86*** (-2.69)	-1.16*** (-3.36)	-1.11*** (-3.29)
(3)	-1.31*** (-3.72)	-1.24*** (-3.53)	-1.46*** (-4.05)	-1.24*** (-3.48)	-1.21*** (-3.56)
(4)	-1.48*** (-3.19)	-1.32*** (-3.08)	-1.37*** (-3.26)	-1.49*** (-3.79)	-1.25*** (-3.05)
High Timing	-2.05*** (-3.27)	-1.74*** (-2.90)	-1.81*** (-3.08)	-1.57*** (-3.06)	-2.08*** (-3.89)
High-Low Timing	-1.40** (-2.17)	-0.65 (-1.10)	-0.78 (-1.44)	-0.53 (-1.00)	-1.23*** (-2.64)

This table reports the raw returns (net of fees, panel A) and abnormal returns (panel B) of fund portfolios sorted on factor timing activity. Each week we sort funds into five quintiles by either a single factor timing measure τ_{mkt}, τ_{smb}, τ_{hml} or τ_{mom} or by the aggregated Timing Indicator. We keep those portfolios constant for one week and calculate the equal weighted net portfolio return. Panel A reports the average annualized returns of these portfolios. Each column represents the sorting by a distinct timing measure. We report the returns for each quintile portfolio (Rows 1-5) as well as the difference between the most and the least active portfolios (Row 6). Panel B reports Carhart (1997) alphas of the same portfolios. 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 4: The effect of factor timing on future fund returns: Fama-MacBeth-regressions

Explanatory variables	Dependent variable: annualized return$_{j,t}$ (in %)		Dependent variable: annualized alpha$_{j,t}$ (in %)	
τ_{mkt}	-0.030 (-0.67)		-0.079** (-2.29)	
τ_{smb}	-0.020 (-1.41)		-0.025*** (-2.63)	
τ_{hml}	0.008 (0.78)		0.001 (0.20)	
τ_{mom}	-0.041 (-1.33)		-0.046* (-1.70)	
Timing Indicator		-0.554 (-1.00)		-0.884** (-2.47)
ln(tna)	-0.141 (-1.60)	-0.142 (-1.58)	-0.126* (-1.74)	-0.124* (-1.70)
ln(fund age)	0.115 (0.74)	0.171 (1.12)	0.034 (0.32)	0.059 (0.57)
ln(manager tenure)	0.138 (1.24)	0.128 (1.13)	0.080 (1.02)	0.076 (0.97)
Expenses	-0.811*** (-4.78)	-0.760*** (-4.34)	-0.741*** (-5.49)	-0.730*** (-5.45)
Turnover	-0.296 (-0.88)	-0.338 (-0.97)	-0.437** (-1.98)	-0.456** (-2.03)
Lagged Alpha	0.032 (0.75)	0.025 (0.57)	0.093** (2.16)	0.090** (2.07)
Fund Flows	0533* (1.76)	0578* (1.88)	0.534** (2.44)	0.558** (2.48)
Style Dummies	YES	YES	YES	YES
Average R^2	0.28	0.27	0.14	0.13

This table reports the result of Fama-MacBeth regressions of annualized fund raw returns (columns 1 and 2) or annualized abnormal fund returns (columns 3 and 4) on timing measures and controls. Each week, expected returns are calculated from a Carhart (1997) model where the factor loadings are estimated over the past 3 years of weekly return data from the dynamic factor model. Abnormal returns are the differences between actual weekly returns and the expected returns. Fama-MacBeth regressions are applied on the panel data of weekly abnormal returns. T-statistics are reported in parentheses. We use Newey-West standard errors (max. lag=12) for the regressions with abnormal returns as the dependent variable. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

Table 5: Mid- and long-term return effects of timing activity

Panel A: Average annualized cumulative raw returns

	1 week	1 month	6 months	1 year	3 year
Low Timing	5.94 (1.30)	5.95*** (2.64)	5.87*** (6.28)	5.92*** (9.02)	7.21*** (22.25)
(2)	5.47 (1.20)	5.48** (2.43)	5.49*** (5.87)	5.59*** (8.52)	6.98*** (21.53)
(3)	5.21 (1.14)	5.20** (2.30)	5.18*** (5.55)	5.37*** (8.18)	6.90*** (21.30)
(4)	5.22 (1.15)	5.17** (2.29)	5.31*** (5.68)	5.41*** (8.26)	6.96*** (21.48)
High Timing	5.00 (1.10)	5.08** (2.25)	5.27*** (5.64)	5.52*** (8.42)	7.04*** (21.72)
High-Low Timing	-0.94 (-1.26)	-0.87** (-2.29)	-0.60*** (-3.45)	-0.40*** (-3.52)	-0.17*** (-3.82)

