Journal of Empirical Finance 16 (2009) 55–69

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Timing the investment grade securities market: Evidence from high quality bond funds ☆

Vaneesha Boney [a], George Comer [b,*], Lynne Kelly [c]

[a] University of Denver, USA
[b] Georgetown University, USA
[c] Howard University, USA

ARTICLE INFO

Article history:
Received 25 October 2006
Received in revised form 13 August 2007
Accepted 26 June 2008
Available online 2 July 2008

JEL classification:
G20
G23
G11

Keywords:
Portfolio evaluation
Fixed income funds
Market timing

ABSTRACT

We examine the ability of bond fund managers to shift assets between bonds and cash and across bonds of different maturities in order to capture the changes in their relative returns. As measured by estimated changes in portfolio allocations, we find strong evidence of perverse market timing ability between cash and investment grade securities, and our results indicate additional perverse timing across the bond maturity spectrum. Results are robust to an alternative performance metric. We present evidence that the survival of the majority of these funds despite their negative performance may reflect the value investors place on the portfolio diversification benefits of holding these funds.

1. Introduction

There is an extensive academic literature on the market timing ability of equity and hybrid mutual funds. Yet no published work has examined the timing ability of fixed income funds despite the fact that these funds have come to play a major part in many investors' portfolios, especially given the bear market in stocks that occurred during the early part of this decade.[1] According to data provided by Morningstar, the number of domestic nonmunicipal bond funds has grown by 210% (from 813 to 2517) over the ten year period covering 1994–2003.[2] At the end of 2003, the total amount of assets under management in these funds was just under $800 billion. Often, as individuals near retirement, investment experts suggest that investors shift a substantial portion of their portfolio into bond funds.

☆ This paper is a substantial revision of our earlier work entitled “Maturity Based Timing Skill: An Examination of High Quality Bond Funds.” We are grateful for comments from Edwin Elton, Wayne Ferson, Martin Gruber, Phyllis Keys, Javier Rodriguez, and an anonymous referee. We also thank seminar participants at Georgetown University, the University of Delaware, and the Washington Area Finance Association Conference for their suggestions. Comer acknowledges financial support from the Leo I. Higdon Jr. Endowment Fund from the Georgetown McDonough School of Business.
* Corresponding author. 417 Old North, Georgetown University, Washington, DC 20057, USA. Tel.: +1 202 687 0676.
E-mail addresses: vaneesha.boney@du.edu (V. Boney), gc45@georgetown.edu (G. Comer), ljkelly@howard.edu (L. Kelly).
[1] To the best of our knowledge, the only other studies that address bond fund timing ability are working papers by Chen, Ferson, and Peters (2005) and Comer (2005). Chen et al. examine timing skill with respect to factors related to bond returns. Comer focuses on sector timing skill. Both use extensions of the Treynor and Mazuy (1966) model.
[2] These numbers include all mutual funds classified by Morningstar as having one of the following portfolio objectives: adjustable rate mortgage, convertible bond, general corporate, high quality corporate, high yield corporate, general government, government mortgage, government treasury, or multi-sector bond.

0927-5398/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2008.06.005

Examining the timing ability of bond funds is of particular interest from an academic viewpoint. First, the majority of studies indicate that institutional money managers lack forecasting skill. However, a small but growing literature (e.g. Chance and Hemler, 2001; Bollen and Busse, 2001, 2004; Glassman and Riddick, 2006; Chen, 2005; Comer, 2006) suggest that various groups of managers have more market timing skill than previously believed, specifically when high frequency data and appropriate extensions of traditional timing models are employed. Of interest is whether bond fund managers share this forecasting skill. Second, studies by Blake, Elton, and Gruber (1993) and Ferson, Kisgen, and Henry (2006) indicate that various groups of bond funds underperform their relative indices on a risk adjusted basis. Thus, if the active management of bond fund managers is providing value to investors, the contribution is more likely to be reflected through timing ability rather than asset selection skill.

Various methods of measuring timing ability have been developed. Unfortunately, there is a lack of available bond data to implement holdings based measures of skill pioneered by Grinblatt and Titman (1989).[3] Thus, we infer timing skill from the funds' return series. The unique feature of our study is that we rely on the quadratic programming technique pioneered by Sharpe (1992) nested within the framework of the Henriksson and Merton (1981) models. Blake, Elton, and Gruber (1993) were the first to illustrate that Sharpe's methodology could be used to infer the investment policy of groups of bond funds, but the use of this technique to examine the performance and forecasting skill of money managers has been mostly confined to the hedge fund literature.

Sharpe's technique allows us to represent a manager's actual investment portfolio as a hypothetical portfolio of passively managed asset classes that best replicate the return series of the fund over time. Since Blake et al. (1993), there has been a substantial increase in the number and availability of bond market index return data. In Section 4 of this study, we develop index models that closely replicate individual fund performance and provide results that are consistent with the individual fund's reported investment policy. Thus, instead of inferring skill based on the convexity of fund returns relative to a benchmark factor model as done by Treynor and Mazuy (1966), we are able to provide numerical estimates of timing activity as measured by changes in fund manager allocations across various bond market conditions. We focus on two separate decisions made by the manager 1) the broad allocation between bonds and cash and 2) specific allocations across various bond maturities.

Bond fund managers can engage in timing activity in a variety of ways. In this study, we examine timing skill as measured by a manager's ability to time the investment grade securities market by shifting allocations to bonds of different maturities in order to capture changes in relative returns. Our focus on such skill is motivated by the significant differences in the returns generated by bonds of various maturities over our sample period.

A simple simulation illustrates the value of capturing differences in return across the bond maturity spectrum. We use monthly return data over the 1994–2003 time period from the Lehman 1–5 year, 5–10 year, and long (10+ year) government/credit bond indices. Suppose a manager has perfect timing ability and allocates 100% of his portfolio each month to the index with the highest return. Ignoring expenses and transaction costs, we find that the cumulative return to this strategy would be 379% compared to a −16% return for a manager who always invests 100% in the index with the lowest return. Given that it is very unlikely any manager would turn over his portfolio to such an extent, we conduct a second simulation. We assume that the manager must keep 75% of the portfolio in the 5–10 year index and can only shift 25% of his assets between the other two indices. In this case, a perfect market timer would generate a cumulative return of 151% compared to a return of 62% for a perverse timer. Clearly, the results are economically significant and are more than enough to cover the transaction costs that would be incurred.

We examine the timing ability of high quality corporate bond funds. According to the December 31, 2003 Morningstar Principia Pro CD, these funds have total net assets of $124 billion which represents approximately 17% of the total investment in the domestic nonmunicipal bond fund universe. Although this represents a subset of total bond fund investment, we focus on this group of funds because they are most likely to engage in timing activity by adjusting the average maturity of the fund. Unlike funds with a general corporate or general government classification, our sample of funds is restricted to investing the majority of its assets in investment grade issues. The main mechanism by which a manager within this group can distinguish himself is by capturing the differences in relative returns across various bond maturities rather than exploiting return differences due to credit quality.

We use Sharpe's methodology to examine whether changes in fund investment policy are consistent with the ability to positively time the investment grade securities market. Our results indicate that as a group our fund sample demonstrates statistically significant perverse timing ability as measured by higher allocations to cash when bonds of all maturities outperform Treasury bills. We compare our timing results to the results from synthetic fund samples which employ passive investment strategies. Evidence indicates that the perverse timing ability we detect is not spurious. Within the investment grade securities market, results indicate additional perverse timing across the bond maturity spectrum. Negative performance is widespread throughout the sample, and our results are robust to an alternative performance metric.

Given that 67% of our sample survives for the entire time period, our perverse timing results raises the question of the value provided to investors by high quality bond funds. Although we do not have a definitive answer to this issue, evidence from fund investor survey data, correlations between fund returns and the stock market, and the time series pattern of cash flows to the funds

[3] To the best of our knowledge, there is no source of complete bond fund portfolio holdings. The Thomson Financial/CDA Spectrum database used in the Grinblatt and Titman (1989) study does not include fixed income holdings. Some bond funds voluntarily provide portfolio data to Morningstar Principia, but the frequency and consistency of the disclosure is not constant. CRSP provides data on some bond portfolio attributes, such as average weighted maturity. But a check of this data series indicates that only 33% of our sample has a complete time series available.

