Measuring the Timing Ability of Fixed Income Mutual Funds

by
Yong Chen
Wayne Ferson
Helen Peters

first draft: November 25, 2003
this revision: October 2, 2008

* Ferson is the Ivadelle and Theodore Johnson Chair of Banking and Finance and a Research Associate of the National Bureau of Economic Research, Marshall School of Business, University of Southern California, 3670 Trousdale Parkway Suite 308, Los Angeles, CA. 90089-0804, ph. (213) 740-5615, ferson@marshall.usc.edu. Peters is Professor of Finance at Boston College, 140 Commonwealth Ave, Chestnut Hill, MA. 02467, (617) 552-3945, helen.peters.1@bc.edu. Chen is Assistant Professor of Finance at the Department of Finance, Insurance and Business Law, Virginia Polytechnic Institute, Blacksburgh, VA. 24061, yong.chen@vt.edu. We are grateful to participants in workshops at Baylor University, Brigham Young, Boston College, DePaul, the University of Connecticut, the University of Florida, the University of Georgia, Georgetown, the University of Iowa, the University of Massachusetts at Amherst, Michigan State University, the University of Michigan, Southern Methodist University, the University of Oklahoma, the University of Kansas, Virginia Tech., York University and the Office of Economic Analysis, Securities and Exchange Commission for feedback, as well as to participants at the following conferences: The 2005 Global Finance Conference at Trinity College, Dublin, the 2006 Journal of Investment Management Conference, the 2008 ISCTE/NOVA Business School conference, the 2005 Southern Finance and the 2005 Western Finance Association Meetings. We also appreciate helpful comments from Vasantha Chigurupati, Walter Dolde, James Hilliard, Scott Linn, Ricardo Rodriguez, Bryan Schmutz, M. Abdullah Sahin, Nadia Voztioblennaia and Hong Yan.

ABSTRACT

Measuring the Timing Ability of Fixed Income Mutual Funds

This paper evaluates the ability of bond funds to "market time" factors related to bond markets. Timing ability generates nonlinearity in fund returns, but there are several non-timing-related sources of nonlinearity. We find that controlling for non-timing-related nonlinearity is important. Funds' returns are more concave than benchmark returns, relative to nine common factors, and this would appear as poor timing ability in naive models. With the controls, the overall distribution of the timing coefficients appears neutral to weakly positive. The timing-adjusted performance of many bond funds appears significantly negative on an after-cost basis, but many funds have positive performance on a before-cost basis.

1. Introduction

The amount of academic research on bond fund performance is small in comparison to the economic importance of bond funds. Recently the total net assets of U.S. bond funds has been about 1/6 the amount in equity-style mutual funds and similar to the value of hedge funds. Large amounts of additional fixed-income assets are held in professionally managed portfolios outside of mutual funds, for example in pension funds, trusts and insurance company accounts. The turnover of a typical bond mutual fund far exceeds that of a typical equity fund, suggesting that active portfolio management is important in bond funds. Thus, it is important to understand the performance of bond fund managers.

Elton, Gruber and Blake (EGB, 1993, 1995) and Ferson, Henry and Kisgen (2006) study US bond mutual fund performance, concentrating on the funds' risk-adjusted returns, or alphas. They find that the typical risk-adjusted, average performance is slightly negative and largely driven by funds' expenses. This might suggest that investors would be better off selecting low-cost passive funds, and EGB draw this conclusion. However, performance may be decomposed into components, such as timing and selectivity ability. If investors place value on timing ability, for example a fund that can mitigate losses in down markets, they would be willing to pay for this insurance with lower average returns. This is one of the first papers to comprehensively study the ability of US bond funds to time their markets.[1]

Timing ability on the part of a fund manager is the ability to use superior information about the future realizations of common factors that affect bond market returns.[2] Selectivity

[1] Brown and Marshall (2001) develop an active style model and an attribution model for fixed income funds, isolating managers' bets on interest rates and spreads. Comer, Boney and Kelly (2005) study timing ability in a sample of 84 high quality corporate bond funds, 1994-2003, using variations on Sharpe's (1992) style model. Aragon (2005) studies the timing ability of balanced funds for bond and stock indexes.

[2] We do not explicitly study "market timing" in the sense recently taken to mean trading by investors in a fund to exploit stale prices reflected in the fund's net asset values. But we will see that these issues can affect measures of a fund manager's ability.

refers to the use of security-specific information. If common factors explain a relatively large part of the variance of bond returns, it follows that a relatively large fraction of the potential performance of bond funds might be attributed to timing. However, measuring the timing ability of bond funds is a subtle problem.

Traditional models of market timing ability rely on convexity in the relation between the fund's returns and the common factors.[3] In bond funds, perhaps even more clearly than in equity funds, convexity or concavity can arise for various reasons unrelated to timing ability. We find that other sources of nonlinearity are important, and our empirical analysis attempts to control for other sources of nonlinearity.

We find that funds' returns are typically more concave, in relation to a set of nine bond market factors, than are unmanaged benchmarks. Thus, without controls for non-timing-related nonlinearity, funds would appear to have poor (negative) market timing ability. When we introduce the controls the overall distribution of the timing coefficients appears neutral to weakly positive. After adjusting for timing ability the performance of many bond funds is significantly negative on an after-cost basis but many funds have significant ability on a before cost basis.

The rest of the paper is organized as follows. Section 2 describes the models and methods. Section 3 describes the data. Section 4 presents our empirical results and Section 5 offers some concluding remarks.

2. Models and Methods

A traditional view of performance separates timing ability from security selection ability, or

[3] The alternative approach is to directly examine managers' portfolio weights and trading decisions to see if they can predict returns and factors (e.g. Grinblatt and Titman, 1989). Comer (2006) and Moneta (2008) are early steps in this direction for bond funds. Of course, weight-based approaches cannot capture market timing that occurs between weight reporting intervals, which can be up to six months in length.

selectivity. Timing is closely related to asset allocation, where funds rebalance the portfolio among asset classes and cash. Selectivity essentially means picking good securities within the asset classes.

Like equity funds, bond funds engage in activities that may be viewed as selectivity or timing. Bond funds may attempt to predict issue-specific supply and demand or changes in credit risks associated with particular bond issues. Funds can also attempt to exploit liquidity differences across bonds, for example on-the-run versus off-the-run issues. These trading activities can be classified as security selection. In addition, managers may tune the interest rate sensitivity (e.g., duration) of the portfolio to time changes in interest rates in anticipation of the influence of economic developments. They may vary the allocation to asset classes differing in credit risk or liquidity, and which are likely to have different exposures to economic factors. Since these activities relate to anticipating market-wide factors, they may naturally be considered as market timing.

Classical models of market-timing use the convexity in the relation between the fund's return and the "market" return to indicate timing ability. In these models the manager observes a private signal about the future performance of the market and adjusts the market exposure of the portfolio. If the response is assumed to be a linear function of the signal as modelled by Admati, Bhattacharya, Ross and Pfleiderer (1986), the portfolio return is a convex quadratic function of the market return as in the regression model of Treynor and Mazuy (1966). If the manager shifts the portfolio weights discretely, as modelled by Merton and Henriksson (1981), the convexity may be modelled with call options. Our models modify the classical setup for bond market factors and to control for nonlinearities that are unrelated to managers' timing ability.

2.1 Nonlinearity Unrelated to Timing

There are many reasons apart from timing ability that a fund's return could have a nonlinear relation to a market factor. We group these reasons into four general categories. This section explains the intuition for each effect.

First, the underlying assets held by a fund may have a nonlinear relation to market factors. While nonlinearity may occur in equities, it is very likely in bonds. Even simple bond returns are nonlinearly related to interest rate changes. As we show in Table 3 below, unmanaged bond benchmark returns are often convex functions of common factors. Thus, to measure timing ability it is important to control for this nonlinearity.

A second potential cause of nonlinearity is "interim trading," studied by Goetzmann, Ingersoll and Ivkovic (2000), Ferson and Khang (2002) and Ferson, Henry and Kisgen (2006). This refers to a situation where fund managers trade more frequently than the fund's returns are measured. With mutual funds interim trading definitely occurs. A related effect is derivatives or other securities with option-like payoffs. For example, a fund that holds call options bears a convex relation to the underlying asset (Jagannathan and Korajczyk, 1986). Derivatives may often be replicated by high frequency trading, so if we can control for interim trading we also control for these derivatives.[4]

The potential impact of interim trading can be illustrated with an example. For the example we use daily data for 1989-2003 on the Lehman US Government Total Return Index, R_g, and the daily return to "cash," R_c, which we proxy with a 3-month Treasury bill.[5] A hypothetical fund trades each day, rebalancing in response to the flow of new money at the end of the previous

---

[4] Brown et al. (2004) explore arguments that incentives and behavioral biases can induce managers without superior information to engage in option-like trading within performance measurement periods.

[5] Specifically, to form the daily return we take the daily 3-month Treasury yield from the Federal Reserve H.15 release and assume that it corresponds to the purchase price of a newly issued, 91-day bill. On the next trading day the bill is assumed to be sold based on the new yield quote and its now shorter maturity. The price change, relative to the initial price, determines the daily return.

day. The new money flows in response to public expectations about the Government bond's return. This follows Ferson and Warther (1996), who find that new money flows into equity funds when instruments for public expectations of equity returns are high. The public information is modelled for this example by regressing the daily excess return, R_g-R_c, on the lagged change in the 3-month yield, Δy, where we get a slope coefficient b=0.28 (t-statistic = 2.78). The funds weight each morning is 10% in cash plus the percentage flow from the previous night, and the rest in the government index. The fund rebalances to 10% cash at the close of each day, prior to the learning of the new money flow. The 10% target is assumed to represent three standard deviations of the daily percentage cash flows, and the fitted values of the regression, scaled to match, represent the cash flows. The daily return of the fund is therefore $R_p = (1-x)R_g + xR_c$, where $x = 0.10 + F$ and F is the cash flow of the previous day. We compound the daily portfolio returns to monthly returns. Regressing the 177 monthly excess returns on the Government index excess return and its square, the coefficient on the square is 0.09 and its t-ratio is 2.34. The example shows that interim trading can result in a fund return that is nonlinear relative to the benchmark index. Thus, a fund with interim trading can appear to have market timing ability if we fail to control for the interim trading effect.

A third potential reason for nonlinearity unrelated to timing ability is stale pricing of a fund's assets. Thin or nonsynchronous trading has long been known to bias downward the estimates of beta for a portfolio (e.g., Fisher (1966), Scholes and Williams, 1977). If the degree of stale pricing is related to a market factor, we show below that such "systematic" stale pricing can create spurious concavity or convexity in the measured return.

Systematic stale pricing can be illustrated with another example using monthly returns on the Government Index, R_g, for 1989-2003. The measured return on a hypothetical fund is $R_{pt} = (1-\theta_t)R_{gt} + \theta_t R_{gt-1}$, where $\theta_t \in [0,1]$ captures the extent of stale pricing. We set $E(\theta_t)=0.15$, which roughly says that 15% of the measured prices are stale. The stale pricing is systematic, or correlated with the return of the index: $\theta_t = 0.15 + b_1[R_{gt}-E(R_g)]$. We set b_1 so that the minimum value of θ_t is zero, implying $b_1 = -2.71$. Thus, prices become more stale when the index return is low. The Scholes and Williams effect says that the fund's measured beta on the index will be biased toward zero, to a greater extent when the pricing is more stale. Thus, the measured betas will be too low when the index return is low. This creates a spurious impression of market timing ability. Regressing the 177 monthly measured excess returns on the government index excess return and its square, the coefficient on the squared term is 0.23 and its t-ratio is 2.00. Thus, a fund with systematic stale pricing can appear to have timing ability if we fail to control for the stale pricing effect. A fourth reason for nonlinearity unrelated to timing ability arises if there is public information about future asset returns. As shown by Ferson and Schadt (1996), even if the conditional relation between the fund and a benchmark return is linear, the response of managed portfolio weights to public information can induce nonlinearities in the unconditional relation. The portfolio betas may be correlated with market returns because of their common dependence on public information. Other examples of public information effects on equity fund timing coefficients are provided by Ferson and Warther (1996), Becker, et al (1999), Christopherson, Ferson and Turner (1999) and Ferson and Qian (2004).

