Forthcoming, *International Review of Financial Analysis*

The Performance of US Bond Mutual Funds

Andrew Clare[a], Niall O'Sullivan[b], Meadhbh Sherman[b] and Sheng Zhu[b]

Abstract
We evaluate the performance of the US bond mutual fund industry using a comprehensive sample of bond funds over a long time period from January 1998 to February 2017. In this one study, we examine bond fund selectivity, market timing and performance persistence. We evaluate bond funds relative to their self-declared benchmarks and in terms of both gross-of-fee returns and net-of-fee returns. We document considerable abnormal performance among funds both to the fund (gross returns) and to the investor (net returns). Bond fund performance is found to be superior in the post financial crisis period. However, past strong performance cannot be relied upon to predict future performance. Finally, while some funds exhibit market timing ability; we find a predominance of negative market timing among US bond mutual funds.

Keywords: Mutual funds, bond funds, benchmark returns, market timing, persistence.
JEL classification: G11, G12, G14

[a] The Sir John Cass Business School, City University, London, UK.

[b] Cork University Business School and Centre for Investment Research, University College Cork, Ireland.

Corresponding author: Professor Andrew Clare, Cass Business School, 106, Bunhill Row, London. EC1Y 8TZ, UK. Email: a.clare@city.ac.uk

---

We are grateful for financial support from the Irish Research Council.

1. Introduction

Over the last decade or so an increasing proportion of investors, both institutional and retail, have switched their equity investments from actively managed funds to funds that track a financial market index. The gradual, but seemingly inexorable, switch from active to index-based equity funds has been led by US-based investors. At the end of 2016 the Investment Company Institute (ICI) reported in its 2017 Annual Report that \$2.6trn was invested across 421 index mutual funds. In the same report the ICI reported that net inflows into domestic equity index funds and ETFs (including reinvested dividends) was \$1.4trn in total between 2007 and 2016. Over the same period actively managed domestic equity funds haemorrhaged \$1.1trn.

There are a number of factors behind this shift. First, innumerable independent academic papers have shown that, on average, active equity managers offer poor value for money. Indeed, after examining the performance of active US equity funds, Fama and French (2010) concluded: "*In terms of net returns to investors, performance is poor*" (p. 1921). Second, investors have increasingly focused on value-for-money, and tracker funds tend to have far lower fees than active funds aiming to outperform the same tracked benchmark. Finally, though not exhaustively, investors and their advisors have recognised that asset allocation is far more critical to long-term investment performance than the choice of one active manager over another. In this context the active versus index-tracking decision becomes secondary. All of these and other factors, have led to the outflows from active equity managers to index equity funds, not just in the US but elsewhere too. For example, the UK's asset management trade body, The Investment Association, reported that 25% of £6.0trn managed by their members was managed on an indexed basis in 2016.

While there exists much academic evidence with regard to the performance of actively managed equity funds, far less attention has focussed on actively managed bond funds. US equity mutual funds comprise around 52% of the $16.3trn US mutual fund universe, but bond mutual funds comprise 22% of the total, receiving $2.0trn of net inflows and reinvested coupons, since 2007 (Source: ICI, 2017). The growing importance of this segment of the mutual fund industry would therefore seem to be worthy of independent scrutiny.

To this end, in this paper we conduct a comprehensive study of US bond mutual fund performance using a large sample of 884 funds over the period from January 1998 to February 2017. In contrast to the fragmented previous literature, we examine several aspects of bond fund performance over the same fund sample and time period. Our study has a number of distinguishing features.

First, we focus on bond funds that report a self-declared benchmark so that we can examine fund performance both in relation to an aggregate market index and also in relation to its own benchmark. Unlike past studies that attempt to assign an appropriate benchmark, inevitably with some error, the funds in our sample have self-declared benchmarks. Performance evaluation of the cross-section of funds based on own-benchmark-adjusted performance has two key advantages. First, it controls for investment constraints (restrictions on the bond holdings) that may vary across funds. If funds face investment constraints that are embodied in their benchmark but which are not common across funds, then a comparison of fund performance against a common benchmark is incomplete (see Clarke et al. (2002)). Kothari and Warner (2001) and Angelidis et al (2013) argue that standard mutual fund performance measures are unable to identify significant abnormal performance if the fund's style characteristics differ from those of the benchmark portfolio. Second, as highlighted by

Cremers et al. (2013) in the case of equity funds, many fund benchmarks have non-zero alphas when measured against a broad ‘market’ index such as the S&P500. That is, if a benchmark outperforms the S&P500, then passively tracking the benchmark may yield a positive fund alpha. Using funds’ benchmark-adjusted returns yields bias-adjusted alphas.

Second, we evaluate performance both gross and net of fund fees. We are interested in determining whether abnormal performance achieved by the fund manager gross of fees is also achieved by the fund investor net of fees. This is particularly important in active fund management where fees are generally higher compared to passive management and indeed are charged for the skill of the manager in ‘beating the market’.

Third, our paper contributes to a particularly small literature on market timing skill among bond fund managers, i.e., managers’ ability to correctly anticipate fluctuations in the aggregate bond market and to adjust fund holdings accordingly.

Fourth, we employ a large sample of bond funds over a long sample period that includes the financial crisis period from 2008 and we specifically examine the role of this crisis period in bond fund performance.

Finally, we also study whether past performance predicts future fund performance, i.e., the question of performance persistence.

To anticipate our results, we find considerable evidence of abnormal performance among US bond mutual funds in a single factor model when returns are gross of fees. While, unsurprisingly, this finding is diminished somewhat with the use of a multi-factor model and when returns are measured net of fees, it is not eliminated entirely and remains quite evident. We find evidence of market timing ability among some bond funds, although overall, negative market timing dominates in the sample. Finally, we find no evidence of economically significant performance persistence.

The rest of this paper is organised as follows. In Section 2 we review the relatively limited literature on bond mutual fund performance evaluation; in Section 3 we introduce and describe the data set that we use in the study; we report the results of our *ex post* and *ex ante* analysis of bond mutual fund performance in Section 4; and Section 5 concludes.