Table 1
Descriptive statistics of high quality corporate bond fund sample

Number of funds	84
Average annual fund return	5.54%
Average annual return Lehman government/credit index	6.86%
Average total net assets	$523.9 million
Median total net assets	$133.0 million
Annual portfolio turnover	132.1%
Annual expense ratio	0.76%

The table presents descriptive statistics for the sample of high quality corporate bond funds over the sample period 1994–2003. All values are averages across all the funds in the sample. To obtain averages, we first averaged the time series data of each fund and then averaged across funds. Data is from the CRSP mutual fund database. All data represent annual values.

suggest that investors value these funds for the portfolio diversification benefits they provide despite their average negative performance.

The paper is organized as follows: Section 2 discusses our fund sample. Section 3 provides a description of our empirical methodology. Section 4 introduces the index model used to represent the return generating process of our fund sample. Empirical results and robustness tests are presented in Section 5 where we provide evidence of the timing performance of our funds. Sections 6 and 7 provide discussion of results and conclusions.

2. Fund sample

Our fund sample is composed of funds classified by the Morningstar Principia Pro CD as high quality corporate bond funds as of December 31, 1993. We restrict our sample to funds whose main means of generating superior returns is by shifting assets along the bond maturity spectrum. Thus, our focus is strictly on maturity based timing skill, and we make no attempt to measure bond selection ability or timing skill across sectors of the bond market. We start the sample in 1994 in order to have a sufficiently large sample of these funds to examine.

In order to be included in the sample, the fund must survive as a high quality corporate bond fund for a minimum of three years. This allows a long enough time series for our empirical methodology to be implemented. Once this minimum survival period is met, the fund is included in our sample, and we collect returns for the fund from the beginning of the sample period until the fund changes investment objective or until the fund ceases to exist. The fund's classification is checked annually using the December 31st Morningstar Principia CD for each year to verify that the fund remains a high quality bond fund.

The sample is constructed so that it is not subject to survivorship bias, but the sample does suffer from look-ahead bias. Some look-ahead bias is inherent in tests of market timing ability where a minimum survival period is necessary in order to have a sufficient amount of data for the estimation procedure. As discussed in Carhart, Carpenter, Lynch, and Musto (2002) imposing a minimum survival period can potentially bias inferences drawn from the sample because the funds that do not meet the survival criteria tend to be the poorer performing funds. However, the results to be presented in Section 5 indicate that our sample of funds have poor timing performance. Assuming nonsurvivors ceased to exist due to poor performance, the look ahead bias likely causes us to understate the poor timing skill of the funds. Our conclusions concerning perverse ability are not likely to change if the bias did not exist.

Monthly fund returns are obtained from the Center for Research in Security Prices (CRSP) mutual fund database. Goetzmann, Ingersoll, and Ivkovic (2000), and Bollen and Busse (2001) provide evidence that the use of monthly data may fail to detect timing ability if timing decisions occur at a more frequent interval, and all suggest that higher frequency data is preferable when employing the traditional timing specifications. However, unlike stock funds, the data and information necessary to compute a complete daily total return series for bond mutual funds is not readily available, and we must rely on monthly data.[4] For this study, our use of monthly data does not appear to represent a significant problem. In the empirical section, we provide evidence that our methodology has enough power to reject the null hypothesis of no timing ability across the individual funds in our sample.

Table 1 provides descriptive statistics of the funds. Our final sample is composed of 84 high quality corporate bond funds. The data for the fund net assets, portfolio turnover ratio, and expense ratio are the annual numbers reported by CRSP. In cases where the annual data are not available from CRSP, we use the values provided by the Morningstar Principia database.[5] To obtain the averages reported in the table, we first average over the time series of each fund and then average across funds.

The average portfolio turnover ratio for our fund sample is 132.1%. Turnover ratios for bond funds reflect both active portfolio changes by management and the replacement of maturing bond issues. Thus, we need to compare this ratio to the average turnover of bond index funds to determine if the managers in our sample are engaged in actively managing the funds. Over the 1995–2003 period, three comparable bond index funds were available from Vanguard. During this period, the Vanguard Short Term, Intermediate Term, and Long Term index funds have much lower average turnover ratios of 99.4%, 89.7%, and 69.7%

4 Morey and O'Neal (2006) discuss the problems involved with constructing a daily total return series for bond funds. Bond funds have the choice of several different methods for reporting net asset values and distributions. Most bond funds treat income and capital gains distributions differently. In order to calculate a complete daily return series, one must have knowledge of each fund's accounting method.

5 Turnover data are not reported by CRSP nor Morningstar for one of the funds in our sample.

respectively. Thus, the turnover ratio of our sample suggests that the managers are actively trading and tend not to follow a buy and hold strategy.

At first glance, this active management does not result in superior performance. On average, the funds underperformed the Lehman government/credit bond index which is one of the typical benchmarks used to evaluate the performance of high quality corporate bond funds.[6] The Lehman index returned 6.86% annually compared to the 5.54% average return for the funds. Even when expenses are considered, the funds underperformed this benchmark. Only 0.76% of the 1.32% underperformance can be directly attributed to the average expenses charged by these funds. We will turn to more formal tests of timing ability in Section 5 of the paper.

3. Methodology

To examine timing ability, we employ the quadratic programming technique of Sharpe (1992) to infer the investment policy of the fund throughout our estimation period. By using this technique, we are able to provide numerical estimates of changes in allocations across various bond market conditions rather than inferring timing skill from the statistical significance of the coefficients from a factor model.

To apply this technique, we assume the monthly return for each fund can be represented by a k factor model:

$$r_{p,t} = \sum_{i=1}^{k} b_{pi} r_{i,t} + e_{i,t} \tag{1}$$

where

$r_{p,t}$ total return of fund p
b_{pi} exposure of fund p to asset class i
$r_{i,t}$ total return of asset class i
$e_{i,t}$ unexplained component of fund return.

For any individual fund, the average exposures to each asset class during a given observation window can be estimated by solving the following quadratic program:

$$\begin{aligned} &\min \quad \left[\operatorname{var}\left(r_{p,t} - \sum_{i=1}^{k} b_{pi} r_{i,t}\right)\right] \\ &\text{subject to} \\ &\quad 0 \le b_{pi} \le 1 \quad \forall\, i \\ &\quad \sum_{i=1}^{k} b_{pi} = 1 \end{aligned} \tag{2}$$

In this formulation, the b_{pi} can be interpreted as the positive portfolio weights on the passive indices. Thus, we approximate the fund's actual portfolio allocation over any given estimation window with weights on passive market indices that represent the various asset class choices available to the manager. The estimated b_{pi}'s represent the average weights for the indexing strategy that best explains the fund return series over the estimation period.

This technique has been used to infer the investment philosophies of a wide variety of money managers. In their examination of bond fund alphas, Blake, Elton, and Gruber (1993) use the quadratic programming technique to compare the estimated weights from the procedure with annual fund composition data for both general and high quality bond funds. They find a strong correspondence between the quadratic programming weights and the reported annual sector weights for the bond funds in their sample, and they conclude that the technique can be used to infer management's sector allocations. This technique has been employed most extensively in the hedge fund literature (e.g., Fung and Hsieh, 1997; Brownet al., 1999; Brown et al., 2000; Agarwal and Naik, 2000; Dor et al., 2003) given the lack of publicly available data on hedge fund allocations and investment strategies. Most recently, Comer (2006) uses the technique to estimate the average stock allocation of hybrid funds during up and down markets. He finds that the conclusions about timing skill are highly correlated with a properly specified multifactor version of the Treynor and Mazuy (1966) model.

Once we identify an appropriate index model and establish that our proposed model and data meet the requirements specified by Sharpe (1992), we can use the fund return data to estimate manager allocations over various bond market conditions. We address these issues in the following section.

[6] In Section 4, we provide evidence in support of our use of the government/credit index rather than the credit index as an appropriate benchmark for our sample of funds.

4. An appropriate index model

There is no widely accepted fixed income equivalent to the Carhart (1997) four factor model. We propose an index model based on the total return indices available from Lehman Brothers. The Lehman Brothers indices are the most comprehensive market weighted indices available and are the most widely used family of indices for bond portfolio evaluation. We are interested in an index model that represents bonds of various maturities, and Lehman provides government/credit indices partitioned into the following three maturity groupings based on the bonds' final maturity dates: 1) 1–5 years to final maturity, 2) 5–10 years to final maturity, and 3) 10 or more years to final maturity. Lehman also provides an equivalent set of maturity indices based only on their credit index. Although we are examining high quality corporate funds, we choose the broader government/credit maturity indices rather than the credit indices for our index model. According to 2003 portfolio survey data available from Morningstar, 78.4% of all high quality corporate bond funds report that they invest in US government and agency bonds in addition to corporate bonds. As a result, we feel that the broader indices are most appropriate for this study.

