Available online at www.sciencedirect.com
ScienceDirect
Journal of Financial Economics 86 (2007) 724–758
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

Do mutual funds time the market? Evidence from portfolio holdings ☆

George J. Jiang[a,*], Tong Yao[a], Tong Yu[b]

[a]Eller College of Management, University of Arizona, Tucson, AZ 85721, USA
[b]College of Business Administration, University of Rhode Island, Kingston, RI 02881, USA

Received 12 May 2005; received in revised form 22 August 2006; accepted 26 September 2006
Available online 7 July 2007

Abstract

Previous research finds insignificant market-timing ability for mutual funds using tests based on fund returns. The return-based tests, however, are subject to the "artificial timing" bias. In this paper, we propose and implement new measures of market timing based on mutual fund holdings. Our holdings-based measures do not suffer from the artificial timing bias. We find that, on average, actively managed U.S. domestic equity funds have positive timing ability. Market timing funds use non-public information to predict market returns, tend to have high industry concentration, large fund size, a tilt toward small-cap stocks, and are active in industry rotation.

JEL classification: G10; G11

Keywords: Mutual funds; Market timing; Portfolio holdings

1. Introduction

Two strands of the finance literature stand in curious contrast to each other. On the one hand, many studies argue that aggregate stock market returns are predictable, and that such predictability should have a significant impact on investors' optimal asset

☆We would like to thank Tyler Brough, Jeff Busse (AFA discussant), Yong Chen, Wayne Ferson, Kathy Kahle, Chris Lamoureux, Bill Schwert (the editor), an anonymous referee, and seminar participants at University of Arizona and the 2005 AFA meeting for helpful comments and suggestions.
*Corresponding author. Tel.: +1 520 621 337; fax: +1 520 621 1261.
E-mail address: gjiang@eller.arizona.edu (G.J. Jiang).

0304-405X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2006.09.006

allocation.[1] On the other hand, there is scant evidence that investors actually take advantage of such predictability in their portfolio decisions. In particular, earlier studies such as Treynor and Mazuy (1966) and Henriksson and Merton (1981) as well as more recent studies such as Becker, Ferson, Myers, and Schill (1999) and Jiang (2003) find that the average market timing performance of mutual funds is insignificant and sometimes even negative. Fund managers are sophisticated and informed investors. If they cannot exploit market return predictability, it is unlikely that anyone else could.[2]

Existing market timing measures, such as those of Henriksson and Merton (1981) and Treynor and Mazuy (1966), are based on nonlinear regressions of realized fund returns against contemporaneous market returns (hereafter "return-based measures"). As the literature shows, however, a nonlinear relation between fund and market returns can also be induced by factors other than active market timing. Jagannathan and Korajczyk (1986) point out that certain dynamic trading strategies by mutual funds may give rise to option-like features in fund returns. Thus, nonlinear relations between fund and market returns may be due to a "dynamic trading" effect. Jagannathan and Korajczyk (1986) show further that because returns of certain stocks have option-like features, the returns of a passive portfolio investing in these stocks may also have a convex or concave relation with market returns. This is known as the "passive timing" effect. Both effects are often referred to as "artificial timing" in the literature. Moreover, Goetzmann, Ingersoll, and Ivkovich (2000) show that the return-based measures are biased downwards when funds engage in active timing and trade between the observation dates of fund returns. For instance, when funds engage in daily market timing, the return-based timing measures using monthly fund returns tend to underestimate timing ability.

To overcome these problems, we propose alternative market timing measures based on observed mutual fund portfolio holdings (hereafter "holdings-based measures"). Specifically, we estimate a fund's beta as the weighted average of the betas of the individual stocks held in the portfolio, and directly test whether the covariance between fund betas at the beginning of a holding period and the holding period market returns is significant. In contrast to the return-based measures that rely on ex post realized returns to estimate beta shifting, the holdings-based measures use only ex ante information on portfolio holdings. Therefore, the holdings-based measures do not suffer from any bias induced by subsequent trading activities during a holding period or the dynamic trading effect. In addition, we show that the holdings-based measures are also robust to the passive timing effect. When performing market timing tests on passive portfolios formed on stock characteristics, i.e., size, book-to-market, and momentum, we find that the holdings-based timing measures are generally small and insignificant, whereas the return-based timing measures are significantly negative.