2. Literature Review

Despite their growing importance in investor portfolios, there are far fewer studies of bond fund performance than studies of equity fund performance[1]. Studies of bond mutual funds closely mirror those of equity funds in terms of evaluating risk-adjusted performance, style, selectivity, timing and the role of luck in performance. Bond fund studies employ similar performance attribution models and testing methods where single and multi-factor models are prevalent while in fewer cases, conditional performance models are also employed. The relation between fund return and fund characteristics, such as fund turnover, expenses, fund flow and size as well as fund relative performance persistence also feature in the bond fund literature.

A common approach to bond (and equity) fund performance evaluation is to compare the performance of actively managed funds with the performance of index funds (of comparable risk), that is, the estimation of factor models where the factors represent potential sources of

[1] For a review of the former see Elton and Gruber (2011), for the latter see, for example, Cuthbertson, Nitzsche and O'Sullivan (2010).

return and risk. We can classify this as *ex post* analysis of performance. In the case of bond funds, these factors may be interpreted as representing underlying risks in the economy such as term risk and default risk. Alternatively, the multi-factor models may be interpreted as performance attribution models. The main advantage of this latter approach is that it evaluates fund performance against a simple, feasible strategy that could be implemented by a fund manager. Single-index and/or multi-index performance alphas may then be estimated. Blake, Elton and Gruber (1993) is one such early study that employs both a broad market index of government and corporate bonds as well as more specific investment style benchmark indices. Performance alphas are found to be indistinguishable from zero – underperformance is found to be equal to the fees charged by the funds indicating that bond fund performance does not exceed that of the fund benchmark. Blake at al. also examine performance persistence using performance rank correlation tests over two sub-periods and report small rank correlations: past alpha performance does not forecast subsequent alpha performance. Evaluating bond funds against common risk factors including market, term and default risk, Choi and Kronlund (2017) study bond funds' 'reaching for yield' and its relationship to fund performance. In Fama-MacBeth (1973) tests, the authors document that bond funds that reach for yield produce higher returns, but that these are attributable to the risk factors.

The above studies do not depend upon a particular equilibrium model of security returns. Blake, Elton and Gruber (1995) develop an Arbitrage Pricing Theory (APT) model employing both fundamental economic variables as well as return indices to explain both returns and expected returns on bonds and bond mutual funds. In keeping with the requirements of APT, their model also employs forecasts (prepared by economists and investment professionals) to measure unexpected changes in the fundamental economic influences that affect returns. Bond returns are a function of (excess) stock market returns, default risk, term risk, unexpected changes in inflation and unexpected changes in economic performance as well as an index of aggregate bond returns and a measure of mortgage credit risk. The study finds negative and statistically significant net-of-fee alphas in all categories of bond funds examined including corporate, mortgage and government bond funds.

A further dimension of mutual fund performance is that of market timing ability, i.e., altering the sensitivity of the portfolio to an aggregate market index or benchmark in anticipation of future changes in that index or benchmark. Market timing ability among bond mutual funds has attracted little attention in the literature with few exceptions. With origins in the method of Treynor and Mazuy (1966), Chen, Ferson and Peters (2010) test for nonlinearities (timing ability) in the relation between bond fund returns and nine bond market factors – controlling for several non-timing related nonlinearities that could otherwise lead to spurious timing skill inferences. For example, the nonlinearity of a fund's own investment style benchmarks *vis-à-vis* the factors is shown to be an important control. Augmenting the Treynor and Mazuy specification, Chen at al. (2010) model the benchmark return as a non-linear function of the factor changes. They report overall neutral to weakly positive timing among individual bond funds after these non-timing related nonlinearity controls. Adapting the methodology applied by Wermers (2000) to equity funds, Cici and Gibson (2012) test characteristic timing ability among bond funds and again report neutral to weakly positive timing skill.

Adopting the performance evaluation approach of Daniel, Grinblatt, Titman, and Wermers (1997) as applied to equity funds, Moneta (2015) uses asset class weights, i.e., the proportions of the portfolio invested in different sectors, credit quality and maturity (rather than individual security weights) and calculates style, timing and selectivity performance.

This approach has advantages over returns-based analyses because it deals better with the non-linearities and option-like characteristics of bond funds. On selectivity skill, Moneta reports that bond fund managers demonstrate investment ability by holding securities that outperform their benchmarks but not by enough to cover their expenses and transaction costs. Using a data set of bond fund holdings and the method of Daniel at al. (1997), Cici and Gibson (2012) undertake a characteristic-based benchmark portfolio evaluation of bond fund selectivity; in this case based on duration and credit ratings characteristics. Cici and Gibson (2012) report a lack of evidence that bond fund managers can select corporate bonds that outperform other bonds of similar characteristics.

Within the equity mutual fund performance literature many researchers have examined the issue of performance persistence, i.e., the propensity for fund performance rankings to remain consistent over time (see, for example, the seminal papers of Carhart (1997) and Hendricks, Patel and Zeckhauser (1993)). There are few papers that investigate bond mutual fund performance persistence in the literature. From those that do, evidence of persistence is weak in earlier sample periods. Philpot, Hearth, Rimbey and Schulman (1998) and Philpot, Hearth, Rimbey and Schulman (2000) find short term persistence (over one year) in the relative performance (Sharpe ratios) of high yield, global issue and convertible funds based on contingency table tests. However, this finding is driven by funds ranked in the middle and lower end of the cross-sectional distribution of Sharpe ratios. Furthermore, the authors find no evidence of persistence over longer five year periods. A more recent study by Huij and Derwall (2008) however, does provide evidence of performance persistence in bond mutual fund returns. Using a model based on Elton et al. (1993), the authors carry out multiple persistence tests including rank correlation tests, contingency table tests, two-stage cross-sectional and time series tests of future alpha performance on past alpha performance in the spirit of Fama-MacBeth (1973) and the recursive portfolio formation tests of Hendricks et al.

(1993) and Carhart (1997). The authors use a comprehensive sample consisting of 3,549 funds spanning the period from 1990 to 2003 and document strong evidence of persistence across all four test procedures. The rank correlation tests, contingency table tests and two-stage tests reveal statistically significant persistence. The recursive portfolio formation tests, however, indicate that while holding period alphas are generally monotonically decreasing from top to bottom decile, and while the difference between top and bottom decile alphas is a significant 3% p.a., the alphas are generally negative. This indicates that while the persistence is statistically significant it is not of economic significance.