Thus we propose the following index model:

$$r_i = b_{i,\mathrm{tb}} r_{\mathrm{tb}} + b_{i,\mathrm{gc0105}} r_{\mathrm{gc0105}} + b_{i,\mathrm{gc0510}} r_{\mathrm{gc0510}} + b_{i,\mathrm{gclong}} r_{\mathrm{gclong}} + b_{i,\mathrm{hiyld}} r_{\mathrm{hiyld}} + b_{i,\mathrm{mort}} r_{\mathrm{mort}} + e_i \tag{3}$$

where

r_i	total return for fund i
r_{tb}	total return on the short Treasury index
r_{gc0105}	total return on the 1–5 year government/credit bond index
r_{gc0510}	total return on the 5–10 year government/credit bond index
r_{gclong}	total return on the long (10+) government/credit bond index
r_{hiyld}	total return on the corporate high yield bond index
r_{mort}	total return on the fixed rate mortgage backed security index.

The short Treasury index consists of US Treasury bills with a remaining maturity from one to twelve months and is used to represent the cash investment of the funds.

Although we are interested in timing across the bond maturity spectrum, we include two non-maturity based indices in the model to control for additional factors that likely affect bond returns. As discussed in Elton, Gruber, and Blake (1995), options are an important element of bond returns. We follow their work by including the Lehman mortgage backed securities index in the model to proxy for this effect since options have an important impact on mortgage returns. We also include a high yield index. Work by Comer (2006) and Comer, Larrymore, and Rodriguez (in press) indicate that a high yield index is a significant factor in explaining the returns of the bond portion of hybrid mutual funds despite the fact that most hybrid fund bond portfolios are rated as high quality by Morningstar.

Sharpe (1992) specifies that the technique works optimally under the following three conditions: 1) the asset classes included in the index model are mutually exclusive, 2) the asset classes exhaust the investment set of the fund manager, and 3) the return series of the included indices are weakly correlated or may be highly correlated but have significant differences in variance. As discussed below, the indices included in our proposed model meet each of these conditions. The government/credit maturity indices are constructed such that no security is included in more than one of the indices. In addition, short Treasury securities, mortgage backed securities, and high yield bonds are considered separate classes of fixed income instruments. The securities included in those indices do not overlap with each other or with the securities in the government/credit indices. Together, the six indices cover the typical investment set of a domestic nonmunicipal bond fund.

Panel A of Table 2 provides descriptive statistics of the monthly index returns used in the model while Panel B lists the correlation matrix of the six indices. The short Treasury and high yield bond indices are weakly correlated with each other and with the remaining four indices, thus fitting Sharpe's criteria. But the government/credit maturity indices along with the mortgage backed index are highly correlated with correlations ranging from .81 to .95.

Given these high correlations, we test for significant differences in the standard deviations across the four indices using Levene's homogeneity of variance test. Panel C of Table 2 reports the test statistic which has an F distribution with $k-1$ and $N-k$ degrees of freedom where k represents the number of indices and N represents the total number of observations across both indices. All p values are less that 0.02 so we reject the null hypothesis of equal variance for all pairs of indices that we test. Although some of our indices are highly correlated, there is statistically significant evidence that their standard deviations differ. This difference allows the quadratic program to estimate weights across the highly correlated indices.

Given that the index return data meet the necessary criteria, the main issue is whether this model effectively represents the return generating process of our sample of high quality bond funds. To consider the model appropriate for measuring timing ability, the results from the quadratic program should demonstrate that the funds have significant exposure to the portfolios defined in the model, and the model should have substantial explanatory power.

To test the model, we create a fund of funds by averaging all fund returns during a given month. Using this time series of returns, the quadratic program defined in Eq. (2), and the index model defined in Eq. (3), we estimate the quadratic program for this equal weighted portfolio. The Gauss–Newton method is used to obtain parameter estimates that minimize the objective function. The statistical significance of the weights is based on t tests which use the estimated standard error of the coefficients. Because of the nonlinear

Table 2
Descriptive statistics for bond index returns

Panel A — return statistics

Bond index	Mean	Standard deviation
Short treasury	0.37	0.15
Government/credit 1–5 year	0.51	0.68
Government/credit 5–10 year	0.61	1.55
Long government/credit	0.67	2.38
High yield	0.60	2.13
Mortgage backed	0.56	0.89

Panel B — correlation matrix

	Short treasury	Government/ credit 1–5	Government/ credit 5–10	Long government/credit	High yield	Mortgage
Short treasury	1					
Government/credit 1–5	0.35	1				
Government/credit 5–10	0.22	0.94	1			
Long government/credit	0.18	0.85	0.95	1		
High yield	−0.04	0.10	0.21	0.25	1	
Mortgage	0.34	0.90	0.87	0.81	0.19	1

Panel C — Levene's homogeneity of variance test

Pair of indices tested	*F* statistic	*P* Value
Government/credit 1–5 vs government/credit 5–10	25.21	<.0001
Government/credit 1–5 vs long Long government/credit	29.25	<.0001
Government/credit 1–5 vs mortgage	5.56	.0191
Government/credit 5–10 vs long government/credit	10.02	.0018
Government/credit 5–10 vs mortgage	15.98	<.0001
Long government/credit vs mortgage	25.30	<.0001

The table presents summary statistics for the bond return indices used in the Sharpe style analysis regressions. All of the bond indices and corresponding return data are obtained from Lehman Brothers. The index return data are monthly and covers the ten year time period from 1994–2003. Panel A lists the six indices and their mean monthly return and standard deviation. Panel B lists the correlation matrix across the six indices. A requirement of the Sharpe style regressions is that the indices either 1) have a low correlation with each other or 2) have a high correlation but different standard deviations. Given the high correlations in Panel B, Panel C lists the results of the Levene's homogeneity of variance test across pairs of indices that have correlations greater than .8. The test statistic has an F distribution with $k-1$ and $N-k$ degrees of freedom where k represents the number of indices and N represents the number of observations across both indices.

nature of the estimation, these t tests are only valid asymptotically. Using a Monte Carlo simulation, Lobosco and DiBartolomeo (1997) illustrate that Sharpe estimates are approximately normally distributed, and the traditional t test is a valid approximation.

Results from the estimation are presented in Panel A of Table 3. Despite the constraints imposed by the quadratic program, the method explains nearly 99% of the variation in returns. In addition, the average fund return has a statistically significant nonzero exposure to all of the indices included in the model. The portfolio has the greatest average exposure to the short maturity bond index with an estimated portfolio weight of 30.1%. The lowest average exposure of 1.8% belongs to the high yield index, and the statistical significance of the coefficient indicates the importance of including this index in the model.[7] The results indicate the model is an appropriate representation of the return generating process of an equally weighted portfolio of our sample of bond funds.[8]

By forming a portfolio of funds, we eliminate idiosyncratic risk and generate a much higher coefficient of variation than is likely to be found for the funds individually. Since our empirical tests will be based on the individual fund weights, we also need to examine the individual R^2 to determine the usefulness of our index model. We run the program for each fund. Across the individual funds, the average R^2 is .88 with a median of .92. Over 60% of funds have an R^2 greater than .9 while only two funds have an R^2 of lower than .7. This distribution of the individual R^2s provides additional support for the validity of the model.

Using annual portfolio information from Morningstar, we conduct one additional test of our index model. Morningstar reports a fixed income style based on data they obtain from fund company surveys. They classify funds based on portfolio quality and average weighted maturity. There are three maturity classifications. The short maturity classification includes funds whose portfolios have an average weighted maturity of four years or less. Intermediate funds have an average maturity of four to ten years. Long maturity funds have a maturity of greater than ten years.

Using the December 31st Morningstar Principia CDs over the 1994–2003 time period, we obtain the reported fixed income style for each fund in the sample. We assign a value of one for each annual short maturity classification, a value of two for each intermediate classification, and a value of three for each long maturity classification. Using this time series, we calculate an average maturity style for each fund, and we group funds based on their average style. Funds with values less than 1.5 are classified as short

[7] During our sample period, typical SEC rules require that 65% to 80% of the assets of high quality corporate bond funds are invested in investment grade securities. Thus, detecting some exposure to low quality debt securities is not unexpected.

[8] We also run the quadratic program substituting the credit indices for the government/credit indices. There is a drop in the explanatory power of the regression and the short and long bond maturity indices are no longer statistically significant. This result provides additional support for the use of the broader government/credit maturity indices.