Panel B: Average annualized cumulative abnormal returns

	1 week	1 month	6 months	1 year	3 year
Low Timing	-1.11*** (-7.94)	-1.12** (-8.13)	-1.05*** (-8.94)	-1.04* (-10.75)	-1.04*** (-15.59)
(2)	-1.35*** (-6.96)	-1.41*** (-7.31)	-1.35*** (-8.64)	-1.30*** (-9.92)	-1.25*** (-12.17)
(3)	-1.81*** (-5.26)	-1.80*** (-5.37)	-1.74*** (-5.97)	-1.67*** (-6.25)	-1.49*** (-7.00)
(4)	-2.17*** (-4.62)	-2.22*** (-4.54)	-2.01*** (-4.83)	-1.96*** (-5.41)	-1.69*** (-6.04)
High Timing	-2.46*** (-3.59)	-2.36*** (-3.79)	-2.18*** (-4.26)	-2.02*** (-4.60)	-1.80*** (-7.86)
High-Low Timing	-1.35** (-2.03)	-1.24** (-2.10)	-1.13*** (-2.35)	-0.98** (-2.41)	-0.76*** (-3.30)

This table reports the mid- and long-term effect of factor timing on raw returns and risk adjusted abnormal returns. Each week funds are sorted into five quintile portfolios by the aggregated Timing Indicator. Panel A reports the portfolios' cumulative raw return over the following 1 week, 1 month, 6 month, 1 year or 3 year time horizon. Panel B reports the portfolios' cumulative abnormal returns. Weekly abnormal returns are the difference between actual returns and the expected returns as estimated by a Carhart model. The bottom row of each panel reports the difference between the most and the least actively timing funds. T-statistics are reported in parentheses. We use Newey-West standard errors with a lag length of 156 weeks in panel B. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

Table 6: Determinants of factor timing activity

	τ_{mkt}	τ_{smb}	τ_{hml}	τ_{mom}	Timing Indicator
ln(tna)	-0.000 (-0.59)	0.000 (0.31)	-0.002 (-0.57)	0.000 (0.44)	0.003 (0.27)
ln(fund age)	0.001 (1.52)	-0.002 (-0.53)	0.003 (0.58)	-0.000 (-0.27)	0.015 (0.77)
ln(manager tenure)	0.002*** (3.52)	0.012*** (3.34)	0.014*** (2.74)	0.004*** (4.15)	0.065*** (4.29)
Expenses (in %)	0.017*** (9.03)	0.088*** (9.72)	0.084*** (8.26)	0.022*** (5.92)	0.462*** (10.35)
Turnover ratio	0.004*** (3.85)	0.013* (1.80)	0.020** (2.08)	0.006*** (2.43)	0.086*** (3.40)
Past Alpha	0.013 (0.41)	0.050 (0.51)	0.081 (1.54)	0.050** (2.10)	0.454 (1.08)
Fund Flows	0.001 (0.64)	-0.001 (-0.37)	0.019 (1.64)	0.001 (0.63)	0.036* (1.80)
Style dummy variables					
Growth and Income	–	–	–	–	–
Growth	0.009*** 6.42)	0.040*** (5.42)	0.051*** (4.00)	0.012*** (4.28)	0.221*** (8.47)
Hedged	0.003 (0.62)	0.051*** (3.07)	0.070 (1.05)	-0.002 (-0.53)	0.165 (1.84)
Income	-0.001 (-0.55)	-0.011 (-0.99)	0.009 (0.54)	-0.000 (-0.09)	-0.034 (-0.85)
Micro	0.028*** (6.28)	0.116*** (4.50)	0.110*** (4.37)	0.039*** (6.08)	0.691*** (7.09)
Mid	0.018*** (3.04)	0.077*** (7.15)	0.091*** (2.62)	0.027*** (3.79)	0.470*** (6.72)
Small	0.014*** (4.90)	0.063*** (5.44)	0.093** (2.25)	0.024*** (4.25)	0.422*** (8.46)
R^2	0.16	0.13	0.20	0.27	0.20

This table reports the fund characteristics that determine future factor timing activity. We split our sample into non-overlapping 3-year subperiods, namely 1999-2001, 2002-2004, etc. up to 2014-2016. We regress the fund characteristics at the beginning of these periods on the timing measures estimated from the dynamic factor model during those periods. Standard errors are double clustered on fund level and time period. Z-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