The relationship between bond fund performance on the one hand and fund characteristics such as fund size, trading activity, expenses, fees and fund flow on the other hand is a theme found in both equity and bond fund studies. Bond fund returns are found to benefit from economies of scale (e.g., where the fund belongs to a fund family) and are negatively related to fund fees (Philpot et al. (1998); Dowen and Mann (2004)). In the equity fund literature, there is a generally well established positive but convex relation between performance and fund flow: inflows are more sensitive to past good performance than outflows are to past poor performance (see Cuthbertson et al. (2010) for a review). The sparse literature on bond funds points to the opposite concave relation (Goldstein, Jiang and Ng (2017)). This paper provides evidence to suggest that this is due to the relative illiquidity of corporate bonds compared to equities and that the concavity is accentuated in corporate bond funds that hold illiquid assets and during times of corporate bond market illiquidity. Because of the relative illiquidity, investors in corporate bond funds are more exposed to a poor performance-investor redemptions-poor performance cycle than their counterparts in equity funds.

We provide a comprehensive, up-to-date study of a large sample of 884 funds over a longer sample period than is usually used to study bond fund performance, i.e., from January 1998 to February 2017. This period spans two major financial crises, both of which had an enormous impact on bond markets and bond fund management industry. We examine bond fund performance controlling for both the performance of the fund's self-declared benchmark as well as the performance of the aggregate bond market more generally. We also evaluate fund performance both gross and net of fund fees. Finally, we expand on the extant, small literature on both market timing skill and performance persistence among bond mutual fund managers.

3. Data

Our US bond mutual fund data are taken from Morningstar. This includes the monthly return both net and gross of fees on 884 surviving and non-surviving actively managed funds with self-declared benchmarks from January 1998 to February 2017. These funds were all domiciled in the USA and are US Dollar denominated. Returns relate to the oldest share class of the fund in order to avoid duplicate entries in the dataset. The fund sample, as categorised by Morningstar, are (i) US Short-Term Bond funds (178 funds), (ii) US Fund Intermediate-Term Bond funds (507 funds), (iii) US Long-Term Bond funds (20 funds) and (iv) US High Yield Bond funds (179 funds). The Short-Term category comprises bonds that have between one and three years to maturity; the Intermediate-Term category comprises bonds with three to ten years to maturity; and the Long-Term category comprises bonds with greater than ten years to maturity.

Although Barclays bond indices (formerly Lehman's indices) are the predominant benchmark indices, there is a wide range of self-declared benchmarks, even within each of the four categories. The benchmarks vary by maturity and by credit quality, each one carefully chosen by the fund management company to reflect the risk and return characteristics of their fund. In all we use 74 separate benchmarks for this study. In Table 1 we provide summary information for the 5 most popular benchmarks for each of the four fixed income categories along with the proportion of funds within that category that benchmark their funds to that index. For example, for the Intermediate-Term sector, the Bloomberg Barclays Global-Aggregate Total Return index (USD) is the benchmark for 63.9% of the funds in that section of the market.

As a proxy for the excess return on the 'market', to be used in a single and multi-factor models to risk-adjust bond fund returns, we use the Bloomberg Barclays Global-Aggregate Total Return Index and subtract from this a proxy for the risk free rate (R_f) based upon the one-month TBill yield. In our empirical work we also estimate a multi-factor model, where we add a measure of the term spread (TS) and credit spread (CS). Monthly TS is calculated by subtracting the TBill yield from the US ten-year Treasury yield. Monthly CS is calculated by subtracting the Aaa-rated corporate bond yield from the Baa-rated corporate bond yield, all data collected from the Federal Reserve.

4. Results

In this section, we present empirical results from estimating *ex post* risk-adjusted performance from single and multi-factor models based on both net and gross fund returns. We also examine the impact of the financial crisis period on bond fund performance. Next, we report on findings in relation to market timing ability among bond funds. Finally, we report on fund performance persistence.

4.1 Estimating bond fund alphas

We begin by estimating our baseline single index model for the full sample of funds. The model is of the following form

$$R_{pt} – R_{ft} = \alpha_{1p} + \beta_{1p}\times ERM_t + \varepsilon_{pt} \qquad [1]$$

where R_{pt} is the monthly return at time *t* on mutual fund *p* and R_{ft} is the monthly risk free rate, ERM_t is the excess return (over R_{ft}) in month *t* on the Bloomberg Barclays Global-Aggregate Total Return Index, α_{1p} represents Jensen's alpha for mutual fund *p*, β_{1p} is the market risk of fund *p* and ε_{pt} is a white noise error term. Results are presented in Table 2 where Panel A relates to gross (of fund fees) returns while Panel B relates to net returns.

In each panel we present results separately for the Morningstar bond fund categories of Short-Term, Intermediate-Term, Long-Term and High-Yield, that is 159, 458, 14 and 159 funds respectively. The statistics for the risk-adjusted performance of the 20 Long-Term funds in our sample should be interpreted with some caution, given the small size of this sample. For each category of funds we present a range of summary statistics: the cross-sectional average value of alpha and beta, the standard deviation (across funds) of the coefficient; the proportion of the estimates that are positive (%+) and negative (%-); and the proportion of the estimates that are positive or negative *and* statistically significant at the 95% level of confidence, denoted "% sig +" and "% sig –" respectively.

From Table 2 Panel A, the cross-sectional average alpha in each category of funds is small, although the vast majority of alphas across funds are positive. In the case of Short-Term funds, 60.4% of alphas within this investment style are positive and statistically significant.

Long-Term funds are an exception where a minority of funds exhibit positive alphas while this category of funds also has the largest percentage of negative and significant alphas at 21.4%. In Panel B where we analyse net-of-fee returns, unsurprisingly the cross-sectional distribution of alpha generally shifts to the left compared to gross return alphas and the percentage of positive (negative) alphas decreases (increases). However, in all categories of funds (except Long-Term funds) there remains a considerable proportion of funds that deliver positive and significant risk-adjusted performance even after fees. From the lower panels of Table 2, we can see that the market beta is positive and highly statistically significant across all investment styles, albeit slightly less so in the case of High-Yield funds. Overall, the findings in Table 2 indicate initial strong evidence of abnormal performance among the bond funds that is worthy of deeper investigation. In particular, we investigate the validity of the single factor model. If it is not appropriate for each of the four categories presented in Table 2, then the alphas in that table should be treated with caution.

