MARKET AND STYLE TIMING: GERMAN EQUITY AND BOND FUNDS

This version: 22nd July 2015

Abstract:

We apply parametric and non-parametric estimates to test market and style timing ability of individual German equity and bond mutual funds using a sample of over 500 equity and 350 bond funds, over the period 1990-2009. For equity funds, both approaches indicate no successful market timers in the 1990-1999 or 2000-2009 periods, but in 2000-2009 the non-parametric approach gives fewer unsuccessful market timers than the parametric approach. There is evidence of successful style timing using the parametric approach, and unsuccessful style timing, particularly in the 2000-2009 period. There is evidence of positive and negative bond timing in the 2000-09 period.

Keywords : Mutual funds performance, market timing.

JEL Classification: C14, G11

MARKET AND STYLE TIMING: GERMAN EQUITY AND BOND FUNDS

1. Introduction

After the US, UK, Japan and France, Germany is the 5th largest asset management center in the world. Mutual fund investments in Germany account for around $335 billion under management. With ongoing political and financial restructuring it is expected that individuals will have to become increasingly responsible for future long-term pension savings. Therefore it is expected that the mutual fund industry will grow rapidly over the medium term as reforms to private pension provision place greater emphasis on defined contribution pensions (i.e. 'Riester Rente') and reforms result in a less generous state pension. As in other countries such as the US and UK, mutual fund assets are predominantly held in active funds – this paper examines whether active German equity and bond funds engage in successful market and style timing.

Mutual fund performance is usually discussed in terms of selectivity (alpha) and timing and is analysed using either returns data or (where available) portfolio holdings data. Returns-based studies may be further subdivided into parametric and non-parametric approaches. To model timing effects in the parametric approach, a factor model is augmented with additional non-linear functions of the factors (Treynor and Mazuy 1966 and Henriksson and Merton 1981). Parametric models of timing may be unconditional or conditional on publicly available information, which allows for time-varying alphas and factor loadings (Ferson and Schadt 1996, Christopherson, Ferson and Glassman 1998).

The parametric approach measures both the response to the timing signal and the strength of that response (in terms of the size of the change in beta). The non-parametric returns-based approach provides a measure of the *quality* of the manager's forecast, independent of the *aggressiveness* of the response due to changing factor loadings[1].

In this study we use a large (survivorship-bias free) sample of over 500 equity and 350 bond funds, over the last 20 years (1990-2009). The key contributions of the paper are as follows. First we use both parametric and non-parametric approaches (Jiang 2003) and test for both unconditional and conditional market timing – this to our knowledge has not been done for German funds and provides complementary evidence to the literature on US and UK data – which itself is mainly based on parametric approaches (Treynor-Mazuy, TM 1966, Henriksson-Merton, HM 1981, Ferson and Schadt 1996, Christopherson et al 1998). Second, for the first time, we examine style timing for German equity funds – that is, do managers forecast the future path of small fund

[1] Studies of timing that use holdings data avoid some of the potential biases in parametric factor timing models due to interim trading and passive timing (Jiang, Yao and Yu 2007, Elton, Gruber and Blake 2012, Huang and Wang 2014).

returns relative to large fund returns ("size timing") or returns on high book-to-market relative to low book-to-market firms ("growth timing") and successfully alter their weighting on these factors, to enhance future fund returns[2]. Third, we examine the timing skills of German equity funds with domestic, European and Global mandates – thus providing evidence on the 'home-bias' issue (Coval and Moskowitz 1999, Hong, Kubik and Stein 2005). Finally, we examine the market timing ability of bond funds using parametric and non-parametric models – to the best of our knowledge the latter has not previously been attempted for bond funds and certainly not for German bond funds[3]. Given the paucity of empirical work on German mutual funds this substantially enhances our knowledge of the performance of a large and growing industry in both domestic and foreign markets.

The key results of the paper are as follows. Using a non-parametric measure we find both fewer successful and fewer unsuccessful equity market timers than for the parametric method but overall, both methods give few successful market timers and a larger number of funds that are negative market timers. On *style timing* both approaches indicate that a substantial number of German equity funds with European or Global mandates are unsuccessful timers of "size" and "growth" factors in the later period 2000-09. This suggests that the rapid growth in these international equity funds may have resulted in managers having poor ability in forecasting markets with which they are less familiar.

Overall our non-parametric results suggest that there are few if any equity funds which are successful market or style timers but there is stronger evidence of unsuccessful market and style timers – particularly for European and Global mandates. For bond funds, our preliminary analysis shows a substantial proportion of both positive and negative market timers in the 2000-09 period.

The rest of the paper is organised as follows. Section 2 describes the parametric and non-parametric testing methodologies. In section 3 we discuss previous empirical studies and in section 4 we describe the German fund data set and our empirical results. Section 5 concludes.

2. Parametric and Non-parametric Tests

Our baseline model is the Fama-French three factor (3F) model used on German domestic equity funds by Bessler et al (2009), which we augment with the market timing variables of Treynor and Mazuy (1966), TM and Henriksson and Merton (1981), HM. The 3F+TM model is:

(1) $$r_{t+1} = \alpha + \beta_m r_{m,t+1} + \beta_{SMB} R_{SMB,t+1} + \beta_{HML} R_{HML,t+1} + \gamma_m r^2_{m,t+1} + \varepsilon_{t+1}$$

2 Although we refer to fund managers it is the performance of funds that we examine.

3 The potential role for conditioning information in the predictability of German aggregate bond and stock indexes has been established by Hyde and Kappel (2010) – although no specific combination of variables dominates over different sample periods.

r_t is the fund's excess return, r_m is the market excess return, R_{SMB} and R_{HML} are returns on zero-investment portfolios of high market cap firms minus low cap firms and high book-to-market minus low book-to-market firms, respectively. If β_m is positively related to the forecast market return then the coefficient γ_m > 0 (γ_m < 0) measures successful (unsuccessful) market timing ability and the null hypothesis of no market timing is γ_m = 0. In the HM model the conditional portfolio beta follows a binary response function depending on the manager's forecast of whether next period's market return will exceed the risk free rate and the market timing variable becomes $\gamma_m r^+_{m,t+1}$, where $r^+_{m,t+1}$ = max$\{r^+_{m,t+1}, 0\}$. Using a similar argument, for timing ability with other style factors the TM approach gives:

(2) $$r_{t+1} = \alpha + \beta_m r_{m,t+1} + \beta_{SMB} R_{SMB,t+1} + \beta_{HML} R_{HML,t+1} + \gamma_m r^2_{m,t+1} + \gamma_{SMB} R^2_{SMB,t+1} + \gamma_{HML} R^2_{HML,t+1} + \varepsilon_{t+1}$$

Successful "size timing" occurs when the fund manager increases β_{SMB} based on a forecast increase in the return to small capitalised firms relative to large capitalised firms, resulting in γ_{SMB} > 0. Growth (value) stocks tend to be stocks with lower (higher) than average book-to-market value, B/M. Hence a positive γ_{HML} captures timing on the basis of forecasts of the book-to-market ratio for value relative to growth firms – that is, "growth timing"[4].

Non-Parametric Approach

Jiang (2003) proposes a non-parametric test (initially applied to US mutual funds), which we outline using the market model:

(3) $$r_{t+1} = \alpha + \beta_t r_{m,t+1} + \varepsilon_{t+1}$$

The fund's beta β_t is assumed to vary with the fund manager's market timing information. The fund's timing skill is determined by the ability to correctly predict market movements. Let

---

4 When hypothesis testing, each bootstrap is based on generating fund returns $\tilde{R}$ under a specific null (e.g. γ_m =0), then re-running the regression with the "generated" fund returns $\tilde{R}$, to obtain an estimate of $\tilde{\gamma}_m$ and (Newey-West) $t(\tilde{\gamma}_m)$. This is repeated m-times, to obtain the distribution of $t(\tilde{\gamma}_m)$ under the null. The empirical value for $t(\hat{\gamma}_m)$ is then compared with the tail of the null distribution (to give the p-value of the test). The "generated" fund returns $\tilde{R}$ can be derived using a) "basic bootstrap" where we simply bootstrap the residuals (only) or b) a "factor bootstrap" where we bootstrap both the factors and the residuals c) a block bootstrap (Ledoit and Wolf 2008) d) a contemporaneous bootstrap (Kosowski et al 2006, Fama and French 2010). We find no qualitative difference in results across these alternatives.

$\hat{r}_{m,t+1} = E(r_{m,t+1} \mid I_t)$ be the manager's forecast for next period's market return based on information I_t. Define ν_t as:

(4) $$\nu_t = \Pr(\hat{r}_{m,t_2+1} > \hat{r}_{m,t_1+1} \mid r_{m,t_2+1} > r_{m,t_1+1}) - \Pr(\hat{r}_{m,t_2+1} < \hat{r}_{m,t_1+1} \mid r_{m,t_2+1} > r_{m,t_1+1})$$

Under the null hypothesis of no market timing ability $\nu = 0$, since the probability ($\Pr$) of a correct forecast equals the probability of an incorrect forecast. $\nu \in [-1,1]$ where the two extreme values represent perfect negative and perfect positive (i.e. successful) market timing respectively. Equation (4) may be written as:

(5) $$\nu_t = 2 \times \Pr(\hat{r}_{m,t_2+1} > \hat{r}_{m,t_1+1} \mid r_{m,t_2+1} > r_{m,t_1+1}) - 1$$

The next step is to link the manager's forecast of the market return with their response in adjusting β_t in (3). For any triplet of market return observations $\{r_{m,t_1}, r_{m,t_2}, r_{m,t_3}\}$ sampled from *any* three time periods (not necessarily in consecutive order) with $\{r_{m,t_1} < r_{m,t_2} < r_{m,t_3}\}$, an informed market timer will maintain a higher exposure to the market over the $[r_{m,t_2}, r_{m,t_3}]$ range than in the $[r_{m,t_1}, r_{m,t_2}]$ range. Non-parametric beta estimates for both time ranges are $\beta_{t_1} = (r_{t_2} - r_{t_1}) / (r_{m,t_2} - r_{m,t_1})$ and $\beta_{t_2} = (r_{t_3} - r_{t_2}) / (r_{m,t_3} - r_{m,t_2})$. Grinblatt and Titman (1989) show that for a fund with non-increasing absolute risk aversion and independent timing and selectivity information then $\partial \beta_t / \partial \hat{r}_{m,t+1} > 0$, yielding a convex fund-return, market-return relationship:

(6) $$\frac{r_{t_3} - r_{t_2}}{r_{m,t_3} - r_{m,t_2}} > \frac{r_{t_2} - r_{t_1}}{r_{m,t_2} - r_{m,t_1}}$$

which allows (5) to be written as $\nu = 2 \times \Pr(\beta_{t_2} > \beta_{t_1} \mid r_{m,t_2+1} > r_{m,t_1+1}) - 1$. A sample statistic of a fund's timing ability may be constructed as:

(7) $$\hat{\theta}_n = \binom{n}{3}^{-1} \sum_{r_{m,t_1} < r_{m,t_2} < r_{m,t_3}} \operatorname{sign}\left(\frac{r_{t_3} - r_{t_2}}{r_{m,t_3} - r_{m,t_2}} > \frac{r_{t_2} - r_{t_1}}{r_{m,t_2} - r_{m,t_1}} \right)$$

where sign (·) = (1, -1, 0) for positive, negative and zero market timing respectively. $\hat{\theta}_n$ is the average sign across all triplets taken from *n* observations. $\hat{\theta}_n$ can be shown to be √n-consistent and asymptotically normal (Abrevaya and Jiang 2005, Serfling 1980) with variance:

(8) $$\hat{\sigma}^2_{\hat{\theta}_n} = \frac{9}{n}\sum_{t_1=1}^{n}\left(\binom{n}{2}^{-1}\sum_{t_2,t_3} h(z_{t_1}, z_{t_2}, z_{t_3}) - \hat{\theta}_n\right)^2$$

where

(9) $$h(z_{t_1}, z_{t_2}, z_{t_3}) = \text{sign}\left(\frac{r_{t_3} - r_{t_2}}{r_{m,t_3} - r_{m,t_2}} > \frac{r_{t_2} - r_{t_1}}{r_{m,t_2} - r_{m,t_1}} \mid r_{m,t_1} < r_{m,t_2} < r_{m,t_3}\right)$$

Under the null hypothesis of no market timing $z = \sqrt{n}.\hat{\theta}_n \big/ \hat{\sigma}_{\hat{\theta}_n}$ is asymptotically N(0,1) distributed.

One difficulty in examining a fund's market timing skill is distinguishing the *quality* of the manager's forecast of the future market return from the *aggressiveness* of response in changing the fund's beta. The parametric (TM and HM) market timing measures do not separate out these two elements. The parametric approach measures both the response to the timing signal and the strength of that response (in terms of the size of the change in beta). The non-parametric statistic θ measures the *proportionate number of times* the fund's beta is higher, in a high return period than in a low return period. Hence θ measures only the response to the timing signal and is independent of the aggressiveness of the response. This is because the sign function in (7) assigns a value of 1(-1) if the argument is positive (negative), regardless of the size of the argument[5]. Hence one advantage of the non-parametric procedure over the parametric (regression) approach is that it is based on the quality of a fund manager's timing information rather than the aggressiveness of her response.

The non-parametric test embodies some relatively mild restrictions on behaviour. The test requires $\beta_{m,t}$ be a non-decreasing function of $\hat{r}_{m,t+1}$. This is less restrictive than that of the TM and HM measures which require specific linear and binary response functions respectively[6].

However, the non-parametric and parametric (TM/HM) methods both share potential "problems". For example, both the parametric (TM/HM) and non-parametric methods i) cannot distinguish market timing from option-related spurious timing (Jagannathan and Korajczyk 1986, Jiang, Yao and Yu 2007), ii) require security selection to be independent of information on timing and iii) may be subject to interim trading bias – for example, when daily timing takes place but the empirical data frequency is monthly[7] (Goetzmann, Ingersoll and Ivkovich 2000). Overall, the two

---

5 Reinforcing the above point, simulations show that the non-parametric θ correctly measures differences in information quality but is invariant across different response intensities (aggressiveness) - whereas the parametric γ mainly reflects the manager's aggressiveness (Jiang 2003).

6 Such an assumption is questionable if there is non-linearity in the payment to fund managers in respect of benchmark evaluation (Admati and Pfleiderer 1997), option compensation (Carpenter 2000) and a non-linear performance-flow responses by investors (Chevalier and Ellison 1997).

7 Although both methods suffer from interim trading bias, the non-parametric method results in much less bias than the parametric methods, when timing is daily but monthly data is used in the tests (Jiang 2003).

methods measure slightly different aspects of timing, both have their strengths and weaknesses and both need to be used in empirical work.

Conditional Market Timing

The non-parametric test can be applied as a conditional statistic after allowing for market timing skill attributable to public information (Ferson and Schadt 1996). The null is then a test of the quality of the fund manager's *private* timing signal[8] and is referred to as conditional timing.

This conditional measure involves first calculating both sets of residuals from regressions of the mutual fund returns and market returns on the lagged public information variables. Clearly, these residuals represent the variation in the fund and market returns not explained by the public information. Denoting the pair-wise fund and market regression residuals as $\tilde{r}_t$ and $\tilde{r}_{m,t}$ respectively, the procedure described above may then be applied to the residuals to yield a conditional timing measure:

$$(10) \qquad \tilde{\theta}_n = \binom{n}{3}^{-1} \sum_{\tilde{r}_{m,t_1} < \tilde{r}_{m,t_2} < \tilde{r}_{m,t_3}} \text{sign}\left(\frac{\tilde{r}_{t_3} - \tilde{r}_{t_2}}{\tilde{r}_{m,t_3} - \tilde{r}_{m,t_2}} > \frac{\tilde{r}_{t_2} - \tilde{r}_{t_1}}{\tilde{r}_{m,t_2} - \tilde{r}_{m,t_1}} \right)$$

Note, $\hat{\theta}_n$ in (7) and $\tilde{\theta}_n$ in (10) can clearly be of different magnitudes but may also be of different sign. For example, $\hat{\theta}_n > 0$ but $\tilde{\theta}_n \le 0$ may indicate a successful market timing manager whose skill is attributable entirely to public information.

3. Previous Studies

For domestic equity funds, most US and UK studies using the TM and HM parametric approach find weak evidence of positive market timing and somewhat stronger evidence of negative market timing[9]. Swinkels and Tjong-A-Tjoe (2007) and Chen, Adams and Taffler (2013) also consider style timing variables on US equity funds using only a parametric approach. Swinkels and Tjong-A-Tjoe (2007) find successful timing of the market, growth timing and momentum timing - although they do not test all timing effects simultaneously. Chen, Adams and Taffler (2013) use the parametric approach utilising all style timing variables but only on a subset of "US superior performing growth funds" and find these predominantly exhibit growth timing skills and other style timing effects are largely absent.

---

[8] See also Becker et al (1999) and Ferson and Khang (2002) for further discussion of the effects of conditioning information on timing measures. Portfolio managers may also adjust a fund's exposure to risk factors other than the market or indeed to other benchmark indices according to their year-to-date performance in response to incentives they may face (Chevalier and Ellison 1997, Brown, Harlow and Starks 1996).

[9] For the US see for example, Treynor and Mazuy 1966, Henriksson and Merton 1981, Hendriksson 1984, Lee and Rahman 1990, Ferson and Schadt 1996, Busse 1999, Becker, Ferson, Myers and Schill 1999, Wermers 2000, Bollen and Busse 2001, Aragon 2005, Glassman and Riddick 2006,Swinkels and Tjong-A-Tjoe 2007, Jiang, Yao and Yu, 2007, Chen and Liang 2007. For the UK see Chen, Lee, Rahman and Chan 1992, Fletcher 1995, Leger 1997, Byrne, Fletcher and Ntozi 2006, Cuthbertson et al 2010).

Jiang, Yao and Yu (2007) construct “bottom up” (value weighted) market betas for each US domestic equity fund based on its holdings of particular stocks at the end of each quarter. A time series for the “bottom up” fund beta is then regressed on future market returns (over 1, 3, 6 and 12 months) to provide an estimate of market timing ability over these selected horizons. In contrast to the returns based approach, when using the holdings-based approach they find some evidence of statistically significant positive market timing ability (particularly for aggressive growth and growth objectives) and little evidence of negative timing[10]. These differences they attribute to the increased power and less artificial timing bias of the holdings approach.

Results on the timing ability of US bond funds are mixed, depending on the section of the fixed income market considered and the methodology used. For investment grade bonds, Boney, Comer and Kelly (2009) using a return attribution approach (Sharpe 1992) find negative timing between cash and bonds and across the bond maturity spectrum. Ammann et al (2010) find positive timing ability for convertible bond funds. Chen et al (2013) find around 4% of US bond funds have statistically significant positive or negative timing effects across various systematic factors related to bond returns[11] - and overall the distribution of timing effects is neutral to weakly positive.

Huang and Wang (2014) attempt to isolate timing effects in a homogenous sub-set of bonds, namely US Treasury bond funds. They use portfolio holdings on 146 US Treasury bond funds and find evidence of positive (unconditional) timing skill - but this disappears if conditioning public information is used. In one of the few studies of US market timing for hybrid (equity) funds which contain substantial holdings of bonds, Comer (2006) finds only 4 out of 114 hybrid funds have significant *bond* timing ability, when using a parametric model. Overall, the evidence on US bond funds suggests little positive timing skill[12].

Studies of the German mutual fund industry are rather sparse. Most consider only selectivity, use only a small number of funds and few investigate market and style timing - see for example, Krahnen et al (2006) who use 13 funds (1987-1998), Stehle and Grewe (2001) use 18 funds (1973 to 1998), Griese and Kempf (2003) use 105 funds (1980 to 2000) while Otten and Bams (2002), use 57 funds (1991-1998)[13]. The most comprehensive study of German equity funds is that of Bessler, Drobetz and Zimmermann (2009) who use both conditional and unconditional parametric factor models (CAPM and the 3 factor Fama-French model) as well as a parametric SDF model, on 50 German domestic equity funds (1994-2003). They find the 3-factor

---

10 Kaplan and Sensoy (2010) using US holdings data, find some evidence of positive timing with respect to benchmark betas.

11 The factors include three term structure variables (short rate, slope and curvature), mortgage spread, credit and liquidity spreads, exchange rates and two equity market factors.

12 Although selectivity is not the focus of this paper, US studies of bond funds across various sectors tend find predominantly negative alpha performance after deduction of management fees but some evidence of statistically significant positive and negative persistence. See inter alia, Cornell and Green 1991, Blake et al 1993, , Gruber and Blake 1995, Ferson, Henry and Kisgen 2006, Comer and Rodriguez 2006, Huij and Derwall 2008, Du et al 2009, Boney and Comer 2010, Chen et al 2013, Ammann et al 2010. For European bond performance see Silva et al 2005 and for Canada, Ayadi and Kryzanowski 2011.