Table 3
Sharpe style analysis

Panel A — equal weighted portfolio all funds	
Portfolio	Estimated portfolio percentage
Short treasury	26.2%*
Government/credit 1–5 year	30.1%*
Government/credit 5–10 year	16.9%*
Long government/credit	9.2%*
High yield	1.8%*
Mortgage backed	15.7%*
Number of observations	120
R-square	.987
Adjusted R^2	.986

Panel B — fund groups on morningstar classification			
	Estimated portfolio percentage		
Portfolio	Short maturity	Intermediate maturity	Long maturity
Short treasury	43.6%*	6.0%	21.9%*
Government/credit 1–5 year	44.4%*	27.1%*	0.0%
Government/credit 5–10 year	5.6%	26.6%*	29.8%*
Long government/credit	1.1%	14.4%*	29.2%*
High yield	0.3%	3.1%*	7.2%*
Mortgage backed	4.7%*	22.6%*	11.6%*
Number of funds	35	40	8
Number of observations	120	120	120
R^2	.980	.988	.977
Adjusted R^2	.979	.988	.977

The table presents the results from the Sharpe style analysis technique that is used to estimate the average percentage of assets that various partitions of our high quality bond fund sample allocate to the various bond asset categories. Panel A presents results for an equal weighted portfolio of all funds in the sample. Panel B presents results for equal weighted portfolios of funds groups on their Morningstar classification. Within each partition, a time series of the average monthly returns across the funds is calculated. This return is regressed against the returns of the six bond and cash portfolios listed in the table which are represented by Lehman Brothers bond indices that are detailed in the text. In the estimation, each coefficient for each portfolio is constrained to be between zero and one, and the sum of all six coefficients is constrained to equal one. Each coefficient estimate represents the average percentage of assets allocated to the specific asset category. *denotes portfolio percentages that are significantly different from zero at the five percent level.

funds. Those with values between 1.5 and 2.5 are classified as intermediate funds. Funds with values greater than 2.5 are long funds. The resulting distribution of funds is 35 short maturity funds, 40 intermediate maturity funds, and 8 long maturity funds.[9]

We calculate a time series of monthly returns for an equal weighted portfolio of funds in each of the three groups. We then estimate the quadratic program for each group. It is important to note that the Morningstar classifications do not directly correspond to the maturities of the bond indices used in the index model. However, there still should be certain patterns in the results if the model is an appropriate representation of the return generating process of these funds and assuming the Morningstar classifications accurately reflect the average maturity of the securities held by the fund.

The results are presented in Panel B of Table 3 and are consistent with the Morningstar classifications. Funds classified as short maturity funds have the largest average investment weights in the short Treasury and the 1–5 year government/credit indices. The average total investment weight in these two portfolios is 88%. These funds do not show any statistically significant investment exposure to the 5–10 year or long government/credit indices. At the other end of the maturity spectrum, long maturity funds have the largest exposure (29.2%) to the long government/credit bond portfolio. Nearly 60% of the long maturity funds' portfolio is invested in the 5–10 year and long maturity bond indices, and unlike the short maturity funds, these funds have no exposure to the 1–5 year bond portfolio. The results for the portfolio of intermediate funds fall between the results of the short and long funds. Intermediate funds have a smaller exposure to the 1–5 year bond portfolio than the short funds (27.1% vs 44.4%). They also have a smaller exposure to the long bond portfolio than the long funds (14.4% vs 29.8%). But this group has the largest combined exposure (53.7%) to the 1–5 year and 5–10 year bond portfolios.

These results provide further support for our model and the quadratic programming technique which we will use in the next section to examine timing behavior.

5. Empirical results

In this section, we conduct four sets of tests. First, we examine timing ability between cash and bonds. Second, we focus on timing skill across bond maturities. Third, we test whether the results are robust to an alternative performance metric. Lastly, we examine whether estimated ability is correlated with observable fund characteristics.

[9] For three funds in our sample, Morningstar never reports a maturity classification across the entire life of the fund. We are able to classify two of these funds based on the name of the fund which never changes over the sample period and clearly indicates the average maturity of the bonds held by the funds. There is no way to determine the average maturity of the portfolio of the remaining fund and that fund is not included in the sample for this specific test.

5.1. Timing between cash and bonds

Our empirical tests of timing ability are similar to the portfolio allocation analysis methodology used by Comer (2006). The concept behind portfolio allocation analysis is that a fund which demonstrates positive (perverse) timing ability will have a higher (lower) average percentage of the portfolio allocated to cash during months when cash returns are higher (lower) than bond returns. We partition the data based on the returns to bonds and cash, estimate portfolio allocations across the partitions, and then examine the average portfolio weights held by the fund within each partition.[10]

Our initial test focuses on traditional market timing by examining the ability of the funds to switch between cash and bonds. We define the short Treasury index as our timing portfolio, and we measure ability relative to this portfolio. We partition the data into two groups as follows:

$$S_1 : r_{\mathrm{tb}} = \max\left(r_{\mathrm{tb}}, r_{\mathrm{gc0105}}, r_{\mathrm{gc0510}}, r_{\mathrm{gclong}}\right) \tag{4}$$

$$S_2 : r_{\mathrm{tb}} = \min\left(r_{\mathrm{tb}}, r_{\mathrm{gc0105}}, r_{\mathrm{gc0510}}, r_{\mathrm{gclong}}\right) \tag{5}$$

The first set contains all observations where the total return of the short Treasury index is greater than the return of the three bond maturity indices. The second set contains all observations where the short Treasury index is the worst performing portfolio. Assuming a fund survives for the entire time period, S_1 contains 43 monthly observations while S_2 contains 60 monthly observations.

We employ Sharpe's quadratic program as defined in Eq. (2) and the six index model from Eq. (3) to estimate each fund's exposure to the short Treasury index across the two partitions.[11] We then calculate each fund's average cash adjustment as

$$\text{cash adjustment} = b_{i,\mathrm{tb}}(S_1) - b_{i,tb}(S_2) \tag{6}$$

A positive value for the cash adjustment variable potentially indicates successful timing skill. However, because of the variation in bond market returns, we expect that passive investment strategies would also result in nonzero values of the cash adjustment measure. Thus, to draw any conclusions about timing skill, we must compare the average cash adjustment variable of our fund sample to the average cash adjustment variable of similar funds which employ passive investment strategies.

For each fund in our sample, we construct synthetic funds that have similar characteristics to our fund sample, yet the synthetic funds follow various passive investment strategies. In Section 4, we estimated each fund's average exposure to each of the six indices. As listed in Appendix A, each of the six indices from our model, with the exception of the mortgage index, is composed of two Lehman Brothers subindices. We assume each synthetic fund follows an investment strategy where the fund's initial exposure to each broad index is represented by a random set of investment weights on each subindex where the subindex weights w_i and w_j satisfy $0<w_i$, $w_j<1$ and $w_i+w_j=1$.[12]

Then using each of our actual funds' average exposure to the broad indices, the subindex weights, and the time series of subindex returns, we generate three synthetic funds for each individual fund in our sample. The first synthetic fund is designed to have the same unconditional exposure to the broad indices as the individual funds. Thus, we assume the synthetic fund follows a fixed weight strategy with monthly rebalancing over the entire time period where the monthly investment weights always equal the average weights of the matching fund. The investment weights for the second synthetic fund are initially set to the average weights of the corresponding fund, but evolve over the period according to a buy and hold strategy. The third synthetic fund is assumed to follow a buy and hold strategy with annual rebalancing back to the initial investment weights. For each of the three passive strategies, we repeat this process one hundred times by generating one hundred combinations of subindex investment weights.

This procedure allows us to create synthetic funds with similar characteristics to our actual fund sample since the initial broad index investment weights reflect the investment strategies employed across our sample. In addition, none of the subindex weights used in the simulation perfectly match the actual subindex weights that compose the broad index. For a given synthetic fund, the subindex weights are held fixed for the entire time period. However, as calculated by Lehman Brothers, the actual subindex weights that compose the broader index vary as the available number of bond issues within each subindex changes over time. As a result, the synthetic fund return series are not perfectly correlated with the broad indices from the six index model, and in the case of the fixed weight synthetic funds, we avoid having the simulated returns produce a perfect fit with the model.