4.2 A multi-factor model for bond funds

In this section we extend the previous analysis by augmenting the single-factor model with additional bond market-specific factors. These models allow us to calculate multi-factor alphas, but also provide additional information about the drivers of the returns generated by bond fund managers. In line with Choi and Kronlund (2017), Chen et al. (2010) and others, we add a factor to capture the impact of changes in the steepness of the yield curve (TS) and a factor designed to capture the reward for taking on credit risk (CS). As described previously, we calculate TS as the 10 year US Treasury yield minus the one-month TBill yield. We calculate CS as the yield on Baa rated corporate bonds minus the yield on Aaa rated corporate bonds. In addition, we specify the fund’s self-declared benchmark in the multi-factor model. As discussed previously, this controls for investment constraints that may vary across funds and enables a more valid fund performance comparison. It also allows us to estimate a benchmark-bias adjusted alpha, (i.e., where the fund's benchmark may have a non-zero alpha against the market factor). Our multi-factor model is of the form

$$R_{pt} - R_{ft} = \alpha_p + \beta_{1p} \times (R_{bt} - R_{ft}) + \beta_{2p} \times ERM_t + \beta_{3p} \times TS_t + \beta_{4p} \times CS_t + \varepsilon_{pt} \qquad [2]$$

where R_{bt} is the return on the benchmark of fund *p* at time *t*, TS_t and CS_t are the term spread and credit spread at time *t*. Our market factor is ERM_t, the excess return over month *t* on the Bloomberg Barclays Global-Aggregate Total Return Index. Some of the funds in our sample declare this to be their benchmark. In this section we limit our analysis to the 525 funds with a self-declared benchmark that do not have the Barclays Global-Aggregate index as a benchmark.

Results of the multi-factor estimation are reported in Table 3. Again Panel A presents gross-of-fee results while Panel B refers to net-of-fee results. Comparing the alpha results in each panel of Table 3 with the corresponding alpha results in Table 2, we see that the percentage of funds exhibiting positive and significant alphas is reduced in the multi-factor model across each investment style, except in the case of Long-Term funds where the percentage remains broadly unchanged. Nevertheless, all fund styles indicate a level of significant alphas that exceeds that which might be expected at the 5% significance level under the null hypothesis (note there are only 16 Long-Term funds in the analysis presented in Table 3). This finding continues to be evident even after fund fees.

The market index, ERM_t, exhibits strong statistical significance in all maturity categories of bonds but is less significant in the case of the High-Yield sector. The excess returns of those funds that have a positive β_3 coefficient are positively correlated with a steepening of the yield curve while those with a positive β_4 coefficient are positively correlated with a widening of credit spreads. There is some evidence that the credit spread factor has a stronger role to play in bond fund returns compared to the term spread, particularly in the case of Short-Term funds where 41.8% of these funds have a positive and significant loading on credit spread. However, 21.3% of High-Yield funds have a negative and significant loading on the credit spread factor.

Overall, even after controlling for the additional explanatory risk factors of term spread and credit spread, a notable percentage of funds continue to achieve a positive and significant alpha – this is as high as almost 14% of Short-Term funds, using gross-of-fee returns. This value falls to 11% when we use net-of-fee returns.

Given the scale of the crisis that befell financial markets from 2008 and the resulting volatility in corporate and government bond markets, it is interesting to examine the role of the crisis period on bond fund performance, i.e., it is possible that the crisis had an impact on manager skill. We explore this possibility by estimating the coefficient on a dummy variable within the previous multi-factor model framework. We estimate the following dummy variable-augmented model:

$$R_{pt} - R_{ft} = \alpha_p + \lambda_p \times D + \beta_{1p} \times (R_{bt} - R_{ft}) + \beta_{2p} \times ERM_t + \beta_{3p} \times TS_t + \beta_{4p} \times CS_t + \varepsilon_{pt} \qquad [3]$$

where D is a dummy variable that takes the value of 1 from January 1998 to September 2008 (a date that marks the beginning of the crisis period) and a value of zero over the remainder of the sample period to February 2017. A positive (negative) value for the coefficient on D,

λ_p, indicates that overall alpha performance was higher (lower) in the pre-crisis period than in the post-crisis period. In the interests of brevity, Table 4 presents summary statistics for the α_p and λ_p statistics only. As before, Panel A presents findings for gross returns while Panel B relates to net returns.

From Table 4, there is evidence that alphas are lower pre-crisis compared to post-crisis as indicated by the higher percentage of negative over positive significant lamda values (λ). This is particularly the case across all maturity sectors. The post-crisis zero interest rate environment is likely to have played a key role here where prices (yields) have generally been rising (falling).

4.3 Market Timing

In this section we turn our attention to the market timing ability of bond mutual funds. In particular, we explore the ability of bond fund managers to anticipate fluctuations in the aggregate bond market, as measured by the Barclays Global-Aggregate Total Return Index, and increase (decrease) the fund's market exposure in advance of higher (lower) market returns. We adopt the testing method originally proposed by Treynor-Mazuy (1966) which tests whether a fund's sensitivity to the market is greater in up-markets compared to down-markets. The Treynor-Mazuy model is estimated as follows

$$R_{pt} - R_{ft} = \alpha_{2p} + \beta_{2p} \times ERM_t + \gamma_p \times [ERM^2]_t + \varepsilon_{pt} \qquad [4]$$

In the Treynor and Mazuy model, a positive and significant value for γ_p may be interpreted as indicating that the manager has timing ability, i.e., increasing (reducing) exposure to market risk as the market rises (falls).

Our findings with respect to market timing ability are presented in Table 5. Panel A presents results for gross-of-fee returns while Panel B reports results for net-of fee returns. On average the timing statistics, $\gamma_{p,}$ are negative, indicating that on average the market timing decisions of managers tend to subtract rather than add value for investors. There is evidence that some Long-Term and Intermediate-Term funds are able to time aggregate bond market fluctuations where 12.6% and 14.3% of funds in these categories respectively have positive and statistically significant timing coefficients (considerably higher percentages than would be expected under a true null hypothesis in the industry). However, despite this ability among some funds, across the different investment style bond funds, perversely, the percentage of funds found to significantly negatively time the market far exceeds the percentage found to significantly positively time the market. This is especially true in the case of High-Yield funds. Long-Term funds are an exception where the percentages of positive and negative significant market timers are equal. These findings are mirrored in the net-of-fee returns. Overall, therefore, the evidence suggests that market timing within the US bond mutual fund industry has generally detracted value from fund performance over time.