13 In addition, Banegas, Gillen, Timmermann and Wermers, BGTW (2008) examine *portfolios* of European domiciled funds.

model and SDF approach "deliver closely related performance measures" with virtually zero positive alpha funds and about 4-6 statistically significant negative alpha funds.

Cuthbertson and Nitzsche (2013) examine the "total performance" (selectivity plus market timing) of around 500 German equity funds (1990-2009) using the returns-based timing model and after correcting for false discoveries. They find no funds with positive performance but a considerable proportion with negative total performance. Using the Fama-French returns-based timing model for 129 German equity funds (1989-2005), Stotz (2007) finds little or no evidence of statistically significant timing effects for the market return and for size, book-to-market and momentum variables. In this study we examine market and style timing using both returns-based and non-parametric timing measures, for German equity and bond funds.

4. Data and Empirical Results

We use monthly mutual fund returns (from Bloomberg) on all recorded German domiciled equity and bond mutual funds between 1990 and 2009. The data set includes both surviving and 'dead' funds and in our analysis we use 555 equity funds and 389 bond funds which have a data history of at least two years (to minimise look-ahead bias).

For equity funds our baseline model is the 3-factor, Fama-French model. Our equity funds have German, European and Global geographic mandates and for the market return we have used the appropriate MSCI total return indices (including dividends) for each geographical region. The SMB variables have been calculated by subtracting the total return index of the small cap MSCI index from the relevant market index for the specific geographic mandate. Similarly, HML is defined as the difference between the total return indices of the MSCI value index less the MSCI growth index for the specific geographic region[14]. The risk-free rate is the 1-month Frankfurt money market rate. All variables are measured in Euros (or German Marks prior to the introduction of the Euro). Fund returns are net of management fees, expenses and brokerage commissions (but before any front-end and back-end loads) and are therefore returns to the investor (ignoring any personal tax implications).

For bond funds, an examination of their prospectuses reveals investments primarily in government and corporate investment grade bonds, with some assets also held in high yield bonds, mortgage bonds and asset backed securities - across the US, Europe and globally. As there is no consensus in the literature on an appropriate bond factor model we considered a variety of indices and we report results for a four factor model consisting of two widely used bond indices compiled by Citigroup and two indices from Bank of America/Merrill Lynch. The indices

[14] Use of the MSCI indices allows consistency across factor definitions for "German", "European" and "Global" mandates. Worldscope has greater coverage for our factors but only for funds with a German mandate. Worldscope aims to cover 95% of market capitalization and MSCI indices target 85% of free-floated market capitalisation. Reneeboog, Horst and Zhang (2004) report little change in results when using Worldscope rather than other data sources. Comer and Rodriguez (2011) use this MSCI 3-factor model for US funds that invest internationally, which is extended in Comer and Rodriguez (2012). Huij and Derwall (2011) for US global funds find little difference between the fit of the MSCI 3F-model and alternative models using sector or country indices.

are the US Overall Broad Investment Grade Index and the European Union Government Bond Market Index (from Citigroup). We also include a High Yield Index and a Global Government Bond index (from Bank of America/Merrill Lynch)[15].

Empirical Results

Table 1 shows summary statistics for our sample of equity funds for 1990-99 (Panel A) and 2000-09 (Panel B) for German domiciled funds that have German, European and Global mandates. The salient features are an increase in the total number of equity funds over the two periods from 195 to 544 mainly due to an increase in funds with European or Global mandates - from around 60 equity funds to 220, for each mandate. In comparison the number of equity funds with a German mandate hardly increases at all between the two periods. The later period covers the large stock market declines of 2000-02 and 2008-09. As expected, average equity returns are lower in the 2000-09 period but with standard deviations largely unchanged between the two periods.

[Table 1 here]

Consistent with earlier results on German equity funds, the funds closely track the market index but with a consistent positive weighting towards small stocks across all three mandates and in both sample periods. There is some evidence of a tilt towards growth stocks for European and Global mandates but not for the domestic mandate. Most funds neither under or outperform their factor benchmarks – there are very few statistically significant positive or negative alphas. The average $\bar{R}^2$ for the 3F model are in the range 0.67-0.85 and there is non-normality in many fund residuals – hence we bootstrap statistical tests in our parametric models.

[Table 2 here]

Before assessing the timing skills of individual funds, table 2 presents results for (equally weighted) portfolios of funds with German, European and Global mandates respectively for the 3F model plus market/style timing variables ($r_m^2, R_{SMB}^2, R_{HML}^2$) – that is, the "three factor style timing" (3F+3ST) model. The market betas β_m , size betas, β_{SMB} and growth betas β_{HML} are generally statistically significant. Statistically significant negative market timing effects (γ_m< 0) are found for funds with a European mandate in 1990-1999 and for the Global mandate in 2000-2009. There are two positive statistically significant growth timing effects ($\gamma_{HML} > 0$) for the Global mandate in each period.

---

15 The US Overall Broad Investment Grade Bond Index comprises US Treasuries, government sponsored, mortgages, asset backed as well as investment grade securities with an S&P rating of at least BBB-. The Citigroup European Government Bond Index comprises government bonds issued by European Union countries with a rating of at least BBB-. The High Yield Index comprises US denominated, US issued fixed income securities rated below investment grade. The Global index consists of investment grade bonds issued by OECD countries. An index of mortgage bonds ("US Mortgage Bond Index" from Bank of America/ML) has a correlation of 0.99 with the US Broad Investment Grade Index and is therefore excluded.

4.1. Equity Funds: Parametric Models

On a fund-by-fund basis, we first discuss market timing and style timing of equity funds using parametric models. We present results for the 3F model plus market and style timing variables ($r_m^2, R_{SMB}^2, R_{HML}^2$) – that is, the "three factor style timing" (3F+3ST) model[16]. Table 3 shows the percentage (number) of positive and negative (statistically significant) market and style timing coefficients ($\gamma_m, \gamma_{SMB}, \gamma_{HML}$) in the 1990-99 (Panel A) and 2000-09 (Panel B) periods for equity funds with German, European and Global mandates (using a one-tail test at a 2.5% significance level, with bootstrap t-statistics based on Newey-West adjusted standard errors). To assess the statistical significance of the fraction of funds which reject the null, we use a binomial t-test (Ferson and Chen 2014, appendix), which is reported in table 3 in square brackets.[17]

[Table 3 here]

For both periods 1990-99 (Panel A) and 2000-09 (Panel B) there are no statistically significant positive market timers $\gamma_m > 0$ (column 3) for funds with a German, European or Global mandate. Table 3 (column 6) shows that across all three mandates there are no funds with statistically significant negative market timing ability $\gamma_m < 0$ in 1990-99 (Panel A, column 6). But in 2000-2009 there is a substantial increase in statistically significant negative market timing, from funds trading in European (17.7%, 40 funds) and Global (14.3%, 35 funds) markets - but no statistically significant unsuccessful timers in the domestic (German) market. Overall it appears that the substantial move into funds with European and Global mandates has led to an increase in poor market timing skills[18].

Hence for equity funds the parametric results indicate that no funds are successful market timers (table 3 column 3) but there is a substantial increase in the 2000-09 period of unsuccessful market timers for funds with European and Global mandates (column 6, table 3).

---

16 The correlation matrix for our three factors for the 1990-99 and 2000-09 periods are low, indicating largely independent factors that affect fund returns. Results are qualitatively similar when using the HM style timing variables $r_m^+, R_{SMB}^+, R_{HML}^+$ in place of the TM style timing variables. Addition of cross-product terms in the returns are found to be statistically insignificant – these results are available on request.

17 If $\hat{p}$ is the estimated proportion of funds which reject the null using a test size of λ then a test that $\hat{p}$ is statistically different from λ when testing N funds is $t = (\hat{p} - \lambda) / \sqrt{\lambda(1-\lambda)(1/N)[1+(N-1)\hat{\rho}]}$ where the average correlation $\hat{\rho} = [N(N-1)]^{-1} \sum_j \sum_{i \neq j} \hat{\rho}_{ij}$ and $\hat{\rho}_{ij}$ are the pairwise sample correlations between the tests for fund-i and fund-j. The $\hat{\rho}_{ij}$ are estimated using the residuals from the return regressions, where we assume that the correlation for funds (i,j) with no overlapping data are zero. We are grateful to an anonymous referee for suggesting this test.

18 As we use monthly data, bias due to stale pricing effects are unlikely to alter these results and on adding lagged values of the factors we find this is the case.

Turning now to size timing, between 1990-99 across all mandates (Table 3, Panel A, column 4) there are no statistically significant successful size-timers ($\gamma_{SMB} > 0$) but in the 2000-09 period (Panel B, column 4) there are 7.4% (18 funds) with a Global mandate with statistically significant size timing.

A similar pattern is found for successful growth timing, $\gamma_{HML} > 0$ (column 5). There are hardly any successful growth timers in the 1990-99 period (Panel A) but in the 2000-09 period there are around 20% of funds across all three mandates with statistically significant $\gamma_{HML} > 0$[19].

What about *unsuccessful size timers*? There are no unsuccessful size timers in 1990-1999 (Panel A) but an increase in the number of statistically significant unsuccessful size timers ($\gamma_{SMB} < 0$) in 2000-09, for funds with European (16.4%, 37 funds) and Global mandates (16.4%, 40 funds). For *unsuccessful growth timers*, there are 32% (22 funds) with a German mandate with $\gamma_{HML} < 0$ in 1990-2000, but this falls to zero in 2000-2009 (Panel B).

To sum up our parametric style timing results. For the later 2000-2009 period the parametric style timing regressions show a relatively large percentage of successful growth timers ($\gamma_{HML} > 0$) of around 20% for each of German, European and Global mandates (column 5, panel B) but also a relatively large percentage of unsuccessful size timers ($\gamma_{SMB} < 0$) of around 16% for funds with European or Global mandates – column 7, panel B.

4.2. Equity Funds: Non-Parametric Approach

In table 4 we present results on market and style timing using the non-parametric approach for 1990-99 (Panel A) and 2000-09 (Panel B), and compare these with the parametric results reported in Table 3. Results differ between the two approaches. We begin with market timing.

[Table 4 - here]

Equity Funds: Market Timing

For both periods and across all investment mandates, the non-parametric market timing measure gives zero statistically significant successful market timers $\theta_m > 0$ (Table 4, column 3) – the same as for the parametric $\gamma_{TM} > 0$ measure (Table 3, column 3). In the 2000-2009 period,

19 The much larger number of successful growth timers than either size timers or market timers is also found for US equity *growth funds* by Chen, Adams and Taffler (2013).

the non-parametric measure indicates far fewer unsuccessful market timers θ_m < 0 (Table 4, column 6) than the γ_m < 0 measure (Table 3, column 6)[20].