Results from the portfolio allocation analysis are presented in Table 4 where we report findings for both our actual fund sample and the synthetic fund samples. Overall, the results suggest perverse timing ability by our sample of funds. The average cash adjustment across the sample is −6.21% indicating that funds have a higher allocation to short Treasury securities when this

[10] We note that our tests are strictly based on estimated portfolio weights and are an unconditional measure of timing skill. We make no attempt to distinguish performance based on the use of public or private information as first advocated by Ferson and Schadt (1996).

[11] An alternative approach would pool the observations and estimate the regression using dummy variables. In general, this would result in more efficient estimates. However, in our situation, we are estimating allocations and are constraining the coefficients. A large number of constraints would be necessary to ensure that the set of estimated coefficients on both the indices and dummy variables reflect investment strategies that could be implemented by managers. Given the limited number of observations, we choose to estimate allocations over the separate regimes to maintain as many degrees of freedom as possible.

[12] The mortgage index is composed of three subindices. Thus, for this index, we generate subindex weights w_i, w_j and w_k that satisfy $0<w_i$, w_j, $w_k<1$ and $w_i+w_j+w_k=1$.

Table 4
Portfolio allocation analysis: timing between cash and bonds

Panel A — results for entire sample

	Actual funds	Synthetic fund fixed weight	Synthetic fund buy and hold	Synthetic fund buy and hold rebalance
Mean cash adjustment	−6.21%	1.17%	3.10%	−1.19%
Median cash adjustment	−5.18%	1.02%	3.23%	−1.08%
# positive	24	50	61	47
# negative	50	34	23	36
# no adjustment	10	0	0	1
# positive and significant	5	8	14	12
# negative and significant	21	1	1	7

Panel B — results based on morningstar classifications

	Short maturity	Intermediate maturity	Long maturity
Number of funds	36	40	8
Mean cash adjustment	−1.55%	−9.80%	−9.25%
Median cash adjustment	−0.80%	−5.81%	−9.90%
# positive	16	6	2
# negative	19	25	6
# no adjustment	1	9	0
# positive and significant	4	0	1
# negative and significant	9	10	2

The table reports results for the portfolio allocation analysis performed to examine the timing ability of our fund sample. Panel A reports results for the entire fund sample while Panel B reports results where the sample is sorted by the fund's Morningstar maturity classification. We define the short Treasury index as our timing portfolio and we measure ability relative to this portfolio. Our first dataset contains all observations where the total return on the short Treasury index is greater than the return on the three remaining bond maturity indices. The second set contains all observations where the short Treasury index is the worst performing portfolio. For each partition, we employ Sharpe's quadratic program as defined in Eq. (2) to estimate the each fund's exposure to the short Treasury index. We calculate the difference between the average short Treasury allocation when short Treasury securities are the best performing portfolio and the average allocation when short Treasuries are the worst performing portfolio. We define this difference as the cash adjustment. The mean cash adjustment reported in the table is the average across all funds. Statistical significance is based on the 5% level of significance. For comparison, we also calculate the cash adjustment for synthetic funds that are designed to match each fund's characteristics but instead follow a passive investment strategy. To create the synthetic funds, we assume each synthetic fund follows an investment strategy where their initial exposure to each broad index is represented by a random set of investment weights on each subindex where the subindex weights w_i and w_j satisfy $0<w_i, w_j<1$ and $w_i+w_j=1$. Then using each actual fund's average exposure to the broad indices, the subindex weights, and the time series of subindex returns, we generate three synthetic funds for each individual fund. The first synthetic fund is assumed to follow a fixed weight strategy over the entire time period where monthly investment weights equal the average style weights of the matching fund. The returns for the second synthetic fund are assumed to evolve according to a buy and hold strategy. The returns for the third fund are generated following a buy and hold strategy with annual rebalancing back to the initial investment weights. We use the same combination of subindex weights across funds, but we repeat the process 100 times by generating 100 combinations of subindex investment weights. The results for each synthetic fund sample reflect the average values across the 100 simulations.

portfolio underperforms all of the bond maturity indices. The average is not heavily influenced by outliers as the median adjustment is −5.18%. A t test of the mean indicates this negative value is statistically significant at the five percent level. Individually, 60% of the funds have a negative cash adjustment, and for 26% of the funds, the cash adjustment is negative and significant. Only 29% of the funds have a positive cash adjustment with the adjustment significant for only 6% of the sample.

Our results indicate that ten of the funds in our sample do not have a nonzero exposure to short Treasury securities across either of the data partitions and thus do not appear to engage in timing activity between cash and bonds. When these funds are excluded from the analysis, the overall performance of the funds that do time the cash and bond markets appears much worse. The average (median) cash adjustment is −7.1% (−6.9%). Over 45% of the funds (34 of 74) have a negative cash adjustment greater than −10%.

When we compare these results to the results from our three synthetic fund samples, the evidence indicates that the negative cash adjustment for our actual fund sample is reflective of perverse skill rather than induced by market conditions. The results reported in the table for each synthetic fund sample reflect the average values across the 100 simulations. In all cases, the mean cash adjustment of the actual funds is worse than the cash adjustment generated by the synthetic funds. Both the fixed weight and buy and hold synthetic funds have positive cash adjustments (1.17% and 3.10% respectively) while the annual rebalancing strategy results in a negative cash adjustment of −1.19%. However, none of the 100 simulations across each of the synthetic fund samples generates a cash adjustment worse than −2.41%. In addition, none of the simulations are able to match the number of funds with negative cash adjustment values or the number of negative and significant cash adjustments.

As a final test, we sort the funds by the Morningstar maturity classifications used in Section 4 and calculate the average cash adjustment for each group.[13] Results are reported in Panel B of Table 4. Evidence indicates that perverse timing is consistent across all maturity classifications. The mean and median cash adjustment is negative for each classification. The worst performance is by the intermediate funds as measured by a mean adjustment of −9.80%. If the nine funds with zero adjustment are excluded, the

[13] As discussed in Section 4, we were unable to determine the average maturity of one of our funds from the data provided by Morningstar. For the purpose of this test and all that follow, we classify this fund as a short maturity fund based on the substantial investment in the short maturity portfolio according to the quadratic regression.

Table 5
Portfolio allocation analysis: timing the bond maturity spectrum

	Short maturity funds	Int maturity funds	Long maturity funds
Average allocation to short bonds bonds−normal yield curve	50.4%	26.5%	0.0%
Average allocation to short bonds−inverted yield curve	42.7%	39.9%	0.0%
Short bond adjustment	7.7%	−13.4%	0.0%
Average allocation to long bonds−normal yield curve	0.0%	11.3%	24.5%
Average allocation to long bonds−inverted yield curve	0.0%	14.8%	29.9%
Long bond adjustment	0.0%	−3.5%	−5.4%

The table reports results for the portfolio allocation analysis performed to examine the government/corporate yield curve timing ability of our fund sample. We group funds by their Morningstar maturity classification (short, intermediate, and long). We then partition our dataset into two groups: 1) all months when the government/corporate yield curve is normal and 2) all months when the yield curve is inverted. For each fund group and each partition, we employ Sharpe's quadratic program as defined in Eq. (2) to estimate each fund's exposure to the short and long bond indices. We define the difference between the short bond allocation during a normal yield curve and the short bond adjustment during an inverted yield curve as the short bond adjustment. We define the long bond adjustment in a similar manner. The mean cash adjustment for each group is reported in the table.

average adjustment for the intermediate funds drops to −12.65%. Long funds have similar poor performance as indicated by a negative adjustment of −9.25%. The perverse performance of the short maturity funds is not as strong (−1.55%), but more than twice as many short funds have significant negative cash adjustments than significant positive adjustments.[14]

Overall, our results indicate that the negative cash adjustment of the funds is consistent across all maturity classification of funds and is the result of perverse timing ability rather than a function of changes in bond market conditions.

5.2. Portfolio allocation analysis revisited: timing across bond maturities

As mentioned in the introduction, a main advantage of our empirical technique is that we can take a more detailed look at timing behavior than can be inferred from traditional approaches. In addition to choosing their broad asset allocation between cash and bonds, managers can engage in a separate timing decision where they determine how to allocate assets across bond maturities. Changes in cash allocations may reflect a combination of timing strategies and cash maintenance decisions while changes in the percentages allocated to the three bond maturity indices are likely immune from such liquidity motivated trading. We want to examine whether we detect similar perverse timing behavior across our three bond maturity indices.

As illustrated in the introduction, an examination of returns from the three maturity indices indicates significant variation in the government/credit yield curve over our time period. Specifically, a normal upward sloping yield curve was in effect for 57 months of our sample period and an inverted curve occurred for 46 months. The yield curve was inverted for most of 1994, 1996, 1999, and for half of 2001 providing ample timing opportunities for a skilled manager.