In summary, in order to measure the market timing ability of bond funds we need to control for nonlinearity in the benchmark assets and for nonlinearity that may arise from interim trading, public information effects and systematic stale pricing. In the following subsections we modify the classical market timing model to allow for nonlinearity in the benchmark returns and develop controls for interim trading, public information effects and systematic stale pricing.

2.2 Classical Market Timing Models

The classical market-timing regression of Treynor and Mazuy (1966) is:

$$r_{pt} = a_p + b_p f_t + \Lambda_p f_t^2 + u_t, \tag{1}$$

where r_{pt} is the fund's portfolio return, measured in excess of a short-term Treasury bill. With equity market timing, as considered by Treynor and Mazuy, f_t is the excess return of the stock market index. Treynor and Mazuy (1966) argue that $\Lambda_p > 0$ indicates market-timing ability. The logic is that when the market is up, the successful market-timing fund will be up by a disproportionate amount. When the market is down, the fund will be down by a lesser amount. Therefore, the fund's return bears a convex relation to the market factor.

It seems natural to replace the equity market excess return with changes in the systematic factors for bond returns, like interest rate levels and spreads. However, if a factor is not an excess return, the appropriate sign for the timing coefficient might not be obvious. For example, bond returns move in the opposite direction as interest rates, so a signal that interest rates are about to rise means bond returns are likely to be low. We show in the Appendix that market timing ability still implies a positive coefficient on the squared factor.

Stylized market-timing models confine the fund to a single risky-asset portfolio and cash. This makes sense theoretically, from the perspective of the Capital Asset Pricing Model (CAPM, Sharpe, 1964). Under that model's assumptions there is two-fund separation and all investors hold the market portfolio and cash. But two-fund separation is generally limited to single-factor term structure models, and there is no central role for a "market portfolio" of bonds in most fixed income models. In practice, however, bond funds often manage to a "benchmark" portfolio that defines the peer group or investment style. We use style-specific benchmarks to replace the market portfolio.

2.3 Addressing Nonlinearity in Benchmark Assets

We model nonlinearity in the relation between the benchmark asset returns and the common factors with a nonlinear regression:

$$r_{Bt} = a_B + b_B(f_t) + u_{Bt}, \tag{2}$$

where $b_B(f)$ is a nonlinear function of the factor changes and we assume that u_B and $b_B(f)$ are normal and independent. The Appendix derives the generalization of the market-timing regression that incorporates the nonlinear benchmark:

$$r_{pt} = a_p + b_p\,[b_B(f_t)] + \Lambda_p\, f_t^2 + u_t. \tag{3}$$

The intuition of Equation (3) is that the nonlinearity of the benchmark return determines the nonlinearity of the fund's return. If there is no market-timing ability (Λ_p=0) the nonlinearity of the fund's return mirrors that of the benchmark through the second term of the regression. A successful timer's return has a more convex relation than the benchmark, and thus Λ_p>0. We combine equations (2) and (3), and estimate the model by the Generalized Method of Moments (Hansen, 1982).

One of the forms for $b_B(f)$ that we consider is a quadratic function, which has an interesting interpretation in terms of systematic coskewness. Asset-pricing models featuring systematic coskewness are studied, for example, by Kraus and Litzenberger (1976). Equation (1) is, in fact, equivalent to the quadratic "characteristic line" used by Kraus and Litzenberger. Under their interpretation the coefficient on the squared factor changes does not measure market timing, but measures the systematic coskewness risk. Thus, a fund's return can bear a convex relation to the factor because it holds assets with coskewness risk. Equations (2) and (3) allow the benchmark to have coskewness risk and measure timing ability as the fund's convexity in excess of its benchmark coskewness risk.

2.4 Addressing Interim Trading

Interim trading means that fund managers trade more frequently than the fund's returns are measured. This can lead to incorrect inferences about market timing ability, as shown by

Jiang, Yao and Yu (2005) and to incorrect inferences about total performance, as shown by Ferson and Khang (2002). Ferson, Henry and Kisgen (2006) propose a solution using a continuous-time asset pricing model, where the time-aggregated model prices all portfolio strategies that may trade within the period as nonanticipating functions of the state variables in the model. Thus, a manager with no ability will not record abnormal performance. If the manager wastes resources by interim trading that generates trading costs, the portfolio return will be low and this should be detected as negative average performance. If the continuous-time model can price derivatives by replication with dynamic strategies, the use of derivatives is also covered by this approach.

Ferson, Henry and Kisgen show that the time-aggregated stochastic discount factor (SDF) from a set of popular term structure models is:[6]:

$$ {}_{t-1}m_t = \exp(a - A^r_t + b'A^x_t + c'[x(t) - x(t-1)]). \tag{4} $$

In Equation (4) x(t) is the vector of state variables in the model at time t. The terms $A^x_t = \Sigma_{i=1,\dots 1/\Delta}\, x(t-1+(i-1)\Delta)\Delta$ approximate the integrated levels of the state variables over the period from t-1 to t. The monthly measurement period is divided into $(1/\Delta)$ intervals of length Δ=one trading day. A^r_t is the time-averaged level of the short-term interest rate. The empirical "factors," f_t, in the SDF thus include the usual discrete monthly changes in the state variables, but also include their time averages and the time-averaged short term interest rate: $f_t = \{[x(t) - x(t-1)], A^x_t, A^r_t\}$.

With the approximation $e^f \approx 1+f$, which is accurate for numerically small f, the SDF is linear in the expanded set of empirical factors. Since a linear SDF is equivalent to a beta pricing model, this motivates including the time-averaged variables $\{A^x_t, A^r_t\}$ as additional "factors" to control for interim trading. That is the approach we take in this paper.

[6] A stochastic discount factor is a random variable, ${}_{t-1}m_t$, that "prices" assets through the equation $E_{t-1}\{{}_{t-1}m_t(1+R_t)\}=1$.

2.5 Addressing Public Information

Conditional timing models control for public information effects by allowing funds' betas to vary over time with public information. Ferson and Schadt (1996) and Becker, et al. (1999) find that conditional timing models for equity funds are better specified than models that do not control for public information. In particular, Ferson and Schadt (1996) propose a conditional version of the market timing model of Treynor and Mazuy (1966):[7]

$$r_{pt} = a_p + b_p f_t + C_p'(Z_{t-1} f_t) + \Lambda_p f_t^2 + u_t, \tag{5}$$

where the interaction term $C_p'(Z_{t-1} f_t)$ controls for nonlinearity due to the public information, Z_{t-1}. We include similar interaction terms in this paper to control for public information effects.

2.6 Addressing Stale Prices

Thin or nonsynchronous trading in a portfolio biases estimates of the portfolio beta (e.g., Scholes and Williams, 1977), and a similar effect occurs when the measured value of a fund reflects stale prices, possibly due to illiquid assets (e.g. Getmansky, Lo and Makarov, 2004). If the extent of stale pricing is related to a common factor we call it *systematic stale pricing*. To address systematic stale pricing we use a simple model generalizing Getmansky, Lo and Makarov. Let r_t be the true return on a fund's assets. The true return would be the observed return if no prices were stale. We assume r_t is independent over time with mean μ. The measured return on the fund, r_t^*, is given by:

[7] Ferson and Schadt (1996) also derive a conditional version of the market timing model of Merton and Henriksson (1981), which views successful market timing as analogous to producing cheap call options. This model is considerably more complex than the conditional Treynor-Mazuy model, but they find that it produces similar results.

$$r_t^* = \theta_t r_{t-1} + (1-\theta_t) r_t, \qquad (6)$$

where the coefficient $\theta_t \in [0,1]$ measures the extent of stale pricing at time t. Getmansky, Lo and Makarov allow K>1 lagged returns in Equation (6) to capture longer term smoothing in hedge fund returns, but they assume that the smoothing coefficients, θ_t, are constant over time. Our model allows for time-varying smoothing coefficients but we restrict to a single lag. With K lags, the measured returns have n K-th order moving average structure. In our bond fund portfolio returns we find significant first order autocorrelations, but the second order autocorrelations are insignificant, suggesting a first-order moving average structure.

Assume that the market or factor return r_{mt}, with mean μ_m and variance σ_m^2, is independent and identically distributed over time and measured without stale prices.[8] To model systematic stale pricing consider a regression of θ_t on the market factor:

$$\theta_t = \delta_0 + \delta_1(r_{mt} - \mu_m) + \varepsilon_t, \qquad (7)$$

where we assume that ε_t is independent of the other variables in the model. We are interested in moments of the true return like $\mathrm{Cov}(r_t, r_{mt}^2)$, which measures the timing ability. Straightforward calculations relate the moments of the observable variables to the moments of the unobserved variables, as follows.

$$E(r^*) = \mu \qquad (8a)$$

$$\mathrm{Cov}(r_t^*, r_{mt}) = \mathrm{Cov}(r, r_m)(1- \delta_0 + 2\delta_1\mu_m) - \delta_1 \{\mu\, \sigma_m^2 + \mathrm{Cov}(r, r_m^2)\} \qquad (8b)$$

[8] While this modelling assumption is certainly false, it might be a reasonable approximation when r_m is a highly-traded Treasury index or one of the Lehman bond indexes. Lehman's pricing staff employs a daily matrix pricing approach based on traders' daily quotes and bond characteristics, and they market several products that focus on providing timely prices to clients (see Lehman Brothers, 2006).

$$Cov(r_t^*, r_{mt}^2) = Cov(r, r_m^2)(1 - \delta_0 + \delta_1\mu_m)$$

$$- \delta_1\{ \mu E(r_m^3) + Cov(r, r_m^3) - [Cov(r, r_m) + \mu\mu_m](\mu_m^2 + \sigma_m^2)\} \quad (8c)$$

$$Cov(r_t^*, r_{mt}) + Cov(r_t^*, r_{mt-1}) = Cov(r, r_m) \quad (8d)$$

$$Cov(r_t^*, r_{mt}^2) + Cov(r_t^*, r_{mt-1}^2) = Cov(r, r_m^2) \quad (8e)$$

$$Cov(r_t^*, r_{mt}^3) + Cov(r_t^*, r_{mt-1}^3) = Cov(r, r_m^3) \quad (8f)$$

Equation (8a) shows that stale prices will not affect the measured average returns.[9] Thus, for example we can use funds' average returns net of the benchmark as a simple performance measure. Equation (8b) captures the bias in the measured covariance with the market. The market beta is proportional to this covariance. Equation (8c) shows the measured covariance with the squared market return. When stale pricing is systematic the true beta contaminates the measured timing ability and timing ability contaminates the measured beta.

Fortunately, equation (8e) reveals a simple way to control for a biased timing coefficients due to systematically stale prices. The sum of the covariances of the measured return with the squared factor changes and the lagged squared factor changes delivers the correct timing coefficient. This is similar to the bias correction for betas in the models of Scholes and Williams (1977) and Dimson (1979), as revealed in Equation (8d).

2.7 Combining the Effects

In summary, the general form of the model is a system including Equation (2) and

---

[9] Qian (2006) studies the effects of stale prices on the measured average performance of equity style funds, and finds that if staleness is correlated with new money flows there can be effects on average returns. See also Zeitowitz (2003) and Edelin (1999). Under reasonable assumptions bounding the magnitudes of flows, however, we can show that a model with flows produces the same reduced-form implications for the measurement of market timing as the model presented here.