Overall, both the parametric and non-parametric methods indicate that the number of successful market timers is insignificant in both periods, but there is a significant number of negative timers in the 2000-09 period. Lack of strong evidence supporting successful market timing by managed equity funds may be due to a genuine lack of skill in predicting benchmark returns and the latter is certainly consistent with evidence on daily/monthly predictability and parameter instability in time series forecasting equations for stock market returns – see for example, Ang and Bekaert (2007).

Apparent, negative timing may also be due to bias in both the parametric return-based timing measures and our non-parametric measures. Chen et al (2010) note that controlling for non-timing related sources of non-linearity in bond fund returns leads to much less negative timing and their overall result is that timing is "neutral to weakly positive" for US bond funds. Jiang (2003) notes that possible sources of bias in both approaches arise from option-related spurious timing\"passive timing" (Jagannathan and Korajczk 1986) and interim trading bias (Ferson and Khang 2001). In a later paper Jiang, Yao and Yu (2007) note that a *holdings-based* timing measure does not suffer from these two biases and are therefore able to quantify the negative bias found in the return-based regressions on US data. Overall, using the holdings-based approach, they find some evidence of statistically significant positive timing (over 3 and 6 month holding periods) on average - and no evidence of statistically significant negative timing. In contrast, return-based timing measures exhibit no statistically significant positive or negative timing effects - but the point estimate of the mean or median timing effect is negative. This suggests negative bias for the returns-based tests on US data. However, there are no adjustments made for multiple hypothesis tests (Ferson and Chen 2014) in the above studies. As we do not have holdings data we cannot examine this potential bias in our results on German data.

Another reason for not finding evidence of successful market timing may be due to the "dilution effect". Funds experience an increase in investor cashflows during periods when the *market* return is relatively high (Warther 1995, Edelen and Warner 2001), hence increasing the fund's cash position, leading to a concurrent lower overall portfolio return[21] (Bollen and Busse

---

[20] We also examined non-parametric tests of *conditional* market timing $\tilde{\theta}_m$ using the dividend-price ratio, the short-term (one-month) interest rate and the government yield spread as predictor variables. These are broadly similar to reported results for the unconditional measure. Using detrended information variables (by deducting the average value over the previous 3 months) to take account of persistence in these variables, also produced similar results.

[21] It is also well documented that funds with the highest *relative* returns experience the highest cash inflows and the relationship is non-linear (e.g. Ippolito 1992, Gruber 1996, Chevalier and Ellison 1997, Sirri and Tufano 1998, Massa 2003, Lynch and Musto 2003, Nanda, Wang and Zheng 2004, Barber, Odean and Zheng 2004, Huang, Wei and Yan 2007, Ivkovic and Weisbenner 2009). However, Jiang (2003) by examining the non-parametric timing coefficients θ_m of load versus no-load funds and institutional versus retail funds, infers that there is little evidence that market timing effects (in the US) are influenced by fund flows. But Huang and Wang (2014) using holdings data find that US government bond funds with more significant timing effects (i.e. size of the t-statistic on γ) do attract higher inflows.

2001). Poor market timing is then the price investors pay for liquidity provision[22] - this is discussed further below.

Over our two data periods the number of funds with a European (Global) mandate increases from 57 (73) to 224 (237), whereas those with a domestic mandate only increase from 65 to 83 (Table 1). A deterioration in performance has been found in US data as funds move away from "well understood" local markets to new and wider markets (Coval and Moskowitz 1999, Hong, Kubik and Stein 2005) and we conjecture similar forces may be operating in the German fund industry.

Equity Funds: Style Timing

It is immediately clear from table 4 (column 4) that the non-parametric test indicates zero successful size timing $\theta_{SMB} > 0$ and growth timing $\theta_{HML} > 0$ in both periods.

Results on negative style timing are mixed, when using the non-parametric approach. For 1990-99 there is evidence of unsuccessful size timing for funds with a German mandate where 27.5% (19 funds) have $\theta_{SMB} < 0$ (Table 4 panel A, column 7). However, in the later 2000-09 period there is strong evidence of unsuccessful size timing $\theta_{SMB} < 0$ for the Global mandate (31%, 76 funds – column 7). There are no funds which are unsuccessful growth timers ($\theta_{HML} < 0$) in 1990-99 but a substantial number of funds across all 3 mandates are unsuccessful growth timers over 2000-09 period (Table 4, Panel B, column 8).

Overall, the non-parametric results indicate few successful style timers in either period but a substantial number of unsuccessful style timers in the 2000-09 period.

4.3 Robustness Tests

So far we have analysed the percentage (and number) of funds that have positive or negative market and style timing, using parametric and non-parametric measures. We now examine whether *market timing* effects are similar across different parametric models and the non-parametric approach – thus testing robustness across alternative methodologies.

[Table 5 here]

Table 5 shows the rank correlation between $t(\gamma_m)$ for funds, using four alternative parametric models (CAPM+TM, CAPM+HM, 3F+TM, 3F+HM) and the non-parametric measure θ_m, over our two sample periods. The high correlation coefficients (> 0.9) for the four parametric models show that the measured market timing effect is largely independent of the specific

22 Using data on *all trades* of Canadian mutual funds, Christoffersen, Keim and Musto (2006) find that cash inflows result in flat or negative returns on stocks purchased and positive or flat returns on stocks sold – so transaction costs consequent on cash inflows lead to low fund returns.

parametric model used. The rank correlation between the parametric market timing effect $t(\gamma_m)$ and the non-parametric timing effect θ_m varies between 0.54 and 0.77 which reinforces the importance of using both methods. (This is consistent with the evidence reported in tables 3 and 4 which exhibit different outcomes between the two approaches).

Next, using the non-parametric approach we examine if funds which have strong market timing effects also have strong size-timing and growth-timing – Is the "timing success" of funds correlated across the market, size and growth factors?

[Table 6 here]

Table 6 examines whether market, size and growth timing are correlated across funds[23]. For the period 1990-99 (Panel A), a fund's market timing statistic θ_m has a near zero correlation with its ranking in terms of size timing θ_{SMB} and a correlation of only 0.4 with respect to growth timing, θ_{HML}. These correlations increase in the 2000-09 period and all three "style correlations" are positive and statistically significant. Hence, those funds with positive (negative) market timing also tend to have positive (negative) size timing and growth timing effects - but the rank correlation coefficients are not large (0.24 to 0.44).

Misspecification Tests

We undertake two specification tests. First, a "cross-product test" adds all cross-product terms between the (three) factors and the squared (timing) factors in the 3F-model. Second we test for bias due to stale prices, by adding lead and lagged values of the factors to the 3F-model (Dimson 1979). These tests are applied to the equally weighted portfolios of table 2 and the individual funds in table 3[24].

For the equally weighted portfolios (table 2) when we add the cross-product terms, there is less evidence of statistically significant negative timing effects overall for the period 1900-1999, while for the 2000-2009 period there is a very little change in the overall results for negative timing. Hence, based on equally weighted portfolios, our result suggest some negative bias in timing effects when using the parametric model.

Having looked at EW portfolios of funds we now apply this "cross-product" misspecification test to individual funds. From table 3 we see that in the 2000-2009 period (when there are more funds) there is evidence of considerable negative timing (based on the binomial t-test) - particularly for the European and Global mandates. We now add all the cross-product terms to each fund regression. Testing the parameter restrictions (on the cross-product terms) for each fund separately at a "high" 1% significance level, we find 87% of the funds for the 1990-1999 period *do*

---

23 Results are not reported for conditional parametric models of timing as these are qualitatively similar to our unconditional results.

24 A referee suggested these robustness tests. Details of all the tests in this section are available in an appendix to the paper, available on request.

not reject the null that the additional cross-product terms are zero and for the 2000-2009 period the equivalent figure is 67.1%. So in the 2000-2009 period there is some prima-facie evidence of misspecification.

In the period 1990-1999 there is little change in the proportion of funds with statistically significant positive or negative timing effects when the cross-product terms are added - except for the proportion of funds with $\gamma_{HML} < 0$ which drops dramatically from 31.88% to 4.35%. In the 2000-2009 period there are some dramatic falls in the incidence of statistically significant negative timing but the incidence of positive timing remains broadly unchanged. The theoretical deficiencies of parametric timing effects are well documented (above) and the cross-product misspecification test is suggestive that negative timing effects may be somewhat overstated in the basic parametric timing equation.

The estimated model to account for possible stale pricing (Dimson 1979) is:

$$\begin{aligned} r_t = {} & \alpha + \beta_{m,1} r_{m,t} + \beta_{m,2} r_{m,t+1} + \beta_{m,3} r_{m,t-1} + \beta_{SMB,1} R_{SMB,t} + \beta_{SMB,2} R_{SMB,t+1} + \beta_{SMB,3} R_{SMB,t-1} \\ & + \beta_{HML,1} R_{HML,t} + \beta_{HML,2} R_{HML,t+1} + \beta_{HML,3} R_{HML,t-1} + \gamma_m r_{m,t}^2 + \gamma_{SMB} R_{SMB,t}^2 + \gamma_{HML} R_{HML,t}^2 + \varepsilon_t \end{aligned}$$

When we compare the results using the above equation for each fund, with the results for individual funds in table 3 (without lead and lagged variables) we see no substantial change in the results, suggesting stale prices in monthly data are not a major source of misspecification.

Flow-performance Relationship

We investigate the relationship between fund flows and market timing - whether high inflow funds subsequently have better or worse timing ability than low inflow funds. Each month, for the German, European and Global mandates separately, we sort funds into (equally weighted, EW) quintile portfolios based on their relative cash inflows (over the past month)[25]. The resulting quintile portfolio returns are then regressed against the three Fama and French factors and their timing variables (i.e. Fama-French factors squared).

[Table 7 here]

Because of the relatively small number of funds in the 1990-1999 period we report results only for the 2000-2009 period (Table 7). There is evidence of statistically significant negative market timing for the middle three quintile funds with a Global mandate and for the two lower quintiles of funds with a European mandate (and no statistically significant market timing for quintile funds with a German mandate). There is also evidence of some statistically significant negative size timing, particularly for some quintile sorted funds with a German mandate. In contrast to the above, there is strong evidence of successful growth timing for all quintile sorted

25 Fund flows are measured in the usual way as $\%Flow_t = [NAV_t – NAV_{t-1} (1+r_t)] / NAV_{t-1}$

funds with a Global mandate. However, the market, size and growth timing coefficients for high inflow quintile funds are generally not statistically different from those for low inflow quintile funds - the exception being for size timing for funds with a global mandate, where low inflow funds have a larger negative timing effect than high inflow funds, with t-statistic 2.09). Hence overall, timing ability does not seem to be affected by relative fund flows[26].