Table 7: Factor timing transition matrix

Panel A: 1-year attrition rate and transition matrix

Current Decile	Mean initial / final Timing Indicator	1	2	3	4	5	6	7	8	9	10	Attrition Rate
1	-0.68 / -0.64	73.5	19.3	4.5	1.5	0.6	0.2	0.1	0.0	0.0	0.0	5.6
2	-0.54 /-0.50	19.1	45.0	23.6	8.0	2.9	1.0	0.3	0.1	0.0	0.0	6.5
3	-0.43 / -0.40	4.3	23.2	35.5	22.1	9.4	3.7	1.2	0.4	0.2	0.1	6.9
4	-0.33 / -0.31	1.3	8.2	22.4	31.6	21.3	9.6	3.7	1.2	0.5	0.1	6.7
5	-0.22 / -0.20	0.5	2.9	9.6	22.1	29.8	20.6	9.4	3.6	1.1	0.3	7.0
6	-0.10 / -0.08	0.3	0.9	3.4	9.8	22.4	29.1	21.1	9.4	2.9	0.7	6.5
7	0.06 / 0.06	0.2	0.4	1.4	3.7	10.1	21.0	30.5	22.7	8.4	1.6	6.9
8	0.25 / 0.24	0.1	0.1	0.5	1.1	3.3	9.9	23.9	33.6	22.8	4.7	7.3
9	0.55 / 0.49	0.1	0.0	0.1	0.3	1.1	3.1	8.9	24.0	43.2	19.2	8.1
10	1.38 / 1.13	0.1	0.0	0.1	0.1	0.3	0.6	1.7	5.4	21.6	70.3	15.1

Panel B: 3-year attrition rate and transition matrix

Current Decile	Mean initial / final Timing Indicator	1	2	3	4	5	6	7	8	9	10	Attrition Rate
1	-0.68 / -0.56	52.6	20.7	11.2	6.4	3.8	2.3	1.4	0.8	0.3	0.5	15.3
2	-0.54 /-0.42	20.5	26.6	21.0	13.5	7.6	4.8	3.2	1.7	0.7	0.4	17.7
3	-0.43 / -0.33	11.0	20.5	20.9	16.6	12.0	7.4	5.7	3.2	1.8	0.9	18.9
4	-0.33 / -0.25	5.4	13.2	17.6	18.4	16.1	11.4	8.4	5.1	3.2	1.3	17.8
5	-0.22 / -0.15	2.6	8.7	12.7	15.3	17.2	16.0	11.7	8.5	5.1	2.3	18.8
6	-0.10 / -0.05	1.7	5.5	8.9	11.9	15.6	17.0	15.1	13.1	7.6	3.8	19.1
7	0.06 / 0.08	1.2	3.4	5.8	8.3	12.0	14.9	16.6	17.5	13.2	7.1	19.2
8	0.25 / 0.21	0.7	1.6	3.2	5.9	8.8	13.0	17.8	19.3	18.5	11.3	19.6
9	0.55 / 0.40	0.5	0.9	1.5	3.2	4.8	8.3	14.7	19.3	25.8	21.0	20.1
10	1.38 / 0.81	0.3	0.3	0.8	1.4	2.2	4.0	7.7	12.9	22.9	47.4	26.8

This table displays the transitions of mutual funds between deciles sorted on the aggregated Timing Indicator over a period of 1 year (panel A) and 3 years (panel B). Each week we sort funds by their aggregated Timing Indicator which is calculated from the past 156 weeks using the dynamic factor model. The first column reports the average Timing Indicator of funds within each decile upon decile formation as well as 1 or 3 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 3 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 1 year or 3 years later.

Table 8: Impact of Timing Indicators on risk characteristics

Panel A: Independent variables: Timing activity measures				
	β_{CAPM}	$\sigma(r_{weekly})$	$\sigma(\varepsilon)_{OLS}$	$\sigma(\varepsilon)_{DFM}$
τ_{mkt}	0.327*** (3.49)	0.0156*** (5.46)	0.0229*** (8.78)	$4.36 \cdot 10^{-25}$*** (4.03)
τ_{smb}	0.116*** (2.60)	0.0054*** (3.03)	0.0059*** (8.31)	$0.99 \cdot 10^{-25}$ (1.26)
τ_{hml}	0.094*** (3.10)	0.0055*** (3.15)	0.0054*** (4.35)	$1.07 \cdot 10^{-25}$** (2.07)
τ_{mom}	0.372*** (3.08)	0.0139*** (3.07)	0.0151*** (5.83)	$3.28 \cdot 10^{-25}$*** (2.78)
Style Dummies	Yes	Yes	Yes	Yes
Average R^2	0.32	0.45	0.84	0.05