Equation (9):

$$r_{pt}^{*} = a + \beta' X_t + \Lambda_p [f_t^2 + f_{t-1}^2] + u_{pt}, \tag{9}$$

where X_t is a vector of control variables observed at t or before. The nonlinear function $b_B(f_t)$ is included in X_t to control for nonlinearity in the underlying benchmark assets. The time-averaged factor and short-term interest rate, $\{A^f, A^r\}$, control for interim trading. The products of the factors with lagged state variables control for public information effects. Lagged values of the factor changes, f_{t-1}, are included in X_t and in the timing term of (9) to control for systematic stale pricing.

3. The Data

We first describe our sample of bond funds. We then describe the interest rate and other economic data that we use to construct the factors relative to which we study timing ability. Finally, we describe the funds' style-related benchmark returns.

3.1 Bond Funds

The mutual fund data are from the Center for Research in Security Prices (CRSP) mutual fund data base, and include returns for the period from January of 1962 through December of 2007. We select open-end funds whose stated objectives indicate that they are bond funds.[10] We

[10] Prior to 1990 we consider funds whose POLICY code is B&P, Bonds, Flex, GS or I-S or whose OBJ codes are I, I-S, I-G-S, I-S-G, S, S-G-I or S-I. We screen out funds during this period that have holdings in bonds plus cash less than 70% at the end of the previous year. In 1990 and 1991 only the three digit OBJ codes are available. We take funds whose OBJ is CBD, CHY, GOV, MTG or IFL. If the OBJ code is other than GOV, we delete those funds with holdings in bonds plus cash totalling less than 70%. After 1991 we select funds whose OBJ is CBD, CHY, GOV, MTG, or IFL or whose ICDI_OBJ is BQ, BY, GM or GS, or whose SI_OBJ is BGG, BGN, BGS, CGN, CHQ, CHY, CIM, CMQ, CPR, CSI, CSM, GBS, GGN, GIM, GMA, GMB, GSM or IMX. From this group we delete 116 fund years for which the POLICY code is IG or CS.

exclude money market funds and municipal securities funds. We subject the fund data to a number of screens as described in the Appendix. We group the funds into equally-weighted portfolios according to eight mutually exclusive investment styles: Index, Global, Short-term, Government, Mortgage, Corporate, High Yield and Other.[11]

Summary statistics for the style-grouped funds' returns are reported in Panel A of Table 1. The mean returns are between 0.37 and 0.74% per month. The standard deviations of return range between 0.46% and 1.85% per month. The first-order autocorrelations range from 14% for Index funds to 30% for Short Term funds. The minimum return across all of the style groups in any month is -7.3%, suffered in October of 1979 by the Corporate bond funds. The maximum return is almost 11%, also earned by the Corporate funds, in November of 1981.

Table 1 also reports the second order autocorrelations of the fund returns. The stylized stale pricing model assumes that all the assets are priced within two months. This implies that the measured returns have an MA(1) time-series structure, and the second order autocorrelations should be zero. The largest second order autocorrelation in the panel is 14.9%, with an approximate standard error of $1/\sqrt{T} = 1/\sqrt{159} = 7.9\%$. For the portfolio of all funds, where the number of observations is the greatest, the second order autocorrelation is -7.1%, with an approximate standard error of $1/\sqrt{543} = 4.3\%$. Thus, none of the second order autocorrelations is significantly different from zero, consistent with the assumptions of the model.

We group the funds into equally-weighted portfolios according to various fund

[11] Global funds are coded SI_OBJ=BGG or BGN. Short-term funds are coded SI_OBJ=CSM, CPR, BGS, GMA, GMS, GBS or GSM. Government funds are coded OBJ=GS POLICY=GOV, ICDI_OBJ=GS, or SI_OBJ=GIM or GGN. Mortgage funds are coded ICDI_OBJ=GM, OBJ=MTG or SI_OBJ=GMB. Corporate funds are coded as OBJ=CBD, ICDI_OBJ=BQ, POLICY=B&P or SI_OBJ=CHQ, CIM, CGN or CMQ. High Yield funds are coded as ICDI_OBJ=BY, SI_OBJ=CHY or OBJ=CHY or OBJ=I-G and Policy=Bonds. Index funds are identified by searching for the word "index" in the fund name. Other funds are defined as funds that we classify as bond funds (see the previous footnote), but which meet none of the above criteria.

characteristics, measured at the end of the previous year. Since the characteristics are likely to be associated with fund style, we form characteristic groups within each of the style classifications. The characteristics include fund age, total net assets, percentage cash holdings, percentage of holdings in options, reported income yield, turnover, load charges, expense ratios, the average maturity of the funds' holdings, and the lagged return for the previous year. The Appendix provides the details.

3.2 Bond Market Factor Data

We use daily and weekly data to construct monthly empirical factors. Most of the data are from the Federal Reserve (FRED) and the Center for Research in Security Prices (CRSP) databases. The daily interest rates are from the H.15 release. The factors reflect the term structure of interest rates, credit and liquidity spreads, exchange rates, a mortgage spread and two equity market factors. The Appendix provides the details.

Table 2 presents summary statistics for the monthly series starting in January of 1962 or later, depending on data availability, and ending in December 2007. Missing values are excluded and the units are percent per year (except for the US dollar index and Equity values, represented as the price/dividend ratio). Panel A presents the levels of the variables and panel B presents the monthly first differences. The time-averaged values used as controls for interim trading effects look similar to the levels in Panel A and are not shown.

The average term structure slope was positive, at just over 80 basis points during the sample period. The average credit spread was about one percent, but varied between 32 basis points to about 2.8%. The average mortgage spread over Treasuries was just over 2% for the period starting in 1971, and varied between -0.4% to 5.6%. The liquidity spread averaged 0.4%, with a range between -0.15% and 2.2% over the 1971-2007 period.

In their levels the variables shown in Table 2 are highly persistent time series, as indicated by the first order autocorrelation coefficients. Five of the nine autocorrelations exceed 95%. Moving to first differences, the series look more like innovations.[12] We use the first differences of these variables to represent the factor changes in our analysis.

3.3 Style Index Returns

We form style-related benchmark returns for the mutual funds using two alternative methods. The first method is to simply assign a benchmark based on a fund's declared style. This has the advantage that the benchmark is determined *ex ante* and nothing has to be estimated. It has the disadvantage of relying on the fund's self-declared style. If a fund strategically misrepresents its style or is more accurately represented as a hybrid style, then the benchmark will be inaccurate.

We select seven benchmarks based on funds' declared styles. Global funds are paired with the Lehman Global Bond Index. Short-term bond funds are paired with a portfolio of US Treasury bond returns with less than or equal to 48 months to maturity. This best matches the reported maturities of their holdings (see the Appendix). Mortgage funds are paired with the Lehman US Mortgage Backed Securities Index. Corporate bond funds are paired with the Lehman US AAA Credit Index, while High-yield funds are paired with the Merrill Lynch High-Yield US Master Index.[13] For two of the styles we use combinations of Lehman bond indexes,

---

[12] Given the relatively high persistence and the fact that some of the factors have been studied before exposes us to the risk of spurious regression compounded with data mining, as studied by Ferson, Sarkissian and Simin (2003). However, Ferson, Sarkissian and Simin (2008) find that biases from these effects are largely confined to the coefficients on the persistent regressors, while the coefficients on variables with low persistence are well behaved. Thus, the slopes on our persistent control variables and thus the regression intercepts may be biased, but the market timing coefficients should be well behaved.

[13] We splice the Blume, Keim and Patel (1991) low grade bond index prior to 1991, with the Merrill Lynch High Yield US Master Index after that date.

weighted in proportion to their contributions to the Lehman US Aggregate bond index. Because Government style funds hold significant amounts of mortgage-backed securities in the latter part of the sample period, we pair the Government bond funds with a combination of the Lehman mortgage backed index and a long-term Treasury bond index.[14] Finally, we pair the catch-all Other bond funds and the Index funds with a combination of the Treasury bond index, the mortgage backed index and the corporate bond index.[15]

Our second, alternative method follows Sharpe (1992). Historical returns are used to estimate a fund-specific tracking portfolio of passive asset class returns. This has the advantages of not relying on a fund's self-declared style and of allowing a fund to be represented as a hybrid style. It has the disadvantage that the portfolio weights must be estimated, and the estimates will be imprecise for funds with a limited sample of returns. If the portfolio weights for a particular period are estimated using any future returns data, there may be a look-ahead bias in the analysis for the future period. The details of this approach are discussed in the Appendix.

Panel B of Table 1 presents the Sharpe style-index weights for the style-based portfolios. The weights present sensible patterns, suggesting that both the style classification of the funds and Sharpe's procedure are reasonably valid. The Global funds load most heavily on global bonds.

---

[14] We splice the Ibbotson Associates 20 year government bond return series for 1962-1971, with the CRSP greater than 120 month government bond return after 1971.

[15] The aggregate weights are from Lehman (2006), in percent:

Year	Government	Credit	Mortgage
1976	49	42	5
1986	56	19	23
1996	52	17	30
2001	34	27	35.

We use the most recent ex ante weight to form the benchmark in a given year. Prior to 1977 we pair the Government funds with the long-term Treasury return index and we pair the Other and Index funds with the Corporate bond index.

Short term funds have most of their weight in bonds with less than 48 months to maturity. Mortgage funds have their greatest weights on mortgage backed securities. Corporate funds have more than 70% of their weight in high or low grade corporate bonds. High yield funds have 77% of their weight in low grade corporate bonds. Government funds do load highly on mortgage-backed securities (21%), consistent with the observations of Comer (2006) and Moneta (2008).

4. Empirical Results

We first examine the empirical relations between the factors and passive investment strategies proxied by the style benchmarks. We evaluate the effects of the controls for nonlinearities on these portfolios and on broad portfolios of the mutual funds. We then apply the models to individual funds.

4.1 Factor Model Regressions

We begin the empirical analysis with regressions of the style benchmark returns on changes and squared changes in the factors, looking for convexity or concavity. Nonlinearities in the relations between the factors and the benchmarks suggest what would happen given a naive application of the timing regression (1) for funds, if funds simply held the benchmark portfolios. If the benchmark returns are nonlinearly related to the factors, it suggests that controls for nonlinearities could be important.

Table 3 summarizes the t-ratios for the regression coefficients on the squared factors. Only t-ratios that exceed 1.6 in absolute value are shown; otherwise a zero is recorded. Panel A of Table 3 shows that out of 72 cases (8 styles times 9 factors) there are 25 heteroskedasticity-consistent t-ratios with absolute values larger than 1.6. Nine of the absolute t-ratios are above 2.0. Thus, there is significant evidence of nonlinearity in the benchmark returns. Most of the large coefficients are positive, indicating convexity. Using the Sharpe style benchmarks (not shown in the table) the evidence for convexity is even stronger. This implies that we would measure positive timing ability, based on regression (1), if funds simply held the benchmarks. Thus, controlling for nonlinearity of the benchmark returns is likely to be important for measuring the timing ability of bond funds.

Comparing the funds with the benchmarks, we see the effects of active management. The coefficients for the mutual funds are summarized in Panel B of Table 3. Here we find 11 absolute t-ratios larger than two and 19 in excess of 1.6. This is more than expected by chance; thus, the mutual fund returns are significantly nonlinearity related to the factors.[16] The coefficients on the squared factors are negative in about half of the cases. Thus, the fund returns appear more concave in relation to the factors than do the benchmark returns. This might suggest poor market timing ability on the part of the mutual funds, but we would not expect to find much evidence of market timing for an entire style portfolio of mutual funds. More likely, nonlinearity that differs from the benchmark reflects derivatives, interim trading, public information or stale pricing effects that are not found in the benchmarks. Note that marginally significant t-ratios are found for the index funds on three of the nine factors, as shown in the first row of Panel B. This further suggests that the nonlinearity is unrelated to timing, as index funds are unlikely to be actively market timing.