4.4 Bond Funds

Table 8 shows summary statistics for our sample of bond funds for 1990-99 (Panel A) and 2000-09 (Panel B) for German domiciled funds that have German, European and Global mandates. There is an increase in the number of bond funds with a European or Global mandate from around 80 to 165 funds - but no increase in bond funds with a German mandate. (This mirrors the expansion in equity funds with non-German mandates - see Table 1). Average bond returns are lower in the 2000-09 period (with standard deviations somewhat smaller).

[Table 8 here]

As expected, bond funds with a German or European mandate have a relatively heavy weighting on the European bond index of around 0.55 - 0.60 which does not vary between the two sub-periods. Bond funds with a Global mandate have a lower weight on the European index of around 0.29 and 0.46 in the two sub-periods. The average betas on the other three bond factors (broad investment grade, high yield and global) are much smaller in both sub-periods (the largest being 0.16). However, the correlation between the broad investment grade index with either the high yield or global index is relatively high, at 0.52 and 0.84, respectively - which may mask each individual factor's contribution to explaining bond returns (table 9).

[Table 9 here]

There are virtually no bond funds with statistically significant positive alphas in either data period but 70 (out of 215) funds and 153 (out of 389) funds have statistically significant negative alphas in the periods 1990-1999 and 2000-2009, respectively. The average $\bar{R}^2$ for the 4-factor bond model are in the range 0.50-0.54[27] (consistent with other international bond factor models) and non-normality in many funds' residuals motivates the use of bootstrap t-values).

Before considering the timing skill of *individual* bond funds, table 10 examines the 4F model with the addition of timing variables, for an *equally weighted* portfolio over the period 1990-2009. The beta coefficient of the European index is 0.54 (t = 17.1) and although the betas of the other indices are much smaller, that for the high yield index is statistically significant (β = 0.088, t

---

[26] Including all the cross-product terms in these quintile regressions does not change this general conclusion – although there are changes in a few timing coefficients.

[27] This is comparable to the R-squared of around 0.5 found for US corporate bond funds by Amihud and Goyenko (2012). Results for the Henriksson-Merton model are qualitatively similar and are not reported.

= 8.4) – these results are broadly consistent with those in table 8. The timing ability of the average bond fund shows positive (and statistically significant) timing of the broad investment grade index and a negative (statistically significant) timing effect from the global index – but the timing coefficients for the European and high yield indices are statistically insignificant. This provides prima facie evidence of some timing skill amongst bond funds but these "average effects" could mask considerable variation across individual funds - an issue we now address.

[Table 10 here]

Table 11 summarises market timing effects estimated fund-by-fund for the 4F Treynor-Mazuy parametric model $\gamma(TM)$ and the non-parametric approach θ_m, with respect to the investment grade, high yield, European and Global bond indices.

In the 1990-99 period (Panel A) for both approaches, the proportion of successful timers is insignificant - but for the non-parametric method there is a statistically significant proportion of unsuccessful timers, $\theta_m < 0$, for the European government (13.5%, 29 funds) and global government (30.7%, 66 funds) bond indexes.

[Table 11 here]

In the 2000-09 period (Panel B), there are statistically significant positive timing results $\gamma(TM) > 0$, for the parametric model for 10% to 20% of funds across the 4 indexes[28], while the non-parametric method gives statistically significant positive timing $\theta_m > 0$ only for the broad investment grade (9.8%, 37 funds) and high yield (16%, 60 funds) bond indexes. For negative timing, the parametric method gives a statistically significant $\gamma(TM) < 0$ for 11% to 20% of funds across the 4 indexes, while the non-parametric has statistically significant values of $\theta_m < 0$ only for the European government (6.9%, 26 funds) and global government (8.5%, 32 funds) indexes. Hence overall, the non-parametric method suggests there are both fewer successful and unsuccessful timers in the 2000-09 period, compared with the parametric results.

Part of the reason for this difference (particularly noticeable in the 2000-09 period) may be that the parametric approach is based on both the quality of a fund manager's timing information and the aggressiveness of her response, whereas the non-parametric method depends only on the quality of the information. However overall, the parametric and non-parametric results for the later 2000-09 period both provide some evidence of successful and unsuccessful timing by German bond funds.

[28] Addition of lagged factor returns to the TM model to account for stale pricing effects are found to be statistically insignificant and do not qualitatively change our reported results.

5. Conclusions

Using a parametric and non-parametric approach, we assess both market timing and style timing of individual German equity and bond funds which have domestic, European and global mandates.

In both sample periods (1990-1999 and 2000-2009), the proportion of positive market timers is insignificant under both the parametric and non-parametric approaches, across all three fund mandates. In the 2000-2009 period the parametric method indicates a substantial proportion of statistically significant negative market timers but the non-parametric measure indicates far fewer unsuccessful market timers. Both approaches show an increase in unsuccessful equity market timers in the 2000-09 period which is mainly due to funds with European and Global mandates rather than domestic German mandates. This suggests that German domiciled asset managers may have less skill in forecasting markets that are less familiar and for which they have less timely information. It is also consistent with Hong, Kubic and Stein (2005) who find that fund managers are more likely to buy/sell a stock if other “local fund managers” are also buying/selling the stock – this could lead to bad timing decisions over “foreign stocks” by a subset of German fund managers spreading to other “local fund managers” with a foreign mandate.

We also investigate style timing, where the parametric and non-parametric tests sometimes give different inferences. For example, over the period January 2000-December 2009 the parametric approach indicates statistically significant successful (positive) size and growth timing whereas the non-parametric approach does not. The non-parametric approach also suggests there are more negative growth timers in the 2000-2009 period than the parametric approach. However, overall, both approaches indicate a substantial number of funds with either unsuccessful (negative) size timing or growth timing - but misspecification tests suggest that the incidence of negative timing may be overstated in the parametric approach.

Results for bond funds using parametric and non-parametric methods show some preliminary evidence of both successful and unsuccessful timing, particularly in the 2000-2009 period. Some fund managers may move into (out of) long duration bonds after successfully forecasting a fall (rise) in interest rates but there are also some bond funds which, unsuccessfully time the market – decreasing (increasing) their market exposure before the market return rises (falls).

Although we have applied a number of alternative methodologies and models, further work is needed on the sensitivity of parametric timing coefficients to alternative benchmark factors particularly for international equity funds (Fletcher and Marshall 2005, Comer and Rodriguez 2012). An alternative parametric approach for bond funds would be to investigate timing with respect to specific “economic” factors (e.g. shape of the yield curve, liquidity, etc - Chen et al 2013) rather than our use of bond market indices. Finally, holdings based tests of market and style timing avoid some of the biases in returns-based and non-parametric based tests – but for German data, holdings data is not yet available.

References

Abrevaya, J. and Jiang, J., 2005. A non-parametric approach to measuring and testing curvature. Journal of Business and Economic Statistics, 23 (1), 1-19.

Admati, A. and Pfleiderer, P., 1997. Does it all add up? Benchmarks and the compensation of active portfolio managers. Journal of Business, 70 (3), 323-350.

Admati, A., Bhattacharya, S and Pfleiderer, P., 1986. On timing and selectivity. Journal of Finance, 41(3), 715-730.

Amihud, Y. and Goyenko, R., 2012. Mutual Fund's R2 as Predictor of Performance. SSRN working paper

Ammann, M., Kind, A. and Seiz, R., 2010. What drives the performance of convertible bond funds? Journal of Banking and Finance, 34, 2006-13.

Ang, A. and Bekaert, G., 2007. Stock return predictability: Is it there? Review of Financial Studies, 20 (3), 651-707.

Aragon, G., 2005. Timing multiple markets: theory and evidence from balanced mutual funds, Working Paper, Arizona State University.

Artmann, S., Finter, P., Kempf, A, Koch, S. and Theissen, E., 2010. The cross-section of German stock returns: new data and new evidence. University of Cologne, CFR working paper 10-12.

Ayadi, M.A. and Kryzanowski, L., 2011. Fixed-income performance: role of luck and ability in tail membership. Journal of Empirical Finance, 18, 379-92.

Banegas, A., Gillen B., Timmermann A. and Wermers R., 2008. The performance of European equity mutual funds. SSRN working paper

Barber, B.M., Odean T. and Zheng L., 2004. Out of sight, out of mind: the effects of expenses on mutual fund flows. Journal of Business, 78 (6), 2095-2120.

Becker, C., Ferson, W., Myers, D. and Schill M., 1999. Conditional market timing with benchmark investors. Journal of Financial Economics, 52 (1), 119-148.

Berk, J.B., and Green R.C., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy, 112, 1269-95.

Bessler, W., Drobetz, W. and Zimmermann, H., 2009. Conditional performance evaluation for German equity mutual funds. European Journal of Finance, 15 (3), 287-316.

Blake, C.R., Elton E.J. and Gruber, M.J., 1993. The performance of bond mutual funds. Journal of Business, 66, 371-403.

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

Boney, V. and Comer, G, 2010. Are reported mutual fund bond yields useful? An analysis of municipal bond funds. SSRN working paper.

Boney, V., Comer, G. and Kelly, L., 2009. Timing the investment grade securities market: evidence from high quality bond funds. Journal of Empirical Finance, 16, 55-69.

Breen, W., Jagannathan R., and Ofer J., 1986. Correcting for heteroscedasticity in tests for market timing ability. Journal of Business, 59 (4), 585-598.

Brown, K.C., Harlow W.V. and Starks, L.T., 1996. Of tournaments and temptations : an analysis of managerial incentives in the mutual fund industry. Journal of Finance, 51, 85-110.

Busse, J., 1999. Volatility timing in mutual funds: evidence from daily returns. Review of Financial Studies, 12 (5), 1009 – 1041.

Byrne, A., Fletcher, J. and Ntozi P., 2006. An exploration of the conditional timing performance of UK unit trusts'. Journal of Business Finance and Accounting, 33, 816-838.

Carpenter, J., 2000. Does option compensation increase managerial risk appetite. Journal of Finance, 55 (5), 2311-2331.

Chen, C.R., Lee, C.F., Rahman S. and Chan, A., 1992. A cross-sectional analysis of mutual fund's market timing and security selection skill. Journal of Business, Finance and Accounting, 19 (5), 659-675.

Chen, H., Estes, J. and Ngo, T., 2011. Tax Calendar Effects in the Municipal Bond Market:Tax-Loss Selling and Cherry Picking by Investors and Market Timing by Fund Managers. SSRN working paper.