Given this variation, we revisit our portfolio allocation methodology and examine whether the funds are successful at shifting allocations to bonds of different maturities across changes in the government/credit yield curve. First, we group funds by their Morningstar classification (short, intermediate, and long) and average the returns of the funds within each group. Then, we partition the data into two sets: months with a normal yield curve and months with an inverted yield curve. We use our index model defined in Eq. (3), and we run Sharpe's quadratic program for each fund group across the two partitions.

For each group of funds, we calculate the short bond adjustment and long bond adjustment as follows:

$$\text{short bond adjustment} = b_{i,\text{gc0105}}(\text{normal yield curve}) - b_{i,gc0105}(\text{inverted yield curve}) \tag{7}$$

$$\text{long bond adjustment} = b_{i,\text{gclong}}(\text{normal yield curve}) - b_{i,gclong}(\text{inverted yield curve}) \tag{8}$$

As defined, a negative short bond adjustment and a positive long bond adjustment reflect successful timing ability.

Results are presented in Table 5 and provide evidence of poor timing as measured by the bond adjustment variables. The short maturity funds have a higher allocation to short bonds during normal yield curves as measured by a positive short bond adjustment of 7.7%. Both the intermediate and long maturity funds have higher allocations to long bonds during months of inverted yield curves as indicated by long bond adjustment measures of −3.5% and −5.4% respectively. Only the short bond adjustment for the intermediate funds has the correct sign (−13.4%). But evidence indicates that the funds are using cash assets to increase the short bond allocation during months with inverted yield curves. The average cash allocation is 10.5% lower during inverted yield curve months which also are the months in which investment in cash is more valuable.

We repeat the analysis by calculating the percentage short and long bond adjustments as the changes in the percentage investment in those indices relative to total investment in the three maturity indices. The results remain the same for each of the

[14] The cash adjustment results for the short maturity classification are similar to the synthetic fund results for the annual rebalancing strategy. To test whether the returns for the short maturity classification are mimicking an annual rebalancing strategy, we compare the average returns for the two groups. The t value of a matched pairs *t*-test is −4.92 and we can reject the null hypothesis of no difference in the average return series. Thus, the cash adjustment values for the short maturity classification do not appear to be the results of a passive annual buy and hold strategy with annual rebalancing.

Table 6
Attribution returns

Classification	# of funds	Average attribution return	Standard deviation
Positive and significant	5	0.10%	0.85%
Positive and nonsignificant	19	−0.27%	0.44%
No change	10	−0.61%	0.45%
Negative and nonsignificant	28	−0.71%	0.55%
Negative and significant	22	−0.85%	0.61%

The table presents the distribution of average annual fund attribution returns based on fund timing ability classification. A fund's attribution return compares the fund's annual return during year t to the return that would be achieved over the same year if the estimated weights from a previous period are in effect. We use the previous 24 months of data to estimate the prior period weights. Thus, the first annual attribution return is calculated for 1996 based on estimated weights from 1994–1995. We calculate an annual attribution return for each fund for each year t that a fund has 24 monthly observations available from $t-2$ to t. Then, we average the annual attribution returns for each fund. Next, we group the funds into five categories based on their cash adjustment measure from the portfolio allocation test: 1) positive and significant, 2) positive and not significant, 3) no adjustment, 4) negative and not significant, and 5) negative and significant. The average attribution return reported in the table reflects the average across the funds in each classification.

maturity classifications. For the overall sample, the percentage short bond adjustment is 10.5% and the percentage long bond adjustment is −8.9% which confirms the inability to successfully time the government/credit yield curve.

We would also like to note that the funds do not appear to change allocations to the mortgage and high yield sectors as a means of generating superior returns. The portfolio allocation analysis results in this section and from the previous section indicate relatively constant allocations to each of these indices. As a specific example, the changes in allocation to the mortgage index across the two yield curves ranges from 0.9% to 3.2% for each of the three maturity classifications. The range for the high yield index is even smaller (−0.1% to 2.1%). Thus, the perverse timing between cash and bonds and across bond maturities does not appear to be driven by changes in allocations across the mortgage and high yield sectors.

5.3. Robustness — attribution returns

We do not have access to the actual portfolios of our fund sample, and as a result, we cannot independently validate the portfolio estimates reported in the previous sections. Thus, we test the robustness of our results by defining an annual attribution return as follows:

$$r_{\mathrm{att},y} = r_{p,y} - \sum_{i=1}^{k} b_{pi,y-1} r_{i,y} \tag{9}$$

A fund's attribution return compares the fund's annual return during year y to the return that would be achieved over the same year if the estimated weights from a previous period are in effect.[15] Attribution returns are designed to measure the value of active fund management. Value is reflected through 1) strategic changes to the portfolio allocation made by the manager (market timing ability) or 2) actual selections that deviate from the benchmark indices included in the model (asset selection ability). Given our focus on high quality corporate bond funds, we hypothesize that it would be difficult for a manager to consistently have significant positive or negative asset selection ability since returns to investment grade bonds of a similar maturity tend to be relatively homogeneous. Thus, the variance in attribution returns across funds is more likely to reflect differences in timing skill. If our previous methodology accurately measures timing ability, there should be a relationship between a fund's cash adjustment measure and its average annual attribution return. Funds that have positive ability should have greater attribution returns than those funds that display perverse skill.

Unlike our previous tests, portfolio weights used in the calculation of the attribution returns are estimated over a continuous time period rather than across partitions reflecting various bond market conditions. To calculate a fund's annual attribution return, we use 24 months of data to estimate the prior period weights. Thus, the first annual attribution return is calculated for 1996 based on estimated weights from 1994–1995.[16] By using the two year estimation window, we have at least one annual attribution return for each fund given the minimum three year survival rule imposed for inclusion in our sample. For each fund, we calculate an annual attribution return for each year y that a fund has 24 monthly observations available from year $y-2$ to y. If a fund did not survive for the entire year y, we calculate a partial year attribution return.

We average the annual attribution returns for each fund. In order to compare the results to our portfolio allocation analysis, we place each fund into one of five categories based on the following classifications of the fund's cash adjustment measure: 1) positive and significant, 2) positive and not significant, 3) no adjustment, 4) negative and not significant, and 5) negative and significant. Then, we calculate the average cash adjustment across funds within each group.

Results are presented in Table 6. There is a strong relationship between the attribution returns and the results from portfolio allocation analysis. The average annual attribution return across all funds is −0.57% indicating negative performance consistent

[15] Sharpe (1992), Ibbotson (1996), and Dor et al. (2003) use this methodology. Sharpe and Dor, Jagannathan, and Meier refer to the calculation as a selection return while Ibbotson uses the term attribution return.

[16] We note that the design of our test excludes timing performance during 1994 and 1995 and implicitly assumes that the overall estimated timing skill of a fund is not dominated by performance during this two year period.

Table 7
Timing ability and fund expense ratios

	Expense ratio	Cash adjustment
Quartile 1	0.41%	1.1%
Quartile 2	0.68%	−6.1%
Quartile 3	0.82%	−3.4%
Quartile 4	1.12%	−16.5%

The table presents a comparison of the average fund cash adjustment for a partition of the fund sample based on fund expense ratios. Funds are sorted by their average annual expense ratio and divided into quartiles with the first quartile representing funds with the lowest expense ratio. For each quartile, we calculate the average fund cash adjustment across all funds within the group. Annual expense ratios are based on data supplied by CRSP.

with the lack of timing ability we previously found for the overall sample. When we examine our fund groups, we find that the attribution return measure decreases monotonically as estimated timing skill declines. According to our cash adjustment measure, five funds were found to have significant positive timing ability. This is the only category to have a positive mean attribution return with a value of 0.10%. Then, the average attribution return falls to −0.27% for funds with positive but insignificant cash adjustment measures. The funds with negative cash adjustment measures have the worst attribution returns. Funds with negative but insignificant timing ability have average attribution returns of −0.71%, while those that have significant negative timing ability have the lowest attribution return of −0.85%.

Thus, the attribution returns are consistent with the results from the portfolio allocation analysis. The combination of the timing results from our synthetic funds and the strong correlation between the attribution returns and average cash adjustment measures suggest that our results reflect perverse timing ability by the funds rather than spurious results based on model misspecification or problems with the design of the empirical tests.

5.4. Timing ability and fund characteristics

Given the evidence that the overall fund sample has perverse timing ability, we are interested in whether timing ability is correlated with observable fund characteristics. Our goal is to determine if the negative performance is widespread or driven by a specific subset of the fund sample.