We also examine regressions like Table 3 where the fund style returns net of the benchmark returns are the dependent variables. The evidence of concavity is much stronger in these regressions, as would be expected. We find 27 or 28 absolute t-ratios larger than 1.6 and all but one or two are negative, depending on the type of style benchmark. Thus, simply measuring fund returns net of a benchmark will not control for the nonlinearities. If we naively ran the

---

[16] Using a simple binomial model assuming independence, the t-ratio associated with finding 11 "rejections," when the probability of observing a rejection is 5%, is $(11/72 - .05)/(.05(.95)/72)^{0.5} = 4.00$.

regression (1) using fund returns net of benchmark returns on the left hand side, we would find strong evidence of negative timing ability.

Finally, Table 3 reports the adjusted R-squares of the regressions when all nine factors and their squares are included. We would expect the R-squares to be higher for the benchmarks than for the funds if active management in the form of security selection creates nonsystematic noise in the funds. We would expect higher R-squares for the funds if managers are timing systematic factors. A large difference in the R-squares could also indicate bad benchmarks. The average R-squared across the funds is 51.6% and the average across the benchmarks is 53.8%. The R-squares vary across the styles with similar patterns for both the funds and the benchmarks; e.g. the index styles have the highest R-squares and High yield styles have the lowest. While there are some differences between the funds and the benchmarks, the magnitudes of these differences suggest no obvious problems.

4.2 Evaluating the Controls

The next step in the analysis is to evaluate the impact of the controls for non-timing-related nonlinearity in the general model represented by Equations (2) and (9). To control for benchmark nonlinearity the function $b_B(f_t)$ is used. We consider three specifications for $b_B(f)$: Quadratic, exponential and piecewise linear. As described earlier, a quadratic function can be motivated by coskewness. An exponential function can be motivated by a continuous-time model, however, since $e^f \approx 1+f$ when f is numerically small, the linear function well approximates the exponential for small factor changes. We find no empirically measurable impact of an exponential function, compared with a linear function.

The piecewise linear specification is $b_B(f) = bf + c\, f^*I(f>0)$, where $I(f>0)$ is an indicator function for a positive change in the factor. A piecewise linear function can be motivated as approximating an option payoff. We use zero as the breakpoint in the piecewise linear function for simplicity and to avoid estimating a breakpoint parameter. This will be useful when we apply the model at the level of the individual funds, where short time series often limit the degrees of freedom.

We revisit the factor model regressions of Table 3 using the piecewise linear function in place of the quadratic function to capture nonlinearity. The results for the benchmark returns are similar. The coefficient, c, is positive, indicating convexity in the benchmarks, and the t-ratio is large in significant fractions of the cases. The coefficients for the fund portfolios are a mix of positives and negatives, with large t-ratios in a significant fraction of the cases. The funds are typically more concave in the factor changes than are the benchmarks.

Table 4 summarizes the impact of the controls for nonlinearity in the fund-style-and-characteristic portfolio returns, as measured by the t-ratios of the timing coefficients. The table reports the fractions of the cases where the t-ratios are above +2.0 or below -2.0. Panel A sorts the cases by fund style. There are 180 possible cases for each style (9 factors times 20 characteristics-sorted portfolios). Panel B sorts by the factors, which enter the models one at a time. There are 160 possible cases for each factor. Results for the Sharpe style benchmarks are shown.

The first column of Table 4 shows the results when there are no controls for nonlinearity in the models. We find negative timing coefficients on the fund-characteristic portfolios, as we would expect from the evidence in Table 3. For example in panel A, the Global, Short Term, Government and High yield portfolios have more than 7.5% of the t-ratios below -2.0. Where we saw the largest differences between the t-ratios for the funds versus the benchmarks in Table 3, we tend to find thicker tails in the same direction in the models with no controls. For example, we find positive coefficients on the US dollar factor and negative coefficients on the short term interest rate factor.

In the second and third columns of Table 4 we introduce the quadratic and piecewise linear controls for benchmark nonlinearity. These have a substantial impact on the distributions of the timing coefficients. In most of the cases the controls for benchmark nonlinearity reduces the incidence of large t-ratios, relative to the models with no controls. With the quadratic specification all of the fractions of large t-ratios are 6.5% or fewer. With the piecewise linear specification only the equity volatility factor produces a fraction above 7.5%, but this is sensitive to the benchmark portfolio used.

The controls for interim trading, stale pricing and public information have a smaller impact on the distributions of the t-ratios. We ran an alternative version of this experiment, where the took the model with all the controls and compared the timing coefficients with a model where one control at a time was removed. Here, we found that the controls for interim trading, stale pricing and public information had marginally significant effects, with interim trading being the largest.

The far right column of Table 4 summarizes models with all of the controls in place. The piecewise linear specification is used for benchmark nonlinearity. Here only three of 32 experiments produce fractions of large t-ratios above 7.5%. With the self-declared benchmarks none of these is larger than 7.5%. Most of the other patterns in the table are not sensitive to which type of style benchmark is used. Interestingly, none of the cases with large t-ratios in this model indicate negative timing coefficients. The full models thus suggest neutral to perhaps slightly positive timing ability at the level of the fund style-and-characteristics portfolios.

We draw several conclusions from the analysis of Table 4. First, it is important to control for nonlinearity in funds' benchmarks with respect to the factors. Second, a model that combines the various controls produces an overall distribution of the timing coefficients that appears neutral to perhaps slightly positive at the level of fund portfolios.

4.3 Fund level Analysis

We expect to find little timing ability when funds are grouped into large portfolios, but there could be individual funds with significant timing ability. We estimate the market timing coefficients by combining equations (2) and (9) for each fund with at least 36 monthly returns. The results using the piecewise linear $b_B(f)$ function and all of the controls are summarized in Table 5 (the results are similar using the quadratic function as noted below).

There are at least 1100 eligible funds for each factor, and we summarize the distributions of the timing coefficients across the funds. The first column lists selected fractiles of a null distribution of estimated timing coefficients. For each factor we set the timing coefficient in regression (9) equal to zero and simulate the each fund return as the fitted values of the regression with this modification, plus the randomly rescrambled residuals. Estimating the model on this simulated data, we sort the timing coefficients and find the critical values at each fractile. The fractions in the remaining columns are the fractions of the estimated timing coefficients in the original fund data that exceed these critical values. (For left tail areas we show the fractions that lie below the critical values.)

To evaluate these figures, consider a binomial random variable that equals 1.0 with probability p, when a timing coefficient is larger than the critical value for the fractile p. If the correlation of the trials is ρ the variance of the fraction of funds above the critical value is $[p(1-p)\{1/n -(1-1/n)\rho\}]$. The correlation ρ depends on the correlation of the funds' returns. We approximate ρ by estimating the pairwise correlations of all the funds' returns using all pairs with at least 36 months in common. The summary statistics in Table 1 suggest that if two funds' return series are separated by more than a month in calendar time, their average correlation is statistically zero.[17] We therefore scale each contemporaneous correlation by the fraction of the sample where the two series overlap. We make the conservative assumption in this calculation,

[17] Recall that the autocovriance of a portfolio return is approximately the average of the lagged cross covariances of the securities in the portfolio. Since the second order autocorrelations in Table 1 are statistically zero, it suggests no lagged cross correlation of individual fund returns beyond the first lag.

that the correlations at one lag are equal to the contemporaneous correlations and we set the correlations beyond lag one to zero. The resulting estimate of ρ is 0.083. This implies that the standard deviations of the fractions in Table 5 are approximately 0.14, 0.085, 0.06 and 0.045 respectively, for the 0.50, 0.10, 0.05 and 0.025 fractiles.

Overall, the timing coefficients are mildly skewed toward positive values, relative to the null distributions, excepting the coefficients on the slope and curvature factors, which are mildly negatively skewed. For most of the factors, however, the distributions of the coefficients are not significantly different from the null distributions. For the US dollar factor, we do find significant positive timing, with more than 82% of the funds' timing coefficients above the median from the null distribution, but the 5% and 2.5% tails conform closely to the null distribution. (Using the quadratic $b_B(f)$ there are significant positive timing coefficients in the tail as well.) There is marginally significant evidence that some funds can time the credit, liquidity and equity value factors. For the curvature factor, we find more negative coefficients, with only 28% of the funds' coefficients above the median from the null distribution. (Using the quadratic $b_B(f)$ the negative coefficients are significant below the median.)

Panel B reports the estimated correlations of the timing coefficients with the ten fund characteristics described earlier. With about 1200 observations, the standard error of these correlations is about 0.03. There are 38 out of 171 correlations larger than 0.06. Assuming a 5% significance of two standard errors, the binomial t-statistic, accounting for correlation across trials equal to 0.08, is 2.7. Thus, there are significant correlations. The strongest correlations with the timing coefficients are found for the short rate, slope and US dollar factors. For these, timing is negatively related to the expense ratio, the lagged return and the global style dummy, but the other correlations are mixed.

4.4 Timing-adjusted Performance

The idea in most measures of investment performance is to compare the average return of a managed portfolio over some evaluation period to the return of a benchmark portfolio. The benchmark portfolio ideally represents a feasible investment alternative to the managed portfolio. If the objective is to evaluate the investment ability of the portfolio manager or management company, the benchmark should be equivalent in all of the return-relevant aspects to the managed portfolio being evaluated, except that it should not reflect the investment ability of the firm or manager. Aragon and Ferson (2007) call such a portfolio an "*Otherwise Equivalent*" (OE) portfolio. In order to operationalize the OE portfolio it is necessary to have a model to determine what aspects of a portfolio or its return should lead to higher or lower expected returns. Perhaps the simplest example is the CAPM, where the relevant characteristic is the market beta, leading to Jensen's (1968) alpha as the measure of performance.

The intercept in the Treynor-Mazuy regression (1) has been naively interpreted as a "timing adjusted" selectivity measure in several studies. This is incorrect, except in unlikely special cases of stylized "perfect" market timing ability, as discussed by Aragon and Ferson (2007). The reason the intercept does not properly measure performance is that f_t^2 is not a portfolio's excess return. However, the model can be modified to capture the difference between the return of the fund and that of an OE portfolio. Our approach to performance measurement follows Aragon and Ferson (2007).

Let r_{h2} be the excess return of the maximum-squared-correlation portfolio for the squared factor changes, f_t^2. This portfolio is estimated by the regression:

$$f_t^2 = a + H'r_t + u_t, \tag{10}$$

where the weights that define the mimicking portfolio r_{h2} are proportional to the regression coefficients, H. The base assets in r_t are the seven assets we use to form the Sharpe style benchmarks, excluding the short term Treasury rate, and the returns are in excess of this Treasury rate. The expected value of the excess return, $E(r_{h2})$ is the risk premium associated with the squared factor, which we call the convexity premium.[18] Our goal is to form an OE portfolio for each fund that has the same loadings on its style benchmark and r_{h2} as does the fund. This makes the simplifying assumption that style benchmark exposure and timing are the return-relevant characteristics. Our measure of alpha is the excess return of the fund net of the excess return of its OE portfolio. The OE portfolio is formed using the following regression:

$$r_{pt} = a_p + b_p\, r_{Bt} + c_p\, r_{Bt-1} + d_p\, r_{h2t} + e_p\, r_{h2t-1} + v_t. \tag{11}$$

The loading on the benchmark, r_B, is estimated following Scholes and Williams (1977) as $\beta_p = b_p + c_p$ to account for stale pricing in the fund returns. Similarly, the loading on the hedge portfolio is $\Omega_p = d_p + e_p$. The OE portfolio return for fund p is $\beta_p\, r_B + \Omega_p\, r_{h2}$. This portfolio has the same benchmark exposure as the fund and the same conditional exposure to the hedge portfolio (given the nonlinearity captured by the style benchmark return). The contribution of r_{h2} to the OE portfolio is the premium in the market that is required to replicate the fund's exposure to the squared factor. The fund's performance is its average return net of what the unmanaged strategy would cost to produce these return characteristics.