Panel B: Independent variables: 3 year lagged timing activity measures				
	β_{CAPM}	$\sigma(r_{weekly})$	$\sigma(\varepsilon)_{OLS}$	$\sigma(\varepsilon)_{DFM}$
τ_{mkt}	0.111 (1.38)	0.0067*** (3.77)	0.0171*** (7.44)	$7.94 \cdot 10^{-26}$** (2.46)
τ_{smb}	0.074*** (3.44)	0.0029*** (5.38)	0.0040*** (10.01)	$3.32 \cdot 10^{-26}$** (2.05)
τ_{hml}	0.072*** (2.74)	0.0028*** (3.24)	0.0036*** (5.12)	$1.84 \cdot 10^{-26}$ (1.14)
τ_{mom}	0.226*** (4.95)	0.0077*** (8.44)	0.0090*** (17.04)	$0.00 \cdot 10^{-26}$ (0.00)
Style Dummies	Yes	Yes	Yes	Yes
Average R^2	0.31	0.42	0.58	0.04

This table displays the effect of market timing on the risk characteristics of fund returns. Four different risk indicators, namely the market risk factor loading from a one-factor model (β_{CAPM}), the weekly return volatility $\sigma(r_{weekly})$, and two versions of the standard deviation of the residuals from a Carhart factor model regression (representing idiosyncratic risk) are regressed on measures of factor timing activity using a Fama-MacBeth regression. The market, size, value and momentum factor timing measures are calculated from the dynamic factor model over a 3 year period as presented in section 2. The risk indicators are obtained from a period of 3 years. Two versions of residual volatility are calculated from OLS regression ($\sigma(\varepsilon)_{OLS}$) and the dynamic factor model ($\sigma(\varepsilon)_{DFM}$), respectively. The panel A results are obtained from Fama-MacBeth regressions where the factor timing measures are calculated from the same 3-year period the risk indicators are estimated from. Panel B results are obtained from Fama-MacBeth regressions where the factor timing measures are calculated from the 3 years prior to the time period over which the risk indicators are estimated. T-statistics are calculated with Newey–West standard errors (max. lag=156) and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.

Table 9: Influence of factor timing on the Amihud/Goyenko R^2-measure and TE

	R^2	R^2	TE	TE
τ_{mkt}	-0.407*** (-4.17)		0.148*** (7.20)	
τ_{smb}	-0.092*** (-7.10)		0.045*** (5.02)	
τ_{hml}	-0.064*** (-4.60)		0.019*** (6.70)	
τ_{mom}	-0.137** (-3.22)		0.107** (3.25)	
Timing Indicator		-0.049*** (-11.39)		0.022*** (5.61)
Time FE	Yes	Yes	Yes	Yes
R^2	0.37	0.37	0.44	0.44

This table reports the results of a regression of timing measures on the R^2 obtained from a regression of fund returns on a Carhart four-factor model as proposed by Amihud/Goyenko (2013) and on the Tracking Error (TE), i.e. the volatility of a funds' excess return over the self-declared benchmark index. We split our sample into non-overlapping 3-year subperiods, namely 1999-2001, 2002-2004, etc. up to 2014-2016. For each subperiod we calculate the timing measures from the dynamic factor model and we obtain the R^2 from an OLS regression on a Carhart model. TE is calculated as the standard deviation of weekly returns.
We then regress the values of R^2 and TE on the obtained factor timing measures. T-statistics are calculated from time-clustered standard errors and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***

Table 10: The effect of factor timing and R^2 on future fund returns: Fama-MacBeth-regressions

Explanatory variables	Dependent variable: $\alpha_{j,t}$ (in %)					
τ_{mkt}	-0.067** (-1.97)		-0.108*** (-2.65)		-0.096** (-2.41)	
τ_{smb}	-0.024** (-2.21)		-0.032*** (-2.74)		-0.032*** (-2.62)	
τ_{hml}	0.005 (0.77)		-0.006 (-0.63)		-0.006 (-0.62)	
τ_{mom}	-0.046* (-1.69)		-0.052* (-1.73)		-0.049 (-1.64)	
Timing Indicator		-0.797** (-2.33)		-1.286*** (-2.58)		-1.211** (-2.54)
TE	-0.027 (-1.55)	-0.028 (-1.55)			-0.030 (-1.59)	-0.031 (-1.61)
R^2			-7.132* (-1.88)	-7.035* (-1.85)	-7.328* (-1.89)	-7.307* (-1.88)
Controlls	YES	YES	YES	YES	YES	YES
Style Dummies	YES	YES	YES	YES	YES	YES
Average R^2	0.14	0.13	0.15	0.14	0.16	0.14

This table reports the result of Fama-MacBeth regressions of annualized abnormal fund returns on timing measures, on the R^2 obtained from a regression of fund returns on a factor model as proposed by Amihud/Goyenko (2013) and on the Tracking Error (TE), i.e. the volatility of a funds' excess return over the self-declared benchmark index, and controlls. Abnormal returns are excess returns over the expected returns calculated from a Carhart model. R^2 is calculated from an OLS regression of fund returns on a Carhart model during the previous 3 years. Control variables include all fund characteristics as in table 3. T-statistics are calculated with Newey–West standard errors (max. lag=12) and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by *, **, and ***.