Chen, L., Adams, A. and Taffler, R., 2013. What style-timing skills do mutual fund "stars" possess?, Journal of Empirical Finance, 21, 156-173.

Chen, Y. and Liang, B., 2007. Do market timing hedge funds time the market?, Journal of Financial and Quantitative Analysis, 42 (4), 1102-1122.

Chen, Y., Ferson, W. and Peters, H., 2009. Measuring the timing ability and performance of bond mutual funds. Journal of Financial Economics, 98, 72-89.

Chevalier, J. and Ellison, G., 1997. Risk taking by mutual funds as a response to incentives, Journal of Political Economy, 105 (6), 1167-1200.

Christoffersen, S., Keim, D. and Musto, D., 2006. Valuable information and costly liquidity: evidence from individual mutual fund trades. SSRN working paper.

Christopherson, J.A., Ferson W.E. and Glassman, D.A., 1998. Conditioning manager alphas on economic information: another look at persistence in performance. Review of Financial Studies, 11, 111-142.

Comer, G. and Rodriguez, J., 2012. International mutual funds: MSCI benchmarks and portfolio evaluation. SSRN working paper.

Comer, G., 2006. Hybrid mutual funds and market timing performance. Journal of Business, 79, 771-97.

Comer, G., Rodriguez, J., 2006. Corporate and government bond funds: an analysis of investment style, performance and cash flows. SSRN working paper.

Comer, G., Rodriguez, J., 2011. Do mutual fund media recommendations hold value? An empirical analysis of the *Wall Street Journal* Smart Money fund screen, SSRN working paper.

Cornell, B. and Green, K., 1991. The investment performance of low-grade bond funds. Journal of Finance, 46, 29-48.

Coval, J.D., and Moskowitz, T.J., 1999. Home bias at home: local equity preference in domestic portfolios. Journal of Finance, 54, 2045-2074.

Cremers, M., Ferreira, M., Matos, P. and Starks, L., 2011. The mutual fund industry worldwide: explicit and closet indexing, fees and performance. SSRN working paper.

Cuthbertson, K., and Nitzsche, D., 2013. Performance, stock selection and market timing of the German equity mutual fund industry. Journal of Empirical Finance, 21, 86-101.

Cuthbertson, K., Nitzsche, D. and O'Sullivan, N., 2010. The market timing ability of UK mutual funds. Journal of Business, Finance and Accounting, 37, 270-289.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics, 7, 197-226.

Du, D., Huang, Z. and Blanchard, P.J., 2009. Do fixed-income mutual fund managers have managerial skills? Quarterly Review of Economics and Finance, 49, 378-397.

Edelen, R.M. and Warner, J.B., 2001. Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. Journal of Financial Economics, 59, 196-220.

Edelen, R.M., 1999. Assessed flows and the performance of open ended mutual funds. Journal of Financial Economics, 53, 439-466.

Elton, E.J., Gruber, M.J. and Blake, C.R., 1995. Fundamental economic variables, expected returns and bond fund performance. Journal of Finance, 40, 1229-1256.

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

Fama, Eugene F. and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics, Vol. 33, 3-56.

Fama, Eugene F. and Kenneth R. French, 2010, Luck versus Skill in the Cross Section of Mutual Fund Returns, Journal of Finance, Vol. 65, No. 5, 1915-1947.

Ferson, W.E. and Chen, Y., 2014. How many good and bad fund managers are there, really? Marshall School of Business, University of California, August 17th 2014.

Ferson, W.E. and Khang, K., 2001. Conditional performance measurement using portfolio weights: evidence for pension funds. Journal of Financial Economics, 65, 249-282.

Ferson, W.E., Henry, T. and Kisgen, D., 2006. Evaluating government bond performance with stochastic discount factors. Review of Financial Studies, 19, 423-56.

Fletcher, J. and Marshall, A., 2005. An empirical examination of UK international unit trust performance. Journal of Financial Services Research, 27 (2), 183-206.

Fletcher, J., 1995, An examination of the selectivity and market timing performance of UK unit trusts. Journal of Business Finance and Accounting, 22 (1), 143-156.

Glassman, D. and Riddick, L., 2006. Market timing by global fund managers. Journal of International Money and Finance, 25, 1029-50.

Goetzmann, W., Ingersoll J. and Ivkovich, Z., 2000. Monthly measurement of daily timers. Journal of Financial and Quantitative Analysis, 35 (3), 257-290.

Griese, K. and Kempf, A., 2003. Lohnt aktives Fondsmanagement? Ein Vergleich von Anlagestrategien in aktiv und passiv verwaltenten Aktienfonds. Zeitschrift fuer Betriebswirtschraft, 73, 201-224.

Grinblatt, M. and Titman, S., 1989. Portfolio performance evaluation: old issues and new insights. Review of Financial Studies, 2 (3), 393-421.

Gruber M., 1996. Another puzzle: the growth in actively managed mutual funds. Journal of Finance, 51 (3), 783–810.

Hendriksson, R.D., 1984. Market timing and mutual fund performance: an empirical investigation. Journal of Business, 57 (1), 73-96.

Henriksson, R. and Merton, R., 1981. On market timing and investment performance II: statistical procedures for evaluating forecasting skills. Journal of Business, 54 (4), 513-33.

Hong, H., Kubik J.D. and Stein, J.C., 2005. Thy neighbor's portfolio : word-of-mouth effects in the holdings and trades of money managers. Journal of Finance, 60 (6), 2801-2824.

Huang, J. and Wang, Y., 2014. Timing ability of government bond fund managers: Evidence from portfolio holdings. Management Science, 60, 2091-2109.

Huang, J., Wei K.D. and Yan, H., 2007. Volatility of performance and mutual fund flows. SSRN working paper

Huij, J. and Derwall, J., 2008. Hot hands in bond funds. Journal of Banking and Finance, 32 (4), 559-27.

Huij, J. and Derwall, J., 2011. Global equity fund performance, portfolio concentration and the fundamental law of active management. Journal of Banking and Finance, 35, 155-65.

Hyde, S. and Kappel, K., 2010. Predicting German stock and bond returns with macro-variables: evidence of market timing. SSRN working paper.

Ippolito, R., 1992. Consumer reaction to measures of poor quality: evidence from the mutual fund industry. Journal of Law and Economics, 35, 45-70.

Ivkovic, Z. and Weisbenner, S., 2009. Individual investor mutual fund flows. Journal of Financial Economics, 92, 223-237.

Jagannathan, R. and Korajczyk, R., 1986. Assessing the Market Timing Performance of Managed Portfolios. Journal of Business, 59 (2), 217-235.

Jiang, G.J., Yao, T. and Yu, T., 2007. Do mutual funds time the market? Evidence from portfolio holdings. Journal of Financial Economics, 86 (3), 724-758.

Jiang, W., 2003, A non-parametric test of market timing. Journal of Empirical Finance, 10 (4), 399-425.

Kaplan, S.N. and Sensoy, B.A., 2010. Do mutual funds time their benchmarks? SSRN working paper.

Kosowski, R., A. Timmermann, H. White and R. Wermers, 2006, Can Mutual Fund 'Stars' Really Pick Stocks? New Evidence from a Bootstrapping Analysis, Journal of Finance, 61(6), 2551-2595.

Kothari, S. and Warner, J., 2001. Evaluating mutual fund performance. Journal of Finance, 56 (5), 1985-2010.

Krahnen, J.P., Schmid, F. and Theissen, E., 2006. Investment performance and market share: a study of the German mutual fund industry. Wolfgang Bessler: Boersen, Banken and Kapitalmaerkte, Duncker & Humblot, Berlin, 471-491.

Ledoit, O. and Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio, Journal of Empirical Finance, 15, 850-859.

Lee, C. and Rahman, S., 1990. Market timing, selectivity, and mutual fund performance: an empirical investigation. Journal of Business, 63, 261–78.

Leger, L., 1997. UK investment trusts: performance, timing and selectivity. Applied Economics

Letters, 4 (2), 207-210.

Lynch, A.W. and Musto, D.K., 2003. How investors interpret past fund returns. Journal of Finance, 58 (5), 2033-2058.

Massa, M., 2003. How do family strategies affect fund performance? When performance-maximization is not the only game in town. Journal of Financial Economics, 67 (2), 249-304.

Nanda, V.Z., Wang, J. and Zheng, L., 2004. The ABCs of mutual funds: a natural experiment on fund flows and performance. University of Michigan working paper.

Otten, R. and Bams, D., 2002. European mutual fund performance. European Financial Management, 8 (1), 75-101.

Pastor, L. and Stambaugh, R.F., 2010. On the size of the active management industry. SSRN working paper.

Reneeboog, L., Horst, J. and Zhang C, 2004. The price of Ethics: Evidence from Socially Responsible Mutual Funds, SSRN working paper.

Sensoy, B., 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry. Journal of Financial Economics, 92 (1), 25-39.

Serfling, R., 1980. Approximation theorems of mathematical statistics, Wiley, New York.

Sharpe, W.F., 1992. Asset Allocation: Management Style and Performance Measurement. Journal of Portfolio Management, 18 (2), 7-19.

Silva, F., Cortes, M.C. and Armada, M.R., 2005. The persistence of European bond fund performance: does conditioning information matter? International Journal of Business, 10 (4), 341-61.

Sirri, E.R. and Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance, 53 (5), 1589-1622.

Stehle, R. and Grewe, O., 2001. The long-run performance of German stock mutual funds. Humboldt University, Berlin, working paper.

Stotz, O. 2007. Selection, market timing and style timing of equity mutual funds: evidence from Germany. Zeitschrift fur Betriebswirtschaft, 77, 51-73.

Swinkels, L. and Tjong-A-Tjoe, L., 2007. Can mutual funds time investment styles? Journal of Asset Management, 8, 123–32.

Treynor, J. and Mazuy, K., 1966. Can mutual funds outguess the market? Harvard Business Review, 44 (4), 131-136.

Warther, V.A., 1995. Aggregate mutual fund flows and security returns. Journal of Financial Economics, 39, 209-235.

Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs and expenses. Journal of Finance, 55 (4), 1655-1695.

Table 1. Equity Funds: Summary Statistics

This table provides summary statistics for equity mutual funds for the periods January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B) for funds with a German, European or Global mandate. Data frequency is monthly. Statistics include the number (#) of funds used, the average number of observations per fund, the average monthly return (% pm), the average standard deviation (% pm), the average Fama-French three-factor alpha, the number (#) of statistically significant positive and negative alphas (based on bootstrap test statistics), the average weights on the market, SMB and HML factors, the average R-squared and the number of funds with non-normal residuals (using the Jarque-Bera JB test statistic, at a 5% significant level).