Table 6 presents the analysis of timing adjusted performance. The second and third

[18] An alternative approach is to use cross-sectional regressions of returns on betas to estimate mimicking portfolios and factor risk premiums. While we have a large cross section of mutual funds, using the funds with this approach would contaminate the premium estimates with the abnormal returns due to manager ability, if any. We have a small cross section of passive benchmark assets. We therefore use time-series as opposed to cross-sectional regressions to estimate the mimicking portfolios. (See Balduzzi and Robotti, 2007 for comparisons of the two methods.)

columns of Panel A summarize the mean excess return over a short term treasury and the mean return over the style benchmark at selected fractiles of their distributions across the funds. The average excess returns are percent per month. The remaining columns of Panel A of Table 6 report the fractions of funds with alphas larger than (for the right tail and median) or smaller than (for the left tail) the critical values for the indicated fractiles from the null distribution in which the true alphas are zero. Panel B summarizes the average risk premiums for the squared factors (denoted convexity premiums), which are the average excess returns of mimicking portfolios for the squared factors. It also summarizes the distributions of loadings on style benchmarks and the hedge portfolios for the squared factors.

The median fund excess return is 0.46% per month, about 20 basis points below the mean in Table 1. This reflects a skewed cross-sectional distribution of funds' average returns. The upper 10% tail of the returns distribution is above 0.71% while the lower 10% tail is below 0.29%. The median return net of benchmark is -0.07%, consistent with previous studies that find bond funds return less than benchmarks on average. The distribution is skewed to the left, with 5% of the funds below -0.28% and 5% above 0.10%.

The third column of Table 6 shows the values of the estimated alphas under the null hypothesis that the true alphas are zero, taken at various fractiles of the distribution across funds. These critical values suggest that the estimates of alpha have smaller variation than the average returns net of the style benchmarks. Thus, they should provide more precision in measuring performance than the returns net of benchmark. For example, the range between the top and bottom 10% of the alphas is 0.13%. The range is 0.23% for the returns net of benchmark.

Many funds appear to have significantly negative timing-adjusted alphas. Depending on the factor, 78-86% of the funds have alphas below the median value of zero under the null. Between 10-18% of the funds have alphas below -0.15% per month, which is at the left 2.5% tail of the null distribution. The alphas in the right tails, by contrast, are not significantly different from the null distributions. The table provides strong evidence of negative timing-adjusted performance on an after-cost basis.

Panel B of Table 6 digs more deeply into the structure of the timing adjusted alpha analysis. We first summarize the funds' loadings on the style benchmarks. Funds' loadings vary widely in the cross section, with the 10% tails spanning loadings between 0.42 and 1.38. This shows that the returns net of benchmark are crude performance measures. They assume that all of the betas equal 1.0.

The average convexity premiums vary from -2.93% per month for equity volatility to 3.5% for equity values, but the range of funds' loadings on these factors is narrow, with 80% of the funds between -0.12 and +0.09. The rest of the convexity premiums are an order of magnitude smaller than these. If market timing or convexity is valuable, we expect negative return premiums for portfolios that are positively correlated with squared factor changes. This implies that the convexity premiums, given by the mean of r_{h2}, multiplied by the sign of the correlation between the maximum-squared-correlation portfolio and the factor changes should be negative. This correlation is shown in the last line of Table 6. The product is negative for each of the nine factors, excepting the interest rate curvature factor.

The largest effect of the timing adjustments on funds' required returns is the short rate convexity factor exposure, which contributes (0.415)(0.134)=0.06% per month for a fund in the upper ten percentile. For the median fund the affect is slightly positive. For a given fund return, these positive timing adjustments imply smaller alphas.

We estimate the correlations of the alphas with the ten fund characteristics described earlier. Statistically significant correlations are found, with 68 out of 171 correlations in excess of 0.06. The strongest correlations are again found when the factors are the short rate and term slope, although significant correlations are found for other factors. We find a strong negative relation of alpha to the fund expense ratio and a positive relation to turnover for all factors.

These correlations are consistent with what other studies have found for equity funds. There is also positive correlation with the Corporate and Global style dummies, and a weak negative relation with the fund's reported yield.

4.5 Before-Cost Performance

Since the OE portfolio pays no trading costs while the bond funds do, the alphas reflect a mixed message. This is consistent with the approach in much of the performance measurement literature. If investors could replicate the OE portfolios at negligible cost, then these alphas proxy for the value added for investors. In this section we replicate the analysis with the funds' returns measured on a before-cost basis. If a fund returns more than the OE portfolio on a before cost basis, we have evidence of investment ability. To obtain the before-cost returns we add back the average expense ratio plus a measure of trading costs. The trading costs for each fund are a round-trip trading cost estimate based on the fund's style multiplied by the average reported turnover of the fund.[19]

Table 7 presents the results. The median fund return is about 0.11% higher than in Table 6, reflecting a transactions cost of about 1.3% per year. The median return net of the style

[19] The round trip transaction cost figures are as follows. For Global bond funds we use 31 basis point, based on figures in Bias and Declerk (2006). This is an average of twice the half-spreads from their Table 5 plus the information content from their Table 10, weighted in proportion to the numbers of eurobonds and Sterling bonds in their sample. For corporate bonds we use 48 basis points and for high yield bonds we use 75 basis points. These figures are averages from Edwards, Harris and Piwowar (2006), Bessembinder, Maxwell and Venkataraman (2006) and Hotchkiss et al. (2006) for intermediate trade sizes. For Government funds we use 12.5 basis points, following Ferson, Henry and Kisgen (2006). For Mortgage funds and Short-term funds we use 20 basis points. For Index funds and Other bond funds, we use the average of these figures, or 34 basis points. We checked these figures with a trader at Smith-Breeden, John Sprow, who suggested the figures for Mortgage and Short-term bond funds and confirmed that the other numbers seemed reasonable for trade sizes typical of mutual funds under average market conditions.

benchmark is 0.03% per month, which suggests some investment ability for the median fund. The distributions of the before-cost alphas are markedly different from those in Table 6. About 75% of the funds generate before-cost alphas above the median under the null hypothesis that the before cost alphas are all zero. Between 10-15% of the funds have alphas above 0.149% per month, which is at the upper 2.5% tail of the null distribution. The alphas in the left tails, by contrast, are not significantly different from the null distributions.

Comparing tables 6 and 7, the story is similar to what the literature finds for equity funds. After paying costs a significant number of funds have poor performance, and few can generate significant positive performance. Before costs are subtracted, just the opposite is true. A significant number of funds have investment ability, and the performance is consistent with the null of no ability in the left tails.

5. Concluding Remarks

A priori, if bond market factors explain a large part of the variance of a typical bond return, a large fraction of the potential performance of fixed income funds could be attributable to timing common factors. Thus, the nonlinear payoff profiles of bond funds, as implied by market timing, could in principle justify the low expected returns of bond funds that previous studies have documented. Models of market timing measure convexity in the relation between the fund's return and the common factors. However, convexity or concavity is likely to arise for reasons unrelated to timing ability. We adapt classical market timing models to bond funds by controlling for other sources of nonlinearity, such as the use of dynamic trading strategies or derivatives, portfolio strategies that respond to publicly available information, nonlinearity in the benchmark assets and systematically stale prices.

We find that controlling for non-timing-related nonlinearity matters, and naive applications of market timing models without these controls would be misleading. Bond funds' returns are typically more concave, in relation to a broad set of bond market factors, than are unmanaged benchmarks. Thus, without controls for non-timing-related nonlinearity, funds would appear to have very poor (i.e., negative) market timing ability. When we introduce the controls the distribution of the timing coefficients appears neutral at the fund style-group level and neutral to slightly positive in the sample of individual mutual funds.

The magnitudes of the timing-related components of funds' average returns seem small, and many funds' timing adjusted alphas are significantly negative. These results appear broaden the "puzzle" of active mutual fund management as posed by Gruber (1996). When we add expense ratios and estimates of trading costs based on fund turnover back to the funds' returns, the adjusted performance is markedly better. About 75% of the funds earn positive before-cost alphas, adjusted for market timing, and the distribution of fund performance is significantly better than would be expected if the null hypothesis that alpha was zero is true. Thus, we find evidence of investment ability in bond funds, but no evidence of value added for investors.

Our paper also contributes to the emerging literature on bond fund performance by laying out a set of methodological issues and suggesting questions that deserve more research. We find that simple returns net of style benchmarks are not likely to be reliable performance measures because funds' loadings on the benchmarks differ substantially from 1.0. More precision is available with alpha estimates that adjust for timing-related nonlinearity.

Goetzmann et. al (2000) find that successful market timing at a daily frequency by equity funds may produce returns that do not show up as convexity in monthly returns; so monthly data may have low power to detect market timing ability. Bollen and Busse (2xxx) find evidence of timing ability for equity funds using daily data. However, the evidence using daily data for bond funds is not known. Future research should adapt our controls to the study of bond fund timing using daily data. Non-timing-related nonlinearities are likely to remain important in daily data, and daily bond fund returns data present new challenges. For example, income distributions are sometimes treated as accrued interest in bond funds (Morey and O'Neal, 2005) which complicates the interpretation of daily returns. However, the potentially improved power motivates more work along these lines.

Appendix

A.1 Market Timing Models

Assume that the fund manager combines a benchmark portfolio with return R_B and a short-term Treasury security or "cash" with known return R_f. The portfolio weight on R_B is x(s), where s is the private timing signal. The managed portfolio return is $R_p = x(s)R_B + [1-x(s)]R_f$. The signal is observed and the weight is set at time t-1, the returns are realized at time t, and we suppress the time subscripts when not needed for clarity. In the simplest example the factor changes and the benchmark's excess returns are related by a linear regression (we allow for nonlinearities below):

$$r_{Bt} = \mu_B + b_B f_t + u_{Bt}, \tag{A.1}$$

where $r_B = R_B - R_F$ is the excess return, $\mu_B = E(r_B)$, the factor changes are normalized to have mean zero and u_{Bt} is independent of f_t. Assume that the signal $s = f + v$, where v is an independent, mean zero noise term with variance, σ_v^2. The manager is assumed to maximize the expected value of an increasing, concave expected utility function, $E\{U(r_p)|s\}$. Finally, assume that the random variables (r,f,s) are jointly normal and let $\sigma_f^2 = Var(f)$. With these assumptions the optimal portfolio weight of the market timer is:[20]

---

[20] The first order condition for the maximization implies:
$E\{U'(.)r_B|s\} = 0 = E\{U'(.)|s\}\ E\{r_B|s\} + Cov(U'(.),r_B|s\}$, where U'(.) is the derivative of the utility function. Using Stein's (1973) lemma, write the conditional covariance as: $Cov(U'(.),r_B|s\} = E\{U''(.)|s\}\ x(s)\ Var(r_B|s)$. Solving for x(s) gives the result, with
$\lambda = -E\{U'(.)|s\}/E\{U''(.)|s\} > 0$.

$$x(s) = \lambda\, E(r_B|s)/\sigma_B^2, \tag{A.2}$$

where $\lambda > 0$ is the Rubinstein (1976) measure of risk tolerance, which is assumed to be a fixed parameter, and $\sigma_B^2 = Var(r_B|s)$, which is a fixed parameter under normality.

The implied regression for the managed portfolio's excess return follows from the optimal timing weight x(s) and the regression (A.1). We have $r_p = x(s)\, r_B$, then substituting from (A.1) and (A.2) and using $E(r_B|s) = \mu_B + b_B\, [\sigma_f^2/(\sigma_f^2 + \sigma_v^2)]\, (f+v)$, we obtain equation (1), where $a_p = \lambda\mu_B^2/\sigma_B^2$, $b_p = (\lambda\mu_B b_B/\sigma_B^2)\, [1 + \sigma_f^2/(\sigma_f^2+\sigma_v^2)]$ and $\Lambda_p = (\lambda/\sigma_B^2) b_B^2\, [\sigma_f^2/(\sigma_f^2+\sigma_v^2)]$. The error term u_{pt} in the regression is a linear function of u_B, v, vu_B, fu_B and vf. The assumptions of the model imply that the regression error is well specified, with $E(u_p\, f) = 0 = E(u_p) = E(u_p\, f^2)$.