There are four characteristics we can examine. Using the fund data described in Section 2, we can sort the funds into quartiles based on their expense ratios, portfolio turnover ratios, and net assets. For these sorts, the first quartile represents funds with the lowest values for the specific characteristic. As a final grouping, we can also sort the funds based on whether the fund survives for the entire time period. For each partition of the funds, we calculate the average cash adjustment for funds that are in a specific grouping.

Table 7 presents results comparing timing performance and expense ratios. The results indicate that positive timing skill is concentrated in the first quartile which contains funds with the lowest average expense ratio. This group has an average expense ratio of 0.41% and a positive cash adjustment of 1.1%. This value is not influenced by outliers as the median adjustment is 0.5%. Within this quartile, only two of the 21 funds have significant negative cash adjustments.

Across the remaining three quartiles, negative performance is widespread and is most negative in the highest expense ratio quartile. Funds in quartile four have the highest average expense ratio of 1.12% and have the most negative cash adjustment of −16.5%. Again, this value is not influenced by outliers as the median is −17.5%. Across the quartiles, this average is the only value statistically different from zero, and it is substantially worse than the values for the other three quartiles. Within this quartile, 16 of the 21 funds have a negative cash adjustment value including seven funds where the adjustment is negative and significant. The cash adjustment difference of 17.6% between quartiles one and four is statistically significant at the five percent level.

We do not find any statistically significant relationships between the cash adjustment variable and the other fund characteristics.[17] Thus, our evidence indicates that the negative timing ability that we have found is relatively widespread across the sample, with the only exception the subset of funds with the lowest expense ratios. We do not have a theory to explain the correlation between the expense ratio and timing ability, although our result is identical to that reported by Comer (2006) for hybrid funds. He finds that better timing performance is directly related to lower expense ratios but does not find any statistically significant relationship between other observable characteristics and stock timing performance.

Studies in the mutual fund tournament literature (e.g. Brown et al., 1996; Koski and Pontiff, 1999) suggest that allocation changes and timing decisions may be related to managerial gaming behavior. Alternatively, Morey and O'Neal (2006) provide evidence that bond fund managers engage in window dressing behavior surrounding disclosure dates. Thus, the relationship we find between expenses and timing skill may be reflective of more complicated behavior by bond fund managers. We leave this issue to future research.

[17] We do find some weak evidence that nonsurvivors perform worse than survivors which is consistent with other results on survivorship in the mutual fund literature. The 56 funds that survive have an average cash adjustment of −5.3% while the 28 nonsurvivors have an average adjustment of −7.9%. However, the difference is not statistically significant.

5.5. Timing and cash flows

Thus far, our results suggest perverse timing ability. As our final empirical test, we want to examine an alternative explanation of the results. Edelen (1999) indicates that the negative market timing performance of equity funds can be explained by liquidity motivated trading related to investor flows. It is possible that when bonds do well, money flows into bond mutual funds. The manager may have trouble investing these flows, and thus it appears that the funds have more cash regardless of the manager's actual investment skill.

To test this hypothesis, we examine the cash flows to the bond funds during the months when the short Treasury index outperforms the three bond maturity indices (S_1 from Eq. (4)) and the months when it underperforms (S_2 from Eq. (4)). For each fund in existence during each month included in S_1 or S_2, we calculate the percentage change in cash flow to the fund as follows:

$$\mathrm{pcf}_{i,t} = \frac{\mathrm{tna}_{i,t} - tna_{i,t-1}(1 + r_{i,t})}{\mathrm{tna}_{i,t-1}} \tag{10}$$

where pcf represents percentage cash flow to the fund, tna represents total net assets of the fund, and r represents the monthly holding period return of the fund. As done by Zhao (2005), we use percentage rather absolute cash flow to control for differences in size across funds since all funds do not survive for all months.

Then, within each data partition, we calculate an equal weighted average of all fund-month observations. Our results indicate that on average, cash flows to bond funds are negative for both data partitions. When the short Treasury index is the best performing index, average percentage change in cash flow to the funds is −0.25% while the percentage is −0.02% when the short Treasury index is the worst. Given that cash outflows are occurring during the months we use to measure timing ability it appears that the return chasing behavior of bond investors is not driving the perverse timing behavior that we have detected.

6. A Suggested answer to a puzzle

Thus far, we have provided three main results. On average, our fund sample underperforms the benchmark index by 1.32% as seen in Table 1 and the entire amount of underperformance cannot be attributed to expenses. Second, the funds demonstrate perverse timing ability between cash and bonds and across bonds of various maturities. Third, the negative performance is relatively widespread and is not driven by a specific subset of the fund sample. Yet, despite these results, 67% of our sample survives for the entire time period. This raises the following question: are investors behaving irrationally by investing in these funds?

Although we are unable to definitively answer the question, we propose a potential explanation based on available data. We believe that investors value these funds not just for the returns generated by the managers but also for the diversification benefits the funds provide within the investors' overall portfolio. Our hypothesis is supported by the following data.

Table 8
Relationship between fund cash flows, average fund returns, and stock market returns

Panel A — time series of fund cash flows, fund returns, and stock market returns

Year	Annual cash flow	Average fund return	Return of S&P 500	Excess average fund return relative to S&P 500
1994	−2921.6	−1.64%	−0.75%	−0.89%
1995	596.8	14.28%	35.67%	−21.39%
1996	−179.7	3.97%	21.16%	−17.18%
1997	868.1	7.85%	30.33%	−22.47%
1998	3631.2	6.86%	22.28%	−15.41%
1999	−419.8	0.74%	25.26%	−24.51%
2000	−3464.3	9.21%	−11.04%	20.26%
2001	2623.9	7.39%	−11.27%	18.66%
2002	4523.7	6.76%	−20.84%	27.61%
2003	1238.9	3.64%	33.14%	−29.50%

Panel B — explanatory power in single variable regression model

Explanatory variable	R^2
Fund return	0.07
S&P 500 return	0.00
Excess average return relative to S&P 500	0.01
Lagged fund return	0.05
Lagged S&P 500 return	0.16
Lagged excess average return relative to S&P 500	0.20

The table presents data on the total cash flows to the 56 surviving funds within our sample and return measures that potentially impact these flows. Panel A presents a time series of each of the variables. Annual cash flow is measured in millions of dollars and is defined as the change in total net asset value minus the appreciation in fund assets. It is calculated on an annual basis for the 56 funds that survive the entire period. Panel B presents the explanatory power of the contemporaneous and lagged values of each of our return measures in a single variable regression where annual cash flow is the dependent variable.

First, the Investment Company Institute publishes a Profile of Mutual Fund Shareholders. The Fall 2004 profile provides the most detailed characteristics of bond fund investors. Nearly 90% of bond fund holders also hold equity funds. Responses to the survey questions indicate that bond fund holders tend to be more risk averse, wealthier, and more sophisticated than equity fund holders. Although the survey does not ask the specific reason that the investor chose a bond fund, 97% of bond fund investors indicate investment diversification as a very or somewhat important reason for owning mutual funds. The combination of information provided by the survey questions suggests that investors view their bond funds as part of a diversified portfolio.

Second, the average return of our sample of funds has a very low correlation with stock returns over the ten year period of our study. We use the annual broad CRSP value weighted market index to represent the average stock return. Its correlation with the average annual return of our fund sample is 0.08. Similar low or negative correlations occur when we compare the average return for each of our maturity groupings to the annual returns from both the CRSP index and the S&P 500 index. Educated investors are likely to be aware of the potential for significant diversification benefits from holding bond funds which have a low correlation with their stock holdings.

Third, over the full sample period of 1994–2003 data from the Investment Company Institute reveals the growth in total net assets for equity and bond funds to be 351% and 116%, respectively. Moreover, a more telling statistic is the growth in total net assets over the 2001–2003 sub-period which marks the timeframe where equities were bearish. Accordingly, if investors do indeed see bond funds as a diversification tool then we would expect the growth in total net assets to be significantly greater than the growth in total net assets for equity funds. Indeed, the growth in total net assets over the 2001–2003 period for equities and bond funds is −16% and 35%, respectively.