Panel A : January 1990 – December 1999				
	All	Germany	Europe	Global
# of Funds	195	65	57	73
Average # observations	66.7	73.3	58.8	66.9
Average Return	1.31	1.14	1.32	1.46
Average SD	6.17	6.18	6.11	6.21
Average alpha	-0.0326	-0.2016	-0.2008	0.2492
# signif. pos./neg. alphas	11/10	2/6	2/2	7/2
Average β_m	0.95	0.93	1.0	0.91
Average β_{SMB}	0.27	0.15	0.32	0.34
Average β_{HML}	-0.05	0.02	-0.04	-0.13
Average $\bar{R}^2$	0.75	0.82	0.75	0.67
# signif. non-normal, JB	55	14	17	24
Panel B: January 2000 – December 2009				
	All	Germany	Europe	Global
# of Funds	544	83	224	237
Average # observations	86.4	99.7	83.7	84.3
Average Return	-0.66	-0.42	0.68	-0.72
Average SD	6.06	6.75	5.84	6.02
Average alpha	-0.24	-0.19	-0.31	-0.20
# signif. pos./neg. alphas	6/89	0/10	2/51	4/28
Average β_m	0.97	0.88	0.99	0.97

Average β_{SMB}	0.35	0.30	0.33	0.39
Average β_{HML}	-0.20	0.07	-0.16	-0.33
Average $\bar{R}^2$	0.80	0.86	0.82	0.77
# signif. non-normal, BJ	241	55	101	85

Table 2. EW Portfolios of funds with same investment mandate

This table reports results for the Fama-French three factor model augmented with the TM market and style timing variables ($r_m^2, R_{SMB}^2, R_{HML}^2$). Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B) for funds with a German, European and Global mandate. Data frequency is monthly. Only funds with at least 24 monthly observations have been included. Newey-West corrected t-statistics are reported in parentheses.

Panel A : January 1990 – December 1999						
	Germany		European		Global	
	coefficient	t-stat	coefficient	t-stat	coefficient	t-stat
α	-0.1709	-1.49	0.2666	1.35	0.0285	0.1181
β_m	0.9434	45.19	0.9869	29.60	0.8014	22.38
β_{SMB}	0.1945	4.88	0.3407	4.83	0.3048	5.09
β_{HML}	0.0157	0.57	-0.1765	-2.21	-0.1051	-1.13
γ_m	0.0025	1.28	-0.0114	-2.43	-0.0085	-1.61
γ_{SMB}	0.0023	0.38	-0.0309	-1.80	0.0046	0.32
γ_{HML}	-0.0053	-1.61	0.0341	2.10	0.0511	2.64
Panel B : January 2000 – December 2009						
	Germany		European		Global	
	coefficient	t-stat	coefficient	t-stat	coefficient	t-stat
α	-0.1805	-1.22	-0.1703	-1.42	-0.0960	-0.63
β_m	0.8881	54.91	0.9972	44.82	0.9780	35.31
β_{SMB}	0.2924	10.08	0.3031	7.86	0.3202	6.70
β_{HML}	0.0822	3.48	-0.2096	-4.78	-0.3887	-7.58
γ_m	-0.0011	-0.77	-0.0044	-1.45	-0.0077	-1.81
γ_{SMB}	-0.0033	-0.62	-0.0032	-0.31	-0.0120	-1.38
γ_{HML}	0.0028	1.25	0.0070	0.79	0.0335	2.63

Table 3. Equity Funds: Market and Style Timing, Parametric Models.

This table reports results for the Fama-French three factor model augmented with the TM market and style timing variables. We report the percentage (number) of funds with significant market timing γ_m, size timing γ_{SMB} and growth timing γ_{HML} coefficients for the Treynor-Mazuy specification for the timing variables ($r_m^2, R_{SMB}^2, R_{HML}^2$). Individual tests are based on bootstrap (Newey-West) t-statistics and the significance level used is a 2.5% one-tail test. To assess the statistical significance of the fraction of funds which reject the null, we use a (binomial) t-statistic (Ferson and Chen 2014), which is reported below in square brackets. ***, ** and * denotes significance at the 0.5%, 2.5% and 5% level, respectively. Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B) for funds with a German, European and Global mandate. Data frequency is monthly. Only funds with at least 24 monthly observations have been included.

		Positive Timing			**Negative Timing**		
1	2	3	4	5	6	7	8
	Number of Funds	$\gamma_m > 0$	$\gamma_{SMB} >0$	$\gamma_{HML} > 0$	$\gamma_m < 0$	$\gamma_{SMB} <0$	$\gamma_{HML} <0$
Panel A : January 1990 – December 1999							
Germany	69	10.14 (7) [1.03]	1.45 (1) [-0.14]	5.80 (4) [0.44]	1.45 (1) [-0.14]	1.45 (1) [-0.14]	31.88*** (22) [3.95]
Europe	59	1.69 (1) [-0.09]	1.69 (1) [-0.09]	8.47 (5) [0.69]	10.17 (6) [0.88]	3.39 (2) [0.10]	1.69 (1) [-0.09]
Global	79	1.27 (1) [-0.19]	1.27 (1) [-0.19]	18.98** (15) [2.53]	2.53 (2) [0.005]	3.80 (3) [0.20]	2.53 (2) [0.005]
Panel B : January 2000 – December 2009							
Germany	83	3.61 (3) [0.18]	7.23 (6) [0.75]	24.10** (20) [3.44]	9.64 (8) [1.14]	7.23 (6) [0.75]	4.82 (4) [0.37]
Europe	226	5.31 (12) [1.12]	5.31 (12) [1.12]	20.35*** (46) [7.12]	17.70*** (40) [6.06]	16.37*** (37) [5.53]	4.87 (11) [0.94]
Global	244	1.23 (3) [-0.54]	7.38** (18) [2.09]	19.67*** (48) [7.35]	14.34*** (35) [5.07]	16.39*** (40) [5.94]	2.05 (5) [-0.19]

Table 4. Equity Funds: Market and Style Timing, Non-Parametric Model.

This table reports the percentage (number) of funds with significant market timing θ_m , size timing θ_{SMB} and growth timing θ_{HML} non-parametric statistics. The significance level used is a 2.5% one-tail test. To assess the statistical significance of the fraction of funds which reject the null, we use a (binomial) t-statistic (Ferson and Chen 2014), which is reported below in square brackets. ***, ** and * denotes significance at the 0.5%, 2.5% and 5% level, respectively. Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B) for funds with a German, European and Global mandate. Data frequency is monthly. Only funds with at least 24 monthly observations have been included.

		Positive Timing			**Negative Timing**		
1	2	3	4	5	6	7	8
	Number of Funds	$\theta_m > 0$	$\theta_{SMB} > 0$	$\theta_{HML} > 0$	$\theta_m < 0$	$\theta_{SMB} < 0$	$\theta_{HML} < 0$
Panel A : January 1990 – December 1999							
Germany	69	7.25 (5) [0.64]	0 (0) [-0.34]	0 (0) [-0.34]	0 (0) [-0.34]	27.54*** (19) [3.36]	0 (0) [-0.34]
Europe	59	1.70 (1) [-0.09]	0 (0) [-0.29]	0 (0) [-0.29]	0 (0) [-0.29]	0 (0) [-0.29]	0 (0) [-0.29]
Global	79	0 (0) [-0.38]	0 (0) [-0.38]	0 (0) [-0.38]	0 (0) [-0.38]	0 (0) [-0.38]	0 (0) [-0.38]
Panel B : January 2000 – December 2009							
Germany	83	1.21 (1) [-0.21]	0 (0) [-0.40]	0 (0) [-0.40]	1.21 (1) [-0.21]	3.61 (3) [0.18]	21.69*** (18) [3.05]
Europe	226	3.10 (7) [0.24]	0.18 (1) [-0.82]	0.18 (1) [-0.82]	2.21 (5) [-0.11]	4.87 (11) [0.94]	39.38*** (89) [14.71]
Global	244	0.82 (2) [-0.72]	0 (0) [-1.07]	0 (0) [-1.07]	9.02*** (22) [2.79]	31.15*** (76) [12.25]	16.39*** (40) [5.94]

Table 5. Equity Funds: Market Timing, Correlation Across Parametric and Non-Parametric Models.

This table reports Spearman rank correlation coefficients for market timing effects in three parametric models and the non-parametric approach. For each fund, the parametric measure of market timing is the t-statistic, t_{γ_m} and for the non-parametric approach it is θ_m . The four parametric models considered are the CAPM Treynor–Mazuy model (CAPM+TM), the CAPM Henriksson-Merton model (CAPM+HM), the 3 factor Treynor – Mazuy model (3F+TM), the 3 factor Henriksson-Merton model (3F+HM). Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B). Only funds with at least 24 monthly observations have been included. In the first sample period 207 funds are included in the ranking and the second sample includes 553 funds. All correlation coefficients are found to be statistically significantly different from zero – at a 5% significance level.

	Panel A : January 1990 – December 1999				**Panel B : January 2000 – December 2009**			
	CAPM+TM	CAPM+HM	3F+TM	3F+HM	CAPM+TM	CAPM+HM	CAPM+TM	CAPM+HM
CAPM+HM	0.9494				0.9584			
3F+TM	0.9416	0.9022			0.9152	0.8907		
3F+HM	0.9101	0.9500	0.9590		0.8757	0.9222	0.9573	
θ_m	0.5439	0.6197	0.5679	0.6318	0.6729	0.7731	0.6822	0.7550

Table 6. Equity Funds: Correlation Across Market and Style Timing, Non-Parametric Approach.

The table reports Spearman rank correlation coefficients across non-parametric test statistics on market timing θ_m , size timing θ_{SMB} and growth timing θ_{SMB} . Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B). Data frequency is monthly. The p-values for the null of a zero rank correlation coefficient are in parentheses.

	Panel A : Jan. 1990 – Dec. 1999		**Panel B : Jan. 2000 – Dec. 2009**	
	θ_m	θ_{SMB}	θ_m	θ_{SMB}
θ_{SMB}	-0.0943 [0.1896]		0.3071 [< 0.0001]	
θ_{HML}	0.3961 [< 0.0001]	-0.2224 [0.0017]	0.2351 [< 0.0001]	0.4380 [< 0.0001]

Table 7: Fund Flows and Timing Performance

This table reports results for the Fama-French three factor model augmented with the TM market and style timing variables ($r_m^2, R_{SMB}^2, R_{HML}^2$). Each month for German, European and Global mandates separately, we sort funds into (equally weighted, EW) quintile portfolios based on their relative cash inflows (over the past month). The resulting quintile portfolio returns are then regressed against the three Fama-French factors for each mandate and their style timing variables. The data period is January 2000 to December 2009. Newey-West corrected t-statistics are reported in parentheses.