4.4 Ex ante analysis of fund returns: performance persistence

In addition to the performance analysis above, it is also important to examine whether past relative performance can help predict future relative performance in the bond mutual fund industry, i.e., the question of persistence. This is a particularly under-explored question in the bond fund industry. There are many tests of fund performance persistence including contingency table tests, rank correlation tests as well as the recursive portfolio formation test of Carhart (1997) and others. A key advantage of the latter over the former is that it is a test of economic significance in fund persistence rather than just a test of statistical significance.

In the testing procedure, we sort funds into decile portfolios based on fund alphas estimated over a backward-looking formation period of 36 months where decile 1 contains the top sorted funds and decile 10 contains the bottom sorted funds. Alphas are estimated based on the following multi-factor model

$$R_{pt}–R_{ft} = \alpha_1 + \beta_1\times ERM_t + \beta_{2p}\times TS_t + \beta_{3p}\times CS_t \quad [5]$$

These decile portfolios are then held for a holding period of one month. In a separate test we repeat this procedure for a holding period of three months. The alpha of the decile portfolios holding period returns are then estimated as follows:

$$R_{dt} =\alpha_1 + \beta_1\times ERM_t + \beta_{2p}\times TS_t + \beta_{3p}\times CS_t \quad [6]$$

where R_{dt} are the holding period returns of each decile. We do not include the self-declared benchmark in [5] and [6] because the holding period decile returns are comprised of funds with different benchmarks. Statistically significant persistence is indicated where the alphas of the forward-looking (or holding period) deciles decline over deciles 1 to 10. Economic significance may be inferred from the sign and t-statistics of the forward- looking alphas.

The performance persistence results are presented in Table 6. The table shows the alpha and t-statistic of alpha of the forward-looking decile returns for one-month and three-month holding periods as indicated. Panel A reports the results of this procedure for gross-of-fee returns while Panel B reports results for net-of-fee returns. From Table 6 we see that for one-month holding periods, the forward-looking alphas are generally positive though not statistically significant at the 5% significance level according to the t-statistic. This indicates that over the sample period, following a strategy of investing in the past top performing funds would not have yielded a positive holding period abnormal performance. This is the case for both the fund manager (i.e., gross of fees) and for the fund investor (i.e., net of fees). The persistence findings are sensitive to the length of the holding period in the procedure. When we extend the holding period from one month to three months it is very evident that the holding period performance declines: all the holding period alphas are negative. We see that the deciles alphas towards the top and bottom are negative and statistically significant while the deciles in the middle are negative but not significant. This indicates that in the case of a three-month holding period there is negative persistence at the top end of the performance distribution. That is, past top-performing bond funds go on to perform relatively poorly in the following period. There is positive persistence at the bottom end of the performance distribution - past poor performing funds remain relatively poor performing in the future. Finally, funds in the middle of the performance distribution over the previous three years remain in the middle of the distribution over the following three months. Overall, our results fail to provide any evidence in support of the proposition that economically significant performance persistence exists among US bond mutual funds.

5. Conclusions

This paper contributes to a much-needed evaluation of the US bond mutual fund industry, which has attracted a dearth of attention compared to the much-studied equity fund industry and which fund flow data reveal is growing in importance. Although performance attribution models are imperfect in bond (and equity) studies, we find strong evidence in support of abnormal performance (alpha) in fund excess returns even after controlling for fund exposures to factors for systematic risk, term spread and credit spread. This abnormal performance is achieved by the fund manager gross-of-fees but, although reduced, it is also delivered to the investor net-of-fees. Our findings suggest that among bond funds as a whole, abnormal performance is superior in the post financial crisis period. While a small but significant proportion of funds exhibit an ability to time aggregate movements in the bond market, we find nevertheless that there is a preponderance of perverse negative over positive market timing among bond funds as a whole. We also conclude that investors should not rely on past positive bond fund performance as an indicator of future performance as there is no evidence of positive persistence. However, poorly performing funds in the past, particularly those in the extreme tail, should be avoided because this poor performance tends to persist – a finding remarkably consistent with similar findings for the equity mutual fund industry.

References

Angelidis, T., Giamouridis, D. and Tessaromatis, N. (2013), Revisiting mutual fund performance evaluation, *Journal of Banking and Finance*,37 (5):1759-1776

Blake, C., Elton, E. and Gruber, M. (1993). The performance of bond mutual funds, *The Journal of Business*, 66, 3, 371-403.

Blake, C., Elton, E. and Gruber, M. (1995). Fundamental economic variables, expected returns, and bond fund performance, *The Journal of Finance*, 50, 4, 1229-1256.

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

Chen, Y., Ferson, W. and Peters, H. (2010). Measuring the timing ability and performance of bond mutual funds, *Journal of Financial Economics*, 98, 72-89.

Choi, J. and Kronlund, M. (2017). Reaching for Yield in Corporate Bond Mutual Funds, *The Review of Financial Studies*, 31, 1930-1965.

Cici, G. and Gibson, S. (2012). The performance of corporate bond mutual funds: evidence based on security-level holdings, *Journal of Financial and Quantitative Analysis*, 47, 1, 159-178.

Clarke, R., de Silva, H. and Thorley, S. (2002). Portfolio constraints and the fundamental law of active management, *Financial Analysts Journal*, 58, 48-66.

Cremers, M., Petajisto, A. & Zitzewitz, E. (2013). Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review*, 2, 1-48.

Cuthbertson, K., Nitzsche, D., O'Sullivan, N. (2010). Mutual Fund Performance: Measurement and Evidence, *Journal of Financial Markets, Instruments and Institutions*, 19, 2, 95-187.

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

Dowen, R. and Mann, T. (2004). Mutual fund behavior, management behavior and investor costs, *Financial Services Review*, 13, 79-91.

Elton, E. and Gruber, M. (2011). Mutual Funds. Available at SSRN: https://ssrn.com/abstract=2088418 or http://dx.doi.org/10.2139/ssrn.2088418

Fama, E. and French, K. (2010). Luck versus skill in the cross-section of mutual fund returns, *The Journal of Finance*, 65,5,1915-1947.