Fourth, unlike the case with equity funds, we find that neither current nor past performance of our fund sample explains a substantial portion of the cash flows to the funds. For the 56 funds that survive over the entire time period, we measure annual fund cash flow to each individual fund. We sum the individual cash flows to get a total annual cash flow to the funds. We then compare annual cash flows to 1) the average fund return, 2) the return of the S&P 500 index, and 3) the excess average fund return relative to the S&P 500. We also compare cash flows to the lagged values of each of the three variables. Panel A of Table 8 reports the time series of each of these variables, while Panel B reports the explanatory power of each of the variables in a single variable linear regression model. We find low correlations between each of the variables and our cash flow measure indicating that other factors that are not performance related impact investor cash flows to these funds. Contemporaneous values of the return measures have virtually no explanatory power, while the lagged values of the variables have stronger explanatory power but only explain 5% to 20% of the variation in cash flows.

Taken together, these data suggest that investors value the diversification benefits of holding these funds within their portfolio despite the overall negative performance of the funds. Bond index funds would seem to be an alternative and more effective solution to achieving such diversification. However, few were in existence during our sample period. According to Morningstar, only nine bond index funds were available at the beginning of our sample period. By 1999, that number had only grown to 14. Even by the end of 2003, there were only 19 index funds available across only 14 fund families. Not all investors, particularly those who invest in bond funds through their institutions' pension or retirement accounts, have access to the small number of available bond index funds and must invest in underperforming actively managed funds to achieve the desired diversification benefits.

7. Conclusion

This study is among the first to examine the timing ability of a sample of bond mutual funds. We choose to examine bond timing performance for two reasons. First, a small but growing literature (Chance and Hemler, 2001; Bollen and Busse, 2001, 2004; Glassman and Riddick, 2006; Chen, 2005; Comer, 2006) suggests that groups of money managers have demonstrated more positive market timing ability than previously believed. Second, an examination of bond market returns over the last decade indicates that positive timing skill by bond funds, specifically the ability to time the investment grade bond maturity spectrum, could generate significantly higher returns than those of a comparable bond index fund.

We focus on the timing skill of high quality bond funds over the 1994–2003 period. The unique feature of our study is that we rely on the quadratic programming technique pioneered by Sharpe (1992) nested within the framework of the Henriksson and Merton (1981) models. We measure skill by comparing estimated fund manager allocations to bonds of different maturities across various bond market conditions. This allows us a more detailed look at timing behavior than can be inferred from the traditional approaches. We obtain estimated allocations using the quadratic programming technique of Sharpe (1992) and run empirical tests based on the portfolio allocation analysis methodology of Comer (2006).

Our results indicate that as a group the sample of funds demonstrates statistically significant perverse timing ability as measured by higher allocations to cash when bonds of all maturities outperform Treasury bills. Within the fixed income class, results indicate additional perverse timing as measured by average allocations to bonds of various maturities. Additional tests indicate that the perverse timing ability we detect is not spurious and is relatively widespread across the sample.

The combination of our perverse timing results and the negative alphas reported by other researchers (Blake et al., 1993; Ferson et al., 2006) raises the question of the value provided by bond funds, particularly given that 67% of our sample survives for the entire 10 year period despite having negative timing ability and underperforming the appropriate benchmark. Although we are unable to provide a definitive answer to this issue, we present evidence that suggests investors value these funds for the diversification benefits they provide within a stock/bond portfolio.

Overall, despite the results from recent studies that indicate superior forecasting skills across various groups of money managers, our evidence indicates that our subset of bond fund managers do not share this skill and that individual investors would be better off with bond index funds.

Appendix A. Lehman bond indices

Our index model in Section 4 uses six Lehman Brothers bond indices. Each of these indices is composed of two subindices while three subindices make up the mortgage index. The subindices are used to generate the returns of the matching synthetic funds as discussed in Section 5.1. Below is a list of each broad index followed by the subindices of which it is composed:

1) Short treasury bill index: 1) Treasury bill index and 2) Treasury coupon index
2) 1–5 year government/credit index: 1) 1–5 year government index and 2) 1–5 year credit index
3) 5–10 year government/credit index: 1) 5–10 year government index and 2) 5–10 year credit index
4) Long (10+ year) government/credit index: 1) long (10+ year) government index and 2) long (10+ year) credit index
5) High yield index: 1) Intermediate (<10 year) high yield index and 2) Long (10+ year) high yield index
6) Mortgage backed security index: 1) GNMA index, 2) FNMA index, and 3) FHLMC index

References

Agarwal, Vikas, Naik, Narayan, 2000. Generalised style analysis of hedge funds. Journal of Asset Management 1, 93–109.
Blake, Christopher, Elton, Edwin, Gruber, Martin, 1993. The performance of bond mutual funds. Journal of Business 66, 371–403.
Bollen, Nicolas, Busse, Jeffrey, 2001. On the timing ability of mutual fund managers. Journal of Finance 56, 1075–1094.
Bollen, Nicolas, Busse, Jeffrey, 2004. Short term persistence in mutual fund performance. Review of Financial Studies 18, 569–597.
Brown, Keith, Harlow, W., Starks, Laura, 1996. Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry. Journal of Finance 51, 85–110.
Brown, Stephen, Goetzmann, William, Ibbotson, Roger, 1999. Offshore hedge funds: survival and performance, 1989–1995. Journal of Business 72, 91–117.
Brown, Stephen, Goetzmann, William, Park, James, 2000. Hedge funds and the Asian currency crisis. Journal of Portfolio Management 95–101.
Carhart, Mark, 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.
Carhart, Mark, Carpenter, Jennifer, Lynch, Anthony, Musto, David, 2002. Mutual fund survivorship. Review of Financial Studies 15, 1439–1463.
Chance, Don, Hemler, Michael, 2001. The performance of professional market timers: daily evidence from executed strategies. Journal of Financial Economics 62, 377–411.
Chen, Yong, 2005. Timing ability in the focus market of hedge funds. working paper.
Chen, Yong, Ferson, Wayne, Peters, Helen, 2005. The timing ability of fixed income mutual funds. working paper.
Comer, George, 2005. Evaluating bond fund sector timing skill working paper.
Comer, George, 2006. Hybrid mutual funds and market timing performance. Journal of Business 79, 771–797.
Comer, George, Larrymore, Norris, Rodriguez, Javier, in press. Controlling for Fixed Income Exposure in Portfolio Evaluation: Evidence from Hybrid Mutual Funds, Review of Financial Studies.
Dor, Arik Ben, Jagannathan, Ravi, Meier, Iwan, 2003. Understanding mutual fund and hedge fund styles using return based style analysis. Journal of Investment Management 1, 97–137.
Edelen, Roger, 1999. Investor flows and the assessed performance of open-ended mutual funds. Journal of Financial Economics 53, 439–466.
Elton, Edwin, Gruber, Martin, Blake, Christopher, 1995. Fundamental economic variables, expected returns, and bond fund performance. Journal of Finance 50, 1229–1256.
Ferson, Wayne, Schadt, Rudi, 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51, 425–461.
Ferson, Wayne, Kisgen, Darren, Henry, Tyler, 2006. Evaluating Fixed Income Performance with Stochastic Discount Factors. Review of Financial Studies 19, 423–455.
Fung, William, Hsieh, David, 1997. Empirical characteristics of dynamic trading strategies: the case of hedge funds. Review of Financial Studies 10, 275–302.
Glassman, Debra, Riddick, Leigh, 2006. Market Timing by Global Fund Managers. Journal of International Money and Finance 25, 1029–1050.
Goetzmann, William, Ingersoll, Jonathan, Ivkovic, Zoran, 2000. Monthly measurement of daily timers. Journal of Financial and Quantitative Analysis 35, 257–290.
Grinblatt, Mark, Titman, Sheridan, 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 393–415.
Henriksson, Roy, Merton, Robert, 1981. On market timing and investment performance. Journal of Business 54, 513–533.
Ibbotson, Roger, 1996. Do winning mutual funds repeat? TMA Journal 51–56.
Koski, Jennfier, Pontiff, Jeffrey, 1999. How are derivatives used? Evidence from the mutual fund industry. Journal of Finance 54, 791–816.
Lobosco, Angelo, DiBartolomeo, Dan, 1997. Approximating the confidence intervals for sharpe style weights. Financial Analysts Journal 53, 80–85.
Morey, Matthew, O'Neal, Edward, 2006. Window Dressing in Bond Mutual Funds. Journal of Financial Research 29, 325–340.
Sharpe, William, 1992. Asset allocation: management style and performance measurement. Journal of Portfolio Management 7–19.
Treynor, Jack, Mazuy, Kay, 1966. Can mutual funds outguess the market? Harvard Business Review 131–136.
Zhao, Xinge, 2005. Determinants of flows into retail bond funds. Financial Analysts Journal 61, 47–59.