Panel A : German funds

Quintile Sort on flows	α	β_m	β_{SMB}	β_{HML}	γ_m	γ_{SMB}	γ_{HML}
1. High	-0.2890 (-1.66)	0.8878 (45.47)	0.2961 (8.68)	0.0939 (3.37)	-0.0005 (-0.31)	-0.0033 (-0.53)	0.0017 (0.64)
2.	-0.3689 (-2.37)	0.8949 (49.87)	0.2884 (9.73)	0.0912 (3.63)	-0.0009 (-0.64)	-0.0132 (-2.33)	0.0025 (1.13)
3.	-0.7788 (-4.59)	0.8959 (45.82)	0.3533 (10.29)	0.0512 (1.79)	-0.0010 (-0.64)	0.0176 (2.80)	0.0050 (1.93)
4.	-0.3497 (-2.20)	0.8992 (49.41)	0.2747 (8.46)	0.0908 (3.40)	-0.0001 (-0.03)	-0.0130 (-2.25)	0.0050 (2.11)
5. Low	-0.3377 (-2.01)	0.8858 (45.48)	0.2938 (9.07)	0.0821 (2.98)	-0.0023 (-1.50)	-0.0071 (-1.19)	-0.0014 (-0.56)
High minus low (t-stat)	-0.0485 (0.40)	0.0020 (0.15)	0.0023 (0.10)	0.0118 (0.58)	0.0018 (1.67)	0.0038 (0.86)	0.0031 (1.69)

Panel B : European funds

Quintile Sort on flows	α	β_m	β_{SMB}	β_{HML}	γ_m	γ_{SMB}	γ_{HML}
1. High	-0.1904 (-1.47)	1.0182 (43.25)	0.3160 (7.68)	-0.1616 (-3.46)	-0.0043 (-1.31)	-0.0195 (-1.80)	0.0097 (1.00)
2.	-0.5537 (-3.90)	1.0412 (39.13)	0.3592 (7.53)	-0.3016 (-5.66)	-0.0017 (-0.50)	-0.0048 (-0.40)	0.0071 (0.66)
3.	-0.4958	0.9746	0.3047	-0.2336	-0.0023	0.0026	-0.0063

	(-3.26)	(35.52)	(6.16)	(-4.10)	(-0.60)	(0.20)	(-0.54)
4.	-0.3715 (-2.95)	1.0035 (46.05)	0.3318 (8.30)	-0.2461 (-5.19)	-0.0055 (-1.79)	0.0034 (0.31)	0.0136 (1.46)
5. Low	-0.4806 (-4.20)	0.9987 (48.75)	0.2511 (7.15)	-0.1263 (-3.02)	-0.0062 (-2.14)	-0.0118 (-1.24)	0.0111 (1.34)
High minus low (t-stat)	0.2902 (3.12)	0.0195 (1.17)	0.0649 (2.20)	-0.0353 (-1.04)	0.0019 (0.84)	-0.0078 (-0.98)	-0.0014 (-0.20)
Panel C : Global funds							
Quintile Sort on flows	α	β_m	β_{SMB}	β_{HML}	γ_m	γ_{SMB}	γ_{HML}
1. High	-0.2439 (-1.55)	1.0007 (34.04)	0.3322 (6.75)	-0.3543 (-6.42)	-0.0036 (-0.81)	-0.0060 (-0.67)	0.0316 (2.38)
2.	-0.3510 (-2.00)	0.9830 (30.35)	0.3653 (6.52)	-0.3782 (-6.53)	-0.0115 (-2.28)	-0.0132 (-1.32)	0.0290 (2.11)
3.	-0.2771 (-1.45)	0.9398 (26.79)	0.3508 (5.62)	-0.4643 (-7.00)	-0.0100 (-1.88)	-0.0110 (-0.99)	0.0377 (2.34)
4.	-0.3086 (-1.88)	0.9790 (31.57)	0.2868 (5.46)	-0.3820 (-6.89)	-0.0119 (-2.57)	-0.0090 (-0.99)	0.0288 (2.07)
5. Low	-0.4989 (-3.15)	1.0201 (33.77)	0.3195 (6.11)	-0.3647 (-6.87)	-0.0030 (-0.63)	-0.0205 (-2.32)	0.0299 (2.35)
High minus low (t-stat)	0.2551 (2.23)	-0.0194 (-0.89)	0.0127 (0.33)	0.0104 (0.25)	-0.0006 (-0.19)	0.0145 (2.09)	0.0017 (0.17)

Table 8. Bond Funds: Summary Statistics

This table provides summary statistics for bond mutual funds for the periods January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B). Data frequency is monthly. Statistics include the number (#) of funds used, the average number of observations per fund, the average return (% pm), the average standard deviation (% pm), the average alpha, the number (#) of statistically significant positive and negative alphas (based on bootstrap test statistics), the average betas on the Broad Investment Grade Index, the High Yield Index, the European Index and the Global Index, the average R-squared and the number of funds with non-normal residuals (using the Jargue-Bera JB test statistic, at a 5% significant level).

Panel A : January 1999 – December 1999				
	All	Germany	Europe	Global
# of Funds	215	51	82	82
Average # observations	68.0	7.7	60.3	75.9
Average Return	0.42	0.38	0.45	0.42
Average SD	1.59	1.02	1.52	2.00
Average alpha	-0.156	-0.177	-0.103	-0.196
# signif. pos./neg. alphas	1 / 70	0 / 26	1 / 22	0 / 22
Average β (Broad Invest. Grade Index)	-0.039	-0.039	-0.093	0.015
Average β (High Yield Index)	0.085	0.038	0.039	0.160
Average β (Europe Index)	0.467	0.546	0.593	0.288
Average β (Global Index)	0.130	0.079	0.144	0.149
Average $\bar{R}^2$	0.52	0.47	0.54	0.54
# signif. non-normal, BJ	52	11	16	25
Panel B: January 2000 – December 2009				
	All	Germany	Europe	Global
# of Funds	389	56	170	163
Average # observations	121.0	156.7	112.7	117.4
Average Return	0.29	0.33	0.32	0.24
Average SD	1.41	1.03	1.22	1.76

Average alpha	-0.115	-0.121	-0.084	-0.146
# signif. Pos./neg. alphas	2 / 153	0 / 35	0 / 57	2 / 61
Average β (Broad Invest. Grade Index)	-0.055	-0.042	-0.078	-0.034
Average β (High Yield Index)	0.096	0.047	0.056	0.153
Average β (Europe Index)	0.528	0.569	0.584	0.456
Average β (Global Index)	0.044	0.028	0.049	0.044
Average $\bar{R}^2$	0.52	0.50	0.52	0.54
# signif. Non-normal, BJ	157	34	71	52

Table 9. Bond Funds: Correlation Matrix of Factors

This table reports the pairwise correlation coefficients for the four bond indexes (factors). Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B). Only funds with at least 24 monthly observations have been included.

	Panel A : January 1990 – December 1999			**Panel B : Jan. 2000 – Dec. 2009**		
	Broad Investment Grade Index	High Yield Index	Global Index	Broad Investment Grade Index	High Yield Index	Global Index
High Yield Index	0.523			0.692		
Global Index	0.840	0.244		0.825	0.537	
European Index	0.099	-0.283	0.358	0.2979	0.058	0.498

Table 10. Bond Portfolio: Equally Weighted

This table provides summary statistics for the equally weighted portfolio of all bond funds for the period January 1990 to December 2009 for the four-factor bond model with market and style timing (squared) variables included. Data frequency is monthly and the reported t-statistics are Newey-West corrected bootstrap t-statistics. The $\bar{R}^2$ for the regression is 75%.

Parameters	Estimated Coefficient	t-stat
Alpha	-0.122	-3.4
β (Broad Invest. Grade)	0.001	0.06
β (High Yield)	0.087	8.4
β (Europe)	0.543	17.1
β (Global)	0.017	0.55
γ (Broad Invest. Grade)	0.008	2.17
γ (High Yield)	-0.001	-0.95
γ (Europe)	-0.004	-0.22
γ (Global)	-0.024	-3.6

Table 11. Bond Funds: Market Timing, Parametric and Non-Parametric Models

This table reports the percentage (number) of bond funds with significant timing coefficients in the parametric 4F Treynor-Mazuy $\gamma(TM)$ model and for the non-parametric timing coefficient, θ_m , for each of the 4 bond indexes. For the parametric model, individual tests are based on bootstrap (Newey-West) t-statistics and the significance level used is a 2.5% one-tail test. To assess the statistical significance of the fraction of funds which reject the null, we use a (binomial) t-statistic (Ferson and Chen 2014), which is reported below in square brackets. ***, ** and * denotes significance at the 0.5%, 2.5% and 5% level, respectively. Data periods are January 1990 to December 1999 (Panel A) and January 2000 to December 2009 (Panel B). Data frequency is monthly. Only funds with at least 24 monthly observations have been included.

Index	**Parametric timing: 4F model**		**Non-parametric timing**	
	$\gamma(TM) > 0$	$\gamma(TM) < 0$	$\theta_m > 0$	$\theta_m < 0$
Panel A : January 1990 – December 1999 (215 Funds)				
Broad Investment Grade	5.58 (12) [1.22]	3.72 (8) [0.48]	1.86 (4) [-0.25]	0.93 (2) [-0.62]
High Yield	0.47 (1) [-0.80]	5.58 (12) [1.22]	0.47 (1) [-0.80]	3.26 (7) [0.30]
European Government	1.86 (4) [-0.25]	3.72 (8) [0.48]	0 (0) [-0.99]	13.49*** (29) [4.33]
Global Government	0.47 (1) [-0.80]	11.16*** (24) [3.42]	0 (0) [-0.99]	30.70*** (66) [11.12]
Panel B : January 2000 – December 2009 (375 Funds)				
Broad Investment Grade	10.40*** (39) [4.98]	4.80 (18) [1.45]	9.87*** (37) [4.65]	2.13 (8) [-0.23]
High Yield	10.67*** (40)	11.20*** (42)	16.00*** (60)	2.93 (11)

	[5.15]	[5.49]	[8.51]	[0.27]
European Government	15.47*** (58) [8.18]	11.47*** (43) [5.65]	1.07 (4) [-0.90]	6.93*** (26) [2.80]
Global Government	20.00*** (75) [11.04]	20.27*** (76) [11.20]	0.80 (3) [-1.07]	8.53*** (32) [3.80]