The model shows that timing ability implies convexity between the fund's return and the systematic factor changes, independent of the direction of the relation between the factor changes and the benchmark return. That is, since $\lambda > 0$ the coefficient $\Lambda_p \geq 0$, independent of the sign of b_B. If the manager does not receive an informative signal then $\Lambda_p = 0$ because $E(r_B|s)$ and x(s) are constants.

A.2 Nonlinearity

The manager's market-timing signal is now assumed to be $s = b_B(f) + v$, where v is normal independent noise with variance, σ_v^2. This captures the idea that the manager focuses on the return implications of information about the factor changes.

The optimal weight function in (A.2) obtains with:
$E(r_B|s) = \mu_B\, [\sigma_v^2/(\sigma_f^{*2} + \sigma_v^2)] + [\sigma_f^{*2}/(\sigma_f^{*2} + \sigma_v^2)][a_B + b_B(f) + v]$, where $\sigma_f^{*2} = Var(b_B(f))$. Substituting as before we derive the nonlinear regression for the portfolio return:

$$r_{pt} = a_p + b_p [b_B(f_t)] + \Lambda_p [b_B(f_t)^2] + u_t, \tag{A.3}$$

with: $a_p = \lambda\, a_B(\mu_B\, \sigma_v^2 + a_B\, \sigma_f^{*2})/[\sigma_B^2(\sigma_f^{*2}+\sigma_v^2)]$,

$b_p = \lambda(\mu_B\, \sigma_v^2 + 2a_B\, \sigma_f^{*2})/[\sigma_B^2(\sigma_f^{*2}+\sigma_v^2)]$,

and $\Lambda_p = (\lambda/\sigma_B^2)[\sigma_f^{*2}/(\sigma_f^{*2}+\sigma_v^2)]$.

The reduced form Equation (3) of the text ignores some of the restrictions in (A.3); for example, if $b_B(f)$ is quadratic there are restrictions involving f^3 and f^4. In our data we find that these higher order terms are empirically negligible. For this reason we replace $b_B(f_t)^2$ with f_t^2 in Equation (3).

A.3 Screening the Fund Sample

There are a total of 40,390 fund-year records in our initial sample. In order to address back-fill bias we remove the first year of returns for new funds, and any returns prior to the year of fund organization, a total of 2,625 records. Data may be reported prior to the year of fund organization, for example, if a fund is incubated before it is made publicly available (see Elton, Gruber and Blake (2001) and Evans, 2006). Extremely small funds are more likely to be subject to back-fill bias. We delete cases where the reported total net assets of the fund is less than $5 million. This removes 5,698 records. We delete all cases where the reported equity holdings at the end of the previous year exceeds 10%. This removes 1,017 records. We identify cases where funds report multiple share classes. Multiple classes are identified when two ICDI codes for the same year have a common fund name and a different share class code. We retain the share class with the largest Total Net Assets and delete the other share classes. This removes another 10,723 records. After these screens we are left with 20,236 fund-years. The number of funds with some monthly return data in a given year is four at the beginning of 1962, rises to 14 at the beginning of 1973, to 564 by 1993 and to 1,054 at the beginning of 2007.

A.4 Funds Grouped by Characteristics

The fund characteristics include age, total net assets, percentage cash holdings, percentage of holdings in options, reported income yield, turnover, load charges, expense ratios, the average maturity of the funds' holdings, and the lagged return for the previous year. Each year we sort the funds of a given style with nonmissing characteristic data from high to low on the basis of the previous year's value of a characteristic and break them into thirds. We form equally weighted portfolio returns from the funds in the high group and the low group for each month of the next year.

A.5 Bond Market Factor Data

Three factors represent the term structure of Treasury yields: A short-term interest rate, a measure of the term slope and a measure of the curvature of the yield curve. The short-term interest rate is the three-month Treasury rate. The slope of the term structure is the ten-year yield less the one-year yield. The curvature measure is: $y_3 - (y_7 + 2y_1)/3$, where y_j is the j-year fixed-maturity yield.

Since our funds hold corporate bonds subject to default risk and mortgage backed securities subject to prepayment risks, we construct associated factors. Our credit spread series is the yield of Baa corporate bonds minus Aaa bonds, from the FRED. These series are measured as the weekly averages of daily yields. We use the averages of the weeks in the month for our time-averaged version of the spread. For the discrete changes in the spread we use the first differences of the last weekly values for the adjacent months. The first difference series may not be as clean as with daily data, but we are limited by the data available to us. Our mortgage spread is the difference between the average contract rate on new conventional mortgages, also available weekly from the FRED, and the yield on a three-year, fixed maturity Treasury bond. Here we use daily data on the Treasury bond and weekly data on the mortgage yield to construct the time averages and discrete changes.

For market timing we are interested in market-wide fluctuations in liquidity. Our measure follows Gatev and Strahan (2006), who advocate a spread of commercial paper over Treasury yields as a measure of short term liquidity in the corporate credit markets. (See also, Bernanke (1983) who interprets the spread as a monetary policy factor.) We use the yield difference between three-month nonfinancial corporate commercial paper rates and the three month Treasury yield. The commercial paper rates are measured weekly, as the averages over business days.

Some of the funds in our sample are global bond funds, so we include a factor for currency risks. Our measure is the value of the US dollar, relative to a trade-weighted average of major trading partners, from the FRED. This index is measured weekly, as the averages of daily figures, and we treat it the same way we treat the other weekly data. Corporate bond funds, and high-yield funds in particular, may be exposed to equity-related factors. We therefore include two equity market factors in our analysis. We measure equity volatility with the VIX-OEX index implied volatility. This series is available starting in January of 1986. We also include an equity market valuation factor, measured as the price/dividend ratio for the CRSP value-weighted index. The dividends are the sum of the dividends over the past twelve months, and the value is the cum-dividend value of the index. The level of this ratio is a state variable for valuation levels in the equity market, and its monthly first difference is used as a factor.

A.6 Sharpe Style Indexes

Following Sharpe (1992) we combine the asset class returns, R_i, using a set of portfolio weights, $\{w_i\}$, to minimize the "tracking error" between the return of the fund group, R_p, and the portfolio, $\Sigma_i w_i R_i$. The portfolio weights are required to sum to 1.0 and must be non-negative, which rules out short positions:

$$\text{Min}_{\{wi\}} \text{ Var}[R_p - \Sigma_i w_i R_i], \tag{A.4}$$

$$\text{subject to:} \quad \Sigma_i w_i = 1, \; w_i \geq 0 \text{ for all i,}$$

where Var[.] denotes the variance. We solve the problem numerically. The asset class returns include US Treasury bonds of three maturity ranges from CRSP (less than 12 months, less than 48 months and greater than 120 months), the Lehman Global bond index, the Lehman US Mortgage Backed Securities index, the Merrill Lynch High Yield US Master index and the Lehman US Credit Aaa bond index.

References

Admati, Anat, Sudipto Bhattacharya, Paul Pfleiderer and Stephen A. Ross, 1986, On timing and selectivity, Journal of Finance 61, 715-732.

Aragon, George, 2005, Timing multiple markets: Theory and evidence from balanced mutual funds, working paper, Arizona State.

Becker, C., W. Ferson, D. Myers and M. Schill, 1999, Conditional Market timing with Benchmark investors, *Journal of Financial Economics* 52, 119-148.

Bernanke, Benjamin, 1983, Nonmonetary effects of financial crisis in the propagation of the Great Depression, American Economic Review 73, 257-276.

Bessembinder, Hendrick, W. Maxwell and K. Venkataraman, 2006, Optimal market transparency: Evidence from the initiation of trade reporting in corporate bonds, Journal of Financial Economics (forthcoming).

Biais, Bruno and Fany Declerk, 2006, Liquidity and price discovery in the european corporate bond market, working paper, Toulouse University.

Blake, Christopher R., Edwin J. Elton and Martin J. Gruber, 1993, The performance of bond mutual funds, *Journal of Business* 66, 371-403.

Blume, Marshall, D. Keim and J. Patel, 1991, Returns and volatility of low-grade bonds: 1977:1989, Journal of Finance 46, 49-74.

Brown, Stephen, David R. Gallagher, Ono W. Steenbeck and Peter L. Swan, 2004, Double or nothing: Patterns of equity fund holdings and transactions, working paper, New York University.

Brown, David T. and William J. Marshall, 2001, Assessing fixed income manager style and performance from historical returns, *Journal of Fixed Income* 10 (March).

Chapman, David A., John B. Long and Neil D. Pearson, 2001, Using proxies for the short rate: When are three months like an instant? *Journal of Finance*

Chiang, I-Hsuan Ethan, 2006, Skewness and coskewness in bond returns, working paper, Boston College.

Comer, George, 2006, Evaluating Bond Fund Sector Timing Skill, working paper, Georgetown University.

Comer, George, Vaneesha Boney and Lynne Kelley, 2005, High quality bond funds: Market timing ability and performance, working paper, Georgetown University.

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

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

Dybvig, Phillip H. and Jonathan Ingersoll, 1982, Mean variance theory in complete markets, Journal of Business 55, 233-251.

Edelen, Roger, 1999, Investor flows and the assessed performance of open-ended mutual funds, Journal of Financial Economics 53, 439-466.

Edwards, A., L. Harris and M. Piwowar, 2006, Corporate bond market transparency and transactions costs, Journal of Finance (forthcoming).

Elton, Edwin, Martin J. Gruber and Christopher R. Blake, 1995, Fundamental Economic Variables, Expected returns and bond fund performance, *Journal of Finance* 50, 1229-1256.

Elton, Edwin, Martin J. Gruber and Christopher R. Blake, 2001, A first look at the Accuracy of the CRSP and Morningstar Mutual fund databases, *Journal of Finance* 56, 2415-2430.

Evans, Richard, 2006, Mutual fund incubation: The market for fund returns, working paper, Boston College.

Farnsworth, Heber K., Wayne Ferson, David Jackson and Steven Todd, Performance Evaluation with Stochastic Discount Factors, 2002, *Journal of Business* 75, 473-504.

Ferson, Wayne E., 1995, Theory and Empirical Testing of Asset Pricing Models, Chapter 5 in Finance, *Handbooks in Operations Research and Management Science*, by Jarrow, Maksimovic and Ziemba (editors), Elsevier, 145-200.

Ferson, W., Tyler Henry and Darren Kisgen, 2006, Evaluating Goverment Bond Funds using Stochastic Discount Factors, *Review of Financial Studies* 19, 423-455.

Ferson, W. and Kenneth Khang, 2002, Conditional performance measurement using portfolio weights: Evidence for pension funds, *Journal of Financial Economics* 65, 249-282.

Ferson, Wayne E. and Meijun Qian, 2004, Conditional Performance Evaluation Revisited, Research Foundation Monograph of the CFA Institute ISBN 0-943205-69-7, 84 pages.

Ferson, W. and Rudi Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425-462.

Ferson, Sarkissian and Simin, 2003, Spurious regressions in Financial Economics? *Journal of Finance* 58, 1393-1414.

Ferson, W., S. Sarkissian and T. Simin, 2008, Asset Pricing Models with Conditional Alphas and Betas: The Effects of Data Snooping and Spurious Regression, *Journal of Financial and Quantitative Analysis* 43, 331-354.

Ferson, W. and Vincent A. Warther, 1996, Evaluating Fund Performance in a Dynamic Market," *Financial Analysts Journal* 52, no. 6, pp.20-28.