Fama, E.F. and MacBeth, J. (1973), Risk, return and equilibrium: Empirical tests, *Journal of Political Economy*, 81, 607-636.

Goldstein, I., Jing, H. and Ng, D. (2017). Investor flows and fragility in corporate bond funds, *Journal of Financial Economics*, 126, 592-613.

Hendricks, D., Patel, J. and Zeckhauser, R. (1993). Hot hands in mutual funds: short-run persistence in relative performance, 1974-1988, *The Journal of Finance*, 48, 93-130.

Huij, J. and Derwall, J. (2008). “Hot Hands” in bond funds, *Journal of Banking and Finance*, 32, 559-572.

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

Moneta, F. (2015). Measuring bond mutual fund performance with portfolio Characteristics, *Journal of Empirical Finance*, 33, 223–242.

Philpot, J., Hearth, D., Rimbey, J. and Schulman, C. (1998). Active management, fund size, and bond mutual fund returns, *The Financial Review*, 33, 2, 115-126.

Philpot, J., Hearth, D., Rimbey, J. and Schulman, C. (2000). Performance persistence and management skill in nonconventional bond mutual funds, *Financial Services Review*, 9, 247-258.

Treynor, J. and K. Mazuy (1966), Can mutual funds outguess the market?, *Harvard Business Review*, Vol. 44, No. 4, pp. 131-136.

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

Table 1: Fund Benchmarks

This Table presents the top five stated fund benchmarks for the full sample of 884 funds, and for each sub-category: Short-Term, Intermediate-Term, Long-Term and High Yield. BBgBarc indicates that the index was constructed by Bloomberg-Barclays; BofAML indicates that the index was constructed by Bank of America-Merrill Lynch; "TR" indicates total returns; while USD indicates that the indices were denominated in US Dollars. All indices are available on Bloomberg and Morningstar.

All	**Proportion of funds benchmarked against index**
BBgBarc US Agg Bond TR USD	40.6%
BBgBarc US Govt/Credit Interm TR USD	7.8%
BBgBarc US Govt/Credit 1-3 Yr TR USD	6.4%
BBgBarc US Corporate High Yield TR USD	4.3%
BofAML US HY Master II TR USD	3.7%
Short-Term	
BBgBarc US Govt/Credit 1-3 Yr TR USD	32.0%
BofAML US Corp&Govt 1-3 Yr TR USD	10.7%
BBgBarc US Agg Bond TR USD	9.0%
BofAML US Treasuries 1-3 Yr TR USD	7.9%
BBgBarc US Govt/Credit 1-5 Yr TR USD	5.6%
Intermediate-Term	
BBgBarc US Agg Bond TR USD	63.9%
BBgBarc US Govt/Credit Interm TR USD	11.4%
BBgBarc US Govt/Credit TR USD	5.5%
BBgBarc US Credit TR USD	3.6%
BBgBarc US MBS TR USD	2.8%
Long-Term	
BBgBarc US Govt/Credit Long TR USD	25.0%
BBgBarc US Agg Bond TR USD	20.0%
BBgBarc US Long Credit TR USD	15.0%
BBgBarc US Govt/Credit TR USD	10.0%
BBgBarc US Credit TR USD	10.0%
High Yield	
BBgBarc US Corporate High Yield TR USD	21.2%
BofAML US HY Master II TR USD	18.4%
BBgBarc US HY 2% Issuer Cap TR USD	17.3%
BofAML US HY Master II Constnd TR USD	17.3%
BBgBarc US Agg Bond TR USD	8.4%

Table 2: Single-Factor Model Estimation of US Bond Mutual Fund Performance

This table presents results of the single factor model estimation in [1]. Panel A presents statistics for gross-of-fee returns while Panel B presents statistics for net-of-fee returns. Columns headed ST, IT, LT and HY present the results for funds in the Short-Term, Intermediate Term, Long-Term and High Yield Morningstar bond fund sectors. "Average" represents the average of an OLS parameter; "ST-Dev" represents the standard deviation of an OLS parameter; "% +" and "% -" represents the proportion of funds with a positive or negative OLS coefficient estimate respectively; "% sig +"and "% sig –" represent the proportion of funds that produce a positive and significant and negative and significant OLS coefficient respectively at the 5% significance level; 'Ave. Adj-R2' represents the average of the adjusted R-squared of each regression.

	ST	IT	LT	HY	ST	IT	LT	HY
	Panel A: Gross-of-fee returns $R_{pt}- R_{ft} = \alpha_{1p} + \beta_{1p}\times ERM_t + \varepsilon_{pt}$				**Panel B: Net-of-fee returns** $R_{pt}- R_{ft} = \alpha_{1p} + \beta_{1p}\times ERM_t + \varepsilon_{pt}$			
# of funds	178	507	20	179	178	507	20	179
$\boldsymbol{\alpha_{1p}}$								
Average	0.001	0.000	0.000	0.004	0.000	0.000	-0.001	0.003
St-Dev	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.003
% +	91.8%	76.9%	21.4%	95.6%	70.4%	48.8%	21.4%	93.1%
% sig +	60.4%	33.2%	7.10%	49.1%	22.0%	7.80%	0.00%	26.4%
% -	8.20%	23.1%	78.6%	4.40%	29.6%	51.2%	78.6%	6.90%
% sig -	1.30%	2.80%	21.4%	0.60%	7.5%	17.9%	28.6%	0.60%
$\boldsymbol{\beta_{1p}}$								
Average	0.386	0.958	1.786	0.401	0.387	0.959	1.786	0.401
St-Dev	0.209	0.230	0.713	0.239	0.209	0.230	0.713	0.239
% +	97.5%	99.8%	100%	96.2%	97.5%	99.8%	100%	96.2%
% sig +	92.5%	99.6%	100%	61.0%	92.5%	99.6%	100%	60.4%
% -	2.50%	0.20%	0.00%	3.80%	2.50%	0.20%	0.00%	3.80%
% sig -	0.00%	0.00%	0.00%	0.60%	0.00%	0.00%	0.00%	0.60%
Ave. Adj-R2	0.449	0.730	0.727	0.029	0.451	0.730	0.728	0.029

Table 3: Multi-Factor Model Estimation of US Bond Mutual Fund Performance

This table presents results of the multi-factor model estimation in [2]. Panel A presents statistics for gross-of-fee returns while Panel B presents statistics for net-of-fee returns. Columns headed ST, IT, LT and HY present the results for funds in the Short-Term, Intermediate Term, Long-Term and High Yield Morningstar bond fund sectors. "Average" represents the average of an OLS parameter; "ST-Dev" represents the standard deviation of an OLS parameter; "% +" and "% -" represents the proportion of funds with a positive or negative OLS coefficient estimate respectively; "% sig +"and "% sig –" represent the proportion of funds that produce a positive and significant and negative and significant OLS coefficient respectively at the 5% significance level; 'Ave. Adj-R2' represents the average of the adjusted R-squared of each regression.

	ST	IT	LT	HY	ST	IT	LT	HY
	Panel A: Gross-of-fee returns				**Panel B: Net-of-fee returns**			
	$R_{pt} - R_{ft} = \alpha_p + \beta_{1p}\times(R_{bt} - R_{ft}) + \beta_{2p}\times ERM_t + \beta_{3p}\times TS_t + \beta_{4p}\times CS_t + \varepsilon_{pt}$				$R_{pt} - R_{ft} = \alpha_p + \beta_{1p}\times(R_{bt} - R_{ft}) + \beta_{2p}\times ERM_t + \beta_{3p}\times TS_t + \beta_{4p}\times CS_t + \varepsilon_{pt}$			
α_p					**α_p**			
Average	0.001	0.001	0.002	0.010	0.001	0.001	0.001	0.009
St-Dev	0.005	0.005	0.005	0.022	0.006	0.005	0.005	0.022
% +	39.7%	50.0%	75.0%	84.0%	30.8%	38.0%	66.7%	78.0%
% sig +	13.7%	10.8%	8.30%	11.3%	11.0%	7.00%	8.30%	11.3%
% -	60.3%	50.0%	25.0%	16.0%	69.2%	62.0%	33.3%	22.0%
% sig -	13.0%	8.90%	0.00%	0.70%	28.1%	12.7%	0.00%	1.30%
β_{2p}					**β_{2p}**			
Average	0.398	0.988	1.715	0.241	0.400	0.987	1.714	0.240
St-Dev	0.203	0.320	0.701	0.177	0.205	0.320	0.700	0.178
% +	97.9%	100%	100%	92.7%	97.9%	100%	100%	92.7%
% sig +	97.3%	100%	100%	18.7%	97.3%	100%	100%	18.7%
% -	2.10%	0.00%	0.00%	7.30%	2.10%	0.00%	0.00%	7.30%
% sig -	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
β_{3p}					**β_{3p}**			
Average	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	-0.001
St-Dev	0.001	0.001	0.002	0.007	0.001	0.001	0.002	0.007
% +	47.9%	57.6%	33.3%	80.0%	48.6%	58.2%	33.3%	80.0%
% sig +	5.50%	7.00%	0.00%	12.7%	5.50%	7.00%	0.00%	12.7%
% -	52.1%	42.4%	66.7%	20.0%	51.4%	41.8%	66.7%	20.0%
% sig -	8.90%	4.40%	0.00%	8.70%	8.90%	3.20%	0.00%	8.70%
β_{4p}					**β_{4p}**			
Average	0.000	-0.001	-0.002	-0.007	0.000	-0.001	-0.002	-0.007
St-Dev	0.005	0.005	0.003	0.010	0.005	0.005	0.003	0.010
% +	69.9%	57.0%	8.30%	9.30%	69.9%	57.6%	8.30%	9.30%
% sig +	41.8%	10.8%	0.00%	2.70%	41.8%	11.4%	0.00%	2.70%
% -	30.1%	43.0%	91.7%	90.7%	30.1%	42.4%	91.7%	90.7%
% sig -	11.6%	8.20%	8.30%	21.3%	11.0%	8.20%	8.30%	20.0%
Ave. Adj-R2	0.529	0.720	0.734	0.292	0.532	0.721	0.734	0.293

Table 4: Bond Mutual Fund Performance: The Role of the Financial Crisis

This Table presents results of the estimation of [3]. Panel A presents statistics for gross-of-fee returns while Panel B presents statistics for net-of-fee returns. Columns headed ST, IT, LT and HY present the results for funds in the Short-Term, Intermediate Term, Long-Term and High Yield Morningstar bond fund sectors respectively.. "Average" represents the average of an OLS parameter; "ST-Dev" represents the standard deviation of an OLS parameter; "% +" and "% -" represents the proportion of funds with a positive or negative OLS coefficient estimate respectively; "% sig +"and "% sig –" represent the proportion of funds that produce a positive and significant and negative and significant OLS coefficient respectively at the 5% significance level.

	ST	**IT**	**LT**	**HY**	**ST**	**IT**	**LT**	**HY**
	Panel A: Gross-of-fee returns $R_{pt}– R_{ft} = \alpha_p + \lambda_p \times D + \beta_{1p} \times (R_{bt}– R_{ft}) + \beta_{2p} \times ERM_t + \beta_{3p} \times TS_t + \beta_{4p} \times CS_t + \varepsilon_{pt}$				**Panel B: Net-of-fee returns** $R_{pt}– R_{ft} = \alpha_p + \lambda_p \times D + \beta_{1p} \times (R_{bt}– R_{ft}) + \beta_{2p} \times ERM_t + \beta_{3p} \times TS_t + \beta_{4p} \times CS_t + \varepsilon_{pt}$			
$\boldsymbol{\alpha_p}$					$\boldsymbol{\alpha_p}$			
Average	0.002	0.002	0.001	0.010	0.001	0.001	0.000	0.009
St-Dev	0.007	0.007	0.010	0.023	0.007	0.007	0.010	0.023
% +	58.9%	68.4%	75.0%	73.3%	50.7%	58.2%	75.0%	68.0%
% sig +	19.9%	14.6%	8.30%	11.3%	15.1%	11.4%	8.30%	11.3%
% -	41.1%	31.6%	25.0%	26.7%	49.3%	41.8%	25.0%	32.0%
% sig -	5.50%	4.40%	0.00%	0.70%	15.8%	7.60%	0.00%	1.30%
λ_p					λ_p			
Average	-0.001	-0.001	0.000	0.000	-0.001	-0.001	0.000	0.000
St-Dev	0.007	0.004	0.010	0.007	0.007	0.004	0.010	0.007
% +	20.7%	18.9%	16.7%	64.3%	20.2%	18.0%	16.7%	63.4%
% sig +	0.90%	2.70%	0.00%	1.80%	0.90%	2.70%	0.00%	1.80%
% -	79.3%	81.1%	83.3%	35.7%	79.8%	82.0%	83.3%	36.6%
% sig -	26.1%	16.2%	16.7%	0.00%	26.6%	16.2%	16.7%	0.00%