Fisher, Lawrence, 1966, An algorithm for finding exact rates of return, Journal of Business 39, 111-118.

Gatev, Evan and Phillip Strahan, 2006, Banks' advantage in supplying liquidity: Theory and evidence from the commercial paper market, Journal of Finance 61, 867-892.

Getmansky, Mila, Andrew W. Lo and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, Journal of Financial Economics 74, 529-610.

Goetzmann, W., Ingersoll J., Ivkovic Z., 2000, Monthly measurement of daily timers, *Journal of Financial and Quantitative Analysis* 35, 257-290.

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

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

Hansen, Lars P., 1982, Large sample properties of the generalized method of moments estimators, Econometrica 50, 1029-1054.

Hotchkiss, E., M. Goldstein and E. Sirri, 2006, Transparency and liquidity: A controlled experiment on corporate bonds, Review of Financial Studies (forthcoming).

Jagannathan, R., Korajczyk, R., 1986. Assessing the market timing performance of managed portfolios. *Journal of Business* 59, 217-236.

Jensen, M. C., 1968, "The Performance of Mutual Funds in the period 1945-1964," *Journal of Finance*, 23, 389-416.

Jiang, George T., Tong Yao and Tong Yu, 2005, Do mutual funds time the market? Evidence from portfolio holdings, Journal of Financial Economics (forthcoming).

Jiang, Wei, 2003, A nonparametric test of market timing, *Journal of Empirical Finance* 10, 399-425.

Krau, Alan and Robert Litzenberger, 1976, Skewness preference and the valuation of risky assets, *Journal of Finance* 31, 1085-1100.

Lehman Brothers, 2006, A Guide to the Lehman Brothers Global Family of Indices, Lehman Brothers Fixed Income Research.

Litterman, R. and J. Sheinkman, 1991, Common factors affecting bond returns, Journal of Fixed Income 1,

Merton, Robert C. and Roy D. Henriksson. 1981. "On market timing and investment performance II: Statistical procedures for evaluating forecasting skills." *Journal of Business*, 54: 513-534.

Moneta, Fabio, 2008, Measuring bond mutual fund performance with portfolio characteristics, working paper, Boston College.

Morey, R. M. and E.S. O'Neal, 2005, Window dressing in bond mutual funds, Journal of Financial Research (forthcoming).

Qian, Meijun, 2006, Stale prices and the performance evaluation of mutual funds, working paper, Boston College.

Rubinstein, Mark, 1973, A comparative static analysis of risk premium, Journal of Business 19, 425-442.

Scholes, Myron and Joseph Williams, 1977, Estimating betas from nonsynchronous data, Journal of Finance 5, 309-327.

Sharpe, William. F., 1992, Asset Allocation: Management style and performance measurement, Journal of Portfolio Management 18, 7-19.

Treynor, Jack, and Kay Mazuy, 1966, "Can Mutual funds Outguess the Market?" *Harvard Business Review*, 44, 131-136.

Zitzewitz, Eric, 2003, Who cares about shareholders? Arbitrage-proofing mutual funds, *Journal of Law, Economics and Organizations* 19, 225-266.

Table 1

Mutual Fund Monthly Returns: Summary Statistics. The sample periods for the fund returns are January of the year indicated under Begin through March of 2007. Nobs is the number of nonmissing time series observations, Begno is the number of funds at the start of the sample period and Endno is the number in March of 2007. The returns are percent per month. Mean is the sample mean, Std is the sample standard deviation, ρ_1 is the first order sample autocorrelation and ρ_2 is the second order autocorrelation. Panel B presents the portfolio weights of Sharpe style benchmarks associated with each group of funds. The benchmarks are formed from the returns to US Treasury bonds with less that or equal to 12 (le12) 48 (le48) months to maturity, greater than 120 months to maturity (gt120), a high-grade corporate bond index (cb), a low-grade corporate bond index (junk), a global bond index (global) and a mortgage-backed securities index (mort). The weights are estimated using all available months from 1976 through 2007.

Panel A: Equally-weighted Portfolios of Mutual Funds

Style	Begin	nobs	Begno	Endno	Mean	Min	Max	Std	ρ_1	ρ_2
All	1962	543	4	1054	0.617	-5.397	9.770	1.511	0.253	-0.007
Index	1991	195	1	18	0.546	-3.609	3.907	1.122	0.138	-0.071
Global	1994	159	47	54	0.436	-3.483	4.557	1.455	0.162	-0.081
Short Term	1993	171	126	217	0.369	-0.856	2.267	0.455	0.300	0.149
Government	1986	255	1	127	0.547	-3.817	4.318	1.248	0.142	-0.120
Mortgage	1991	195	10	61	0.491	-2.411	3.188	0.854	0.204	0.011
Corporate	1962	543	4	308	0.627	-7.339	10.97	1.691	0.177	-0.026
High Yield	1991	195	34	135	0.737	-6.900	7.535	1.835	0.252	0.075
Other	1964	543	1	134	0.661	-5.103	9.609	1.640	0.221	0.001

Panel B: Sharpe Style Weights

Funds	Weights assigned to benchmarks:						
	le12	le48	gt120	cb	junk	global	mort
All	0.00	0.12	0.03	0.35	0.25	0.05	0.19
Index	0.07	0.20	0.21	0.26	0.03	0.02	0.20
Global	0.00	0.00	0.00	0.22	0.17	0.61	0.00
Short Term	0.54	0.19	0.00	0.14	0.02	0.03	0.08
Government	0.06	0.13	0.21	0.36	0.00	0.03	0.21
Mortgage	0.09	0.00	0.02	0.21	0.00	0.00	0.67
Corporat	0.00	0.15	0.07	0.55	0.17	0.02	0.04
High Yield	0.00	0.10	0.00	0.00	0.77	0.00	0.13
Other	0.00	0.00	0.00	0.29	0.39	0.06	0.27

Table 2

Summary Statistics for the bond market factor data. The sample periods begin as indicated under Starts (yyyymm), and all series end in December of 2007. Nobs is the number of time series observations, excluding missing values. The units are percent per year, except the US dollar (an index number) and Equity values (a price to dividend ratio). Mean is the sample mean, std is the sample standard deviation and ρ_1 is the first order sample autocorrelation of the series.

Panel A: Levels of the Factors

Factor	Starts	Nobs	mean	min	max	std	ρ_1
short rate	196201	540	5.847	0.904	16.38	2.842	0.981
term slope	196201	539	0.836	-3.160	3.310	1.115	0.955
curvature	196907	450	0.168	-1.097	0.773	0.286	0.834
credit spread	196201	551	0.989	0.320	2.820	0.418	0.961
mortgage spread	197104	393	2.184	-0.410	5.580	0.821	0.839
liquidity spread	198211	290	0.402	-0.151	2.179	0.345	0.658
US dollar	197101	443	110.3	84.48	167.7	14.77	0.986
Equity Values	196201	540	36.54	16.27	71.18	13.24	0.992
Equity Volatility	198601	252	19.77	10.63	61.41	6.940	0.798

Panel B: First differences of the Factors

short rate	196202	539	0.0041	-4.158	2.611	0.539	0.132
term slope	196202	537	-0.0016	-1.460	2.800	0.334	0.130
curvature	196908	449	-0.0004	-0.807	0.770	0.164	-0.292
credit spread	196202	550	0.0007	-0.550	0.680	0.116	0.063
mortgage spread	197105	392	0.0049	-2.110	3.240	0.461	-0.097
liquidity spread	198212	289	0.0004	-1.845	1.718	0.286	-0.399
US dollar	197102	442	-0.0652	-8.132	8.694	2.318	0.159
Equity Values	196202	539	0.0233	-6.630	6.508	1.615	-0.035
Equity Volatility	198602	251	-0.0319	-15.28	39.03	4.388	-0.193

Table 3

Regressions of bond funds and benchmark index returns on changes in factors and their squares. The sample starts in February of 1962 or later, depending on the factor and fund style, and ends in March of 2007. The t-ratios for the regression coefficients on the squared factor changes are shown when they exceed 1.6 in absolute value; othewise a zero is shown. These are based on regressions with a single factor and its square. R^2 is the adjusted coefficient of determination in a regression featuring the changes in all nine factors and their squared changes.

Style group	short	slope	curve	credit	mort.	liquid	US dollar	equity values	equity volatility	R^2(%)
Panel A: Style Benchmarks										
Index	0.000	1.93	0.000	1.88	0.000	0.000	0.000	0.000	2.20	65.1
Global	0.000	1.64	0.000	0.000	0.000	1.99	0.000	0.000	3.29	57.1
Short Term	1.69	2.69	0.000	2.51	0.000	1.66	0.000	0.000	2.94	60.0
Government	0.000	1.96	0.000	1.88	0.000	0.000	0.000	0.000	2.22	65.4
Mortgage	0.000	2.08	0.000	1.88	0.000	1.74	0.000	0.00	2.63	51.5
Corporate	0.000	1.78	0.000	1.86	0.000	1.83	0.000	0.000	0.000	58.4
High Yield	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-1.92	0.000	30.7
Other	0.000	1.68	0.000	1.71	0.000	2.51	0.000	0.000	0.000	41.9
Panel B: Funds by Style										
Index	-1.81	0.000	0.000	-1.90	0.000	0.000	2.40	0.000	0.000	65.8
Global	-2.08	0.000	0.000	0.000	0.000	-1.75	2.48	0.000	0.000	47.4
Short Term	0.000	0.000	0.000	-1.97	0.000	0.000	0.000	0.000	0.000	43.5
Government	0.000	0.000	0.000	-2.23	0.000	0.000	2.45	0.000	2.38	63.1
Mortgage	0.000	0.000	0.000	-2.03	0.000	0.000	1.63	0.000	0.000	54.5
Corporate	0.000	0.000	0.000	0.000	0.000	1.85	0.000	0.000	0.000	55.3
High Yield	0.000	0.000	0.000	3.31	0.000	0.000	2.44	-2.78	-3.58	42.2
Other	0.000	0.000	0.000	1.62	0.000	2.49	0.000	0.000	0.000	40.9

Table 4

The effects of controls for non-timing-related nonlinearity on fund groups' timing coefficient estimates. The table records the percentage of cases in which the t-ratio for the timing coefficients is above 2.0 (first line) or below -2.0 (second line). In panel A there are 180 possible cases for each fund style and in Panel B there are 160 possible cases for each factor. The cases are based on 20 equally-weighted portfolios of funds grouped within each style according to fund characterstics. The models are estimated on monthly data for 1962-2007, with 543 or fewer observations depending on the fund group and factor combination.

	No controls	quadratic benchmark	piecewise benchmark	stale prices	interim trading	public information	all controls
Panel A: By fund style groups							
Index	3.75	6.25	2.50	1.25	5.00	11.3	2.50
	7.50	0.000	0.625	0.000	13.8	20.6	0.000
Global	7.50	0.625	0.000	21.9	6.88	11.9	0.000
	11.9	2.50	1.25	16.3	14.4	20.6	0.000
Short Term	0.000	0.000	0.000	0.000	0.625	1.25	0.000
	10.0	3.13	2.50	6.88	13.8	16.3	0.000
Government	0.625	0.000	0.000	0.000	0.625	1.25	1.25
	13.8	0.000	0.000	0.000	18.1	21.9	0.000
Mortgage	0.000	0.000	0.625	0.000	0.000	0.000	1.25
	0.625	0.000	0.000	0.000	10.0	22.5	0.000
Corporate	0.000	0.000	1.88	6.25	8.13	0.625	11.9
	2.50	6.25	0.625	0.000	1.88	2.50	0.000
High Yield	21.9	1.25	2.50	22.5	12.5	23.1	0.000
	25.0	0.625	0.000	10.6	15.6	25.0	0.000
Other	0.625	0.000	6.88	8.75	13.8	8.13	8.13
	2.50	0.625	0.000	0.000	1.25	3.13	0.000

Table 4 (continued)

	No controls	quadratic benchmark	piecewise benchmark	stale prices	interim trading	public information	all controls
Panel B: By Factor							
Short Rate	0.000	4.38	0.000	0.000	0.625	0.000	0.000
	27.5	0.625	0.625	17.5	26.3	3.13	0.000
Term Slope	0.000	0.000	0.000	0.625	15.6	0.000	1.88
	0.000	0.625	0.000	5.63	0.000	0.000	0.000
Curvature	0.000	0.000	0.000	3.13	0.000	8.75	2.50
	0.000	0.000	0.000	0.000	0.000	64.4	0.000
Credit Spread	12.5	0.625	0.625	12.5	12.5	12.5	0.000
	11.9	0.000	0.625	0.000	38.8	33.1	0.000
Mortgage Spread	0.000	0.000	0.000	10.0	0.000	0.000	0.000
	0.000	0.000	0.000	0.000	0.625	0.000	0.000
US Dollar	21.3	1.25	1.88	29.4	13.1	35.6	0.000
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Equity Values	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	13.8	0.000	0.000	10.6	3.13	13.8	0.000
Eq. Volatility	0.625	1.88	11.9	5.00	5.63	0.625	20.6
	20.6	11.9	3.75	0.000	20.0	18.1	0.000

Table 5

Timing coefficients of individual bond funds. Panel A summarizes the fractions of funds with timing coefficients larger than (for the right tail and median) or smaller than (for the left tail) the critical values for the indicated fractiles from the null distribution in which the true timing coefficients are zero. Panel B presents the correlations of the timing coefficients with various fund characteristics. The monthly samples start in February of 1962 or later, depending on the factor and fund, and end by March of 2007.

null fractile	short	slope	curve	credit	mortgage	liquidity	dollar	equity value	equity volatility
Panel A: The distributions of the timing coefficients									
0.975	**0.0182**	**0.0086**	**0.0390**	**0.0286**	**0.0597**	**0.0510**	**0.0273**	**0.0296**	**0.0235**
0.950	**0.0316**	**0.0133**	**0.0640**	**0.0715**	**0.0989**	**0.0926**	**0.0640**	**0.0428**	**0.0839**
0.900	**0.0893**	**0.0250**	**0.126**	**0.153**	**0.151**	**0.168**	**0.235**	**0.0763**	**0.252**
0.500	**0.547**	**0.395**	**0.278**	**0.742**	**0.684**	**0.732**	**0.824**	**0.765**	**0.706**
0.100	**0.211**	**0.178**	**0.174**	**0.0469**	**0.0526**	**0.0322**	**0.0406**	**0.0530**	**0.0729**
0.0500	**0.153**	**0.100**	**0.0476**	**0.0167**	**0.0303**	**0.0157**	**0.0257**	**0.0241**	**0.0455**
0.0250	**0.0996**	**0.0469**	**0.0187**	**0.0127**	**0.0214**	**0.0126**	**0.0179**	**0.0078**	**0.0298**
# cases	1265	1278	1282	1259	1122	1274	1282	1284	1275
Panel B: Correlations of timing coefficients with fund characteristics									
expenses	**-0.1520**	**-0.2100**	**-0.0102**	**0.0407**	**-0.1230**	**-0.0130**	**-0.0491**	**0.0570**	**0.0202**
turnover	**0.0388**	**0.0587**	**-0.0105**	**-0.0087**	**-0.0298**	**-0.0242**	**-0.0213**	**-0.0124**	**-0.0010**
flow	**0.0052**	**0.0138**	**0.0005**	**-0.0220**	**0.0916**	**-0.0978**	**-0.0478**	**0.0158**	**-0.0176**
cash	**-0.0587**	**-0.0135**	**0.0339**	**-0.0372**	**-0.0365**	**-0.0301**	**-0.1460**	**0.0542**	**-0.0332**
net assets	**-0.0268**	**0.0462**	**-0.0142**	**-0.0036**	**-0.0017**	**0.0164**	**0.0041**	**-0.0068**	**0.0179**
options	**-0.0497**	**-0.0121**	**-0.0096**	**-0.0042**	**-0.0157**	**0.0274**	**0.0484**	**-0.0044**	**0.0358**
yield	**-0.1430**	**-0.0547**	**-0.0590**	**0.0438**	**-0.1660**	**0.0918**	**0.1540**	**-0.0240**	**-0.0276**
age	**0.0329**	**-0.1860**	**-0.0613**	**0.0669**	**-0.0724**	**-0.0146**	**0.0053**	**-0.0056**	**-0.1050**
lag return	**-0.1570**	**-0.3020**	**-0.1270**	**0.1030**	**0.1340**	**-0.0536**	**-0.0925**	**0.8830**	**0.2430**
maturity	**0.0301**	**-0.0146**	**-0.0794**	**0.0579**	**0.0440**	**-0.0425**	**0.0758**	**0.0018**	**0.0597**
load	**-0.1670**	**-0.0982**	**-0.0200**	**0.0359**	**-0.0577**	**0.0529**	**0.0115**	**0.0249**	**0.0296**
Index	**-0.0021**	**0.0688**	**-0.0091**	**-0.0126**	**0.0147**	**-0.0443**	**0.0247**	**-0.0108**	**0.0053**
Global	**-0.0957**	**-0.1510**	**-0.0409**	**-0.0316**	**0.1650**	**-0.0358**	**-0.158**	**0.0077**	**0.0602**
Short Term	**.0647**	**0.0829**	**-0.0093**	**-0.0150**	**-0.0344**	**0.0470**	**-0.0938**	**-0.0184**	**-0.0956**
Government	**0.0286**	**0.1120**	**-0.0274**	**-0.0163**	**0.0558**	**-0.0343**	**-0.0280**	**-0.0118**	**0.0040**
Mortgage	**0.0493**	**0.0877**	**-0.0020**	**-0.0131**	**-0.0133**	**0.0182**	**-0.0054**	**-0.0063**	**-0.0451**
Corporate	**-0.0221**	**0.1690**	**-0.0008**	**-0.0045**	**0.0216**	**-0.0919**	**-0.0171**	**-0.0285**	**0.0070**
High Yield	**-0.2010**	**0.0004**	**0.0298**	**0.0003**	**-0.0285**	**0.0586**	**0.1800**	**0.0230**	**0.0801**
Other	**-0.0852**	**0.0007**	**-0.0366**	**-0.0076**	**-0.0213**	**0.1960**	**0.1320**	**-0.0032**	**0.2220**

Table 6

Timing adjusted performance of individual bond funds. The first three columns of Panel A summarize the mean excess return over a short term treasury, the mean return over a style benchmark, and the values of the estimated alphas under the null hypothesis that the true alphas are zero, taken at various fractiles of the distribution across funds. The average excess returns are percent per month. The remaining columns report the fractions of funds with alphas larger than (for the right tail and median) or smaller than (for the left tail) the critical values for the indicated fractiles from the null distribution in which the true alphas are zero. Panel B summarizes the average risk premiums for the squared factors (denoted convexity premium), which are the average excess returns of mimicking portfolios for the squared factors. It also summarizes the distributions of loadings on the style benchmarks and on the hedge portfolios for the squared factors. The monthly samples start in February of 1962 or later, depending on the factor and fund, and end by March of 2007.

fractile	mean mean return	net of bench	critical alpha	short	slope	curve	credit	mort.	liq.	dollar	equity value	equity vol.
Panel A: The distribution of Timing Adjusted Alphas												
0.975	**0.818**	**0.171**	**0.150**	**0.023**	**0.031**	**0.024**	**0.018**	**0.020**	**0.010**	**0.021**	**0.023**	**0.017**
0.950	**0.715**	**0.104**	**0.099**	**0.053**	**0.051**	**0.044**	**0.029**	**0.036**	**0.014**	**0.048**	**0.046**	**0.046**
0.900	**0.632**	**0.036**	**0.061**	**0.079**	**0.071**	**0.074**	**0.051**	**0.066**	**0.030**	**0.069**	**0.071**	**0.068**
0.500	**0.463**	**-0.067**	**-0.000**	**0.217**	**0.216**	**0.205**	**0.192**	**0.198**	**0.143**	**0.214**	**0.217**	**0.220**
0.100	**0.285**	**-0.199**	**-0.064**		**0.419**	**0.428**	**0.384**	**0.448**	**0.447**	**0.532**	**0.396**	**0.408**
	0.403											
0.050	**0.193**	**-0.282**	**-0.108**	**0.268**	**0.242**	**0.205**	**0.248**	**0.260**	**0.341**	**0.232**	**0.186**	**0.222**
0.025	**0.100**	**-0.386**	**-0.156**	**0.138**	**0.130**	**0.108**	**0.181**	**0.153**	**0.188**	**0.139**	**0.109**	**0.134**
# cases				1329	1368	1375	1251	1179	697	1357	1376	1376
Panel B: Convexity premiums and loadings												
Fund Loadings on Benchmarks:												
Upper 10%				1.38	1.40	1.35	1.42	1.38	1.40	1.38	1.38	1.33
Median				0.91	0.91	0.91	0.89	0.89	0.91	0.91	0.92	0.90
Lower 10%				0.42	0.43	0.45	0.48	0.43	0.46	0.45	0.42	0.43
Average Convexity premiums				0.415	0.231	-0.332	0.107	0.144	-0.029	-0.244	3.59	-2.93
Fund Convexity Loadings:												
Upper 10%				0.134	0.145	0.054	0.130	0.226	0.103	0.064	0.091	0.093
Median				0.005	0.012	0.001	0.003	0.006	-0.001	-0.0001	-0.006	0.0001
Lower 10%				-0.224	-0.128	-0.065	-0.163	-0.209	-0.129	-0.0775	-0.118	0.072
Correlations of Hedge Portfolios: with Squared factor changes				-0.278	-0.383	-0.308	-0.215	-0.456	0.160	0.510	-0.273	0.532

Table 7

Timing adjusted performance of individual bond funds gross of transactions costs. Transactions costs are estimated as the average expense ratio of each fund plus an assumed round trip trading cost associated with the fund style, multiplied by the average reported turnover. These costs are added back to the fund return. The first three columns of Panel A summarize the mean excess return over a short term treasury, the mean return over a style benchmark, and the values of the estimated alphas under the null hypothesis that the true alphas are zero, taken at various fractiles of the distribution across funds. The average excess returns are percent per month. The remaining columns report the fractions of funds with alphas larger than (for the right tail and median) or smaller than (for the left tail) the critical values for the indicated fractiles from the null distribution in which the true alphas are zero. The monthly samples start in February of 1962 or later, depending on the factor and fund, and end by March of 2007.

fractile	mean return	mean net of bench	critical alpha	short	slope	curve	credit	mort.	liq.	dollar	equity value	equity vol.
The distribution of before cost Timing Adjusted Alphas												
0.975	0.952	0.318	0.149	0.158	0.113	0.144	0.101	0.110	0.141	0.161	0.132	0.155
0.950	0.855	0.239	0.104	0.257	0.198	0.239	0.184	0.197	0.228	0.258	0.251	0.256
0.900	0.768	0.176	0.063	0.430	0.390	0.377	0.360	0.335	0.412	0.409	0.407	0.425
0.500	0.572	0.032	-0.000	0.767	0.754	0.763	0.746	0.730	0.761	0.742	0.753	0.768
0.100	0.389	-0.081	-0.067	0.091	0.091	0.088	0.093	0.097	0.090	0.100	0.094	0.093
0.050	0.303	-0.145	-0.114	0.046	0.059	0.056	0.066	0.060	0.052	0.061	0.055	0.058
0.025	0.241	-0.225	-0.169	0.031	0.042	0.028	0.043	0.044	0.033	0.040	0.034	0.037
# cases				1329	1367	1374	1251	1178	1375	1357	1375	1375