Table 5: Market-Timing Ability Among US Bond Mutual Funds

This Table presents results of the estimation of [4]. Panel A presents statistics for gross-of-fee returns while Panel B presents statistics for net-of-fee returns. Columns headed ST, IT, LT and HY present the results for funds in the Short-Term, Intermediate Term, Long-Term and High Yield Morningstar bond fund sectors. "Average" represents the average of an OLS parameter; "ST-Dev" represents the standard deviation of an OLS parameter; "% +" and "% -" represents the proportion of funds with a positive or negative OLS coefficient estimate respectively; "% sig +"and "% sig –" represent the proportion of funds that produce a positive and significant and negative and significant OLS coefficient respectively at the 5% significance level; 'Ave. Adj-R2' represents the average of the adjusted R-squared of each regression.

	ST	IT	LT	HY	ST	IT	LT	HY
	Panel A: Gross-of-fee returns $R_{pt} - R_{ft} = \alpha_{1p} + \beta_{1p}\times ERM_t + \gamma_{1p}\times[ERM^2]_t + \varepsilon_{pt}$				**Panel B: Net-of-fee returns** $R_{pt} - R_{ft} = \alpha_{1p} + \beta_{1p}\times ERM_t + \gamma_{1p}\times[ERM^2]_t + \varepsilon_{pt}$			
# of funds	178	507	20	179	178	507	20	179
$\boldsymbol{\alpha_{1p}}$								
Average	0.001	0.001	0.000	0.006	0.001	0.000	-0.001	0.005
St-Dev	0.002	0.002	0.003	0.003	0.002	0.002	0.002	0.003
% +	91.2%	81.7%	42.9%	96.9%	73.0%	61.7%	28.6%	96.9%
% sig +	64.8%	51.5%	14.3%	78.6%	35.8%	24.2%	7.10%	66.7%
% -	8.8%	18.3%	57.1%	3.10%	27.0%	38.3%	71.4%	3.1%
% sig -	0.60%	2.40%	14.3%	0.00%	7.50%	14.2%	21.4%	0.60%
$\boldsymbol{\beta_{1p}}$								
Average	0.400	0.977	1.802	0.493	0.401	0.977	1.802	0.493
St-Dev	0.213	0.238	0.661	0.329	0.213	0.238	0.661	0.329
% +	98.1%	99.8%	100%	96.9%	98.1%	99.8%	100%	96.9%
% sig +	93.7%	99.6%	100%	70.4%	93.7%	99.6%	100%	70.4%
% -	1.90%	0.20%	0.00%	3.10%	1.90%	0.20%	0.00%	3.10%
% sig -	0.00%	0.00%	0.00%	0.60%	0.00%	0.00%	0.00%	0.60%
$\boldsymbol{\gamma_{1p}}$								
Average	-3.582	-5.004	-4.626	-22.695	-3.602	-4.958	-4.523	-22.733
St-Dev	7.518	7.960	14.813	20.779	7.497	7.891	14.812	20.856
% +	33.3%	24.9%	35.7%	5.00%	34.0%	24.6%	35.7%	5.00%
% sig +	12.6%	5.00%	14.3%	0.60%	12.6%	5.00%	14.3%	0.60%
% -	66.7%	75.1%	64.3%	95.0%	66.0%	75.4%	64.3%	95.0%
% sig -	32.7%	41.5%	14.3%	48.4%	34.0%	41.4%	14.3%	48.4%
Ave. Adj-R2	0.465	0.741	0.737	0.052	0.467	0.742	0.738	0.052

Table 6: Performance Persistence Among US Bond Mutual Funds

This Table presents results of performance persistence tests. The dependent variables are the monthly returns on the decile portfolios created by the recursive portfolio construction technique described in section 4.4. Funds are sorted into decile portfolios based on fund alphas estimated over a backward looking formation period of 36 months. In separate tests these deciles are held for holding periods of both one month and three months. The alpha of the decile portfolios holding period returns are then estimated. The table shows the alpha and t-statistic of alpha of these estimations for one month and three month holding periods as indicated. Panel A reports the results of this procedure for gross-of-fee returns while Panel B reports results for net-of-fee returns.

Panel A: Gross returns

	Holding period = 1m		Holding period = 3m	
Decile	α_1	t-α_1	α_1	t-α_1
1	0.0007	0.3214	-0.0052	-3.3410
2	-0.0006	-0.2324	-0.0046	-2.7203
3	0.0006	0.2405	-0.0032	-1.8453
4	0.0009	0.3716	-0.0026	-1.4579
5	0.0009	0.4045	-0.0023	-1.3228
6	0.0008	0.4238	-0.0015	-0.8017
7	0.0009	0.5430	-0.0019	-1.2795
8	0.0019	1.1580	-0.0013	-0.9002
9	0.0008	0.6468	-0.0023	-2.0845
10	0.0012	1.5036	-0.0015	-1.7019

Panel B: Net returns

	Holding period = 1m		Holding period = 3m	
Decile	α_1	t-α_1	α_1	t-α_1
1	-0.0009	-0.3550	-0.0062	-3.7579
2	0.0006	0.2153	-0.0042	-2.4329
3	-0.0009	-0.3396	-0.0045	-2.5416
4	0.0003	0.1167	-0.0032	-1.7018
5	0.0012	0.5667	-0.0022	-1.1808
6	0.0005	0.2916	-0.0020	-1.1726
7	0.0005	0.3099	-0.0022	-1.5037
8	0.0017	1.1988	-0.0017	-1.3650
9	0.0005	0.5406	-0.0026	-2.6544
10	0.0003	0.4125	-0.0022	-2.5118