[1]For evidence of market return predictability, see, for example, Campbell (1987), Campbell and Shiller (1988a,b), Cochrane (1991), Fama and French (1987, 1988, 1989), Fama and Schwert (1977), Ferson (1989), Keim and Stambaugh (1986), Lamont (1998), Lewellen (1999), and Pontiff and Schall (1998). For discussions on how such predictability should affect investors' asset allocation decisions, see, for example, Balduzzi and Lynch (1999), Barberis (2000), Campbell, Chan, and Viceira (2003), Campbell and Viceira (1999), and Kandel and Stambaugh (1996).

[2]In addition, Blake, Lehmann, and Timmermann (1999), Coggin, Fabozzi, and Rahman (1993), Fung, Xu, and Yau (2002), and Graham and Harvey (1996) find no evidence of positive market timing ability by pension funds, hedge funds, or investment newsletters.

The holdings-based measures have the further benefit of improved statistical power. Fund returns are volatile with observations often available only at low frequencies (e.g., monthly) and over a relatively short time period. This limits the statistical power of the return-based tests. For the holdings-based measures, fund betas are calculated from stock betas and portfolio weights. When estimating stock betas, we can take advantage of stock return observations at a higher frequency (e.g., daily) and over extended time periods. As a result, even though individual stock betas are estimated with error, fund betas, as weighted averages of a large number of stock betas, are of much higher accuracy. Our simulations confirm that the holdings-based tests have better statistical power than the return-based tests, even when fund holdings are observed less frequently than fund returns.

The use of higher frequency data to enhance statistical power is similar to Bollen and Busse (2001), who perform market timing tests using daily fund returns and find positive timing ability for a sample of 230 domestic equity funds. However, the timing performance documented by Bollen and Busse (2001) is at the daily horizon, whereas in this study we focus on market timing activities at relatively long time horizons that are comparable to those in the market return predictability literature. In addition, the mutual fund portfolio holdings data we examine cover a large sample of actively managed equity funds. Reliable large-sample data on daily fund returns are not yet available (see discussions in Busse and Irvine, 2006).

Mutual fund portfolio holdings are used in a number of existing studies to evaluate fund performance, notably, Grinblatt and Titman (1989, 1993), Grinblatt, Titman, and Wermers (1995), Daniel, Grinblatt, Titman, and Wermers (1997), Wermers (1999, 2000, 2004), and Ferson and Khang (2002). These studies show that performance measures based on portfolio holdings are more powerful in detecting mutual fund stock selection ability. In addition, Daniel, Grinblatt, Titman, and Wermers (1997) and Wermers (2000) examine whether mutual funds exploit time-varying expected returns in characteristic benchmarks based on size, book-to-market, and momentum. Several studies also use portfolio allocation between cash and equity components to measure market timing and find no significant timing performance; see, e.g., Kryzanowski, Lalancette, and To (1996) and Becker, Ferson, Myers, and Schill (1999). As Warther (1995) and Edelen (1999) point out, the cash position of a fund is often affected exogenously by factors such as fund flow shocks, which tend to confound the effect of market timing.

We apply both return-based tests and holdings-based tests to a sample of 2,294 actively managed U.S. equity mutual funds over the period from 1980 to 2002. The fund returns data come from CRSP, and the fund portfolio holdings data come from Thomson Financial. Statistical inference is based on a bootstrapping procedure similar to that proposed by Kosowski, Timmermann, White, and Wermers (2005), which explicitly takes into account the cross-fund correlation and the finite-sample properties of timing measures. Consistent with the existing literature, the return-based timing measures indicate that on average mutual funds have slightly negative but statistically insignificant timing performance. In contrast, the holdings-based measures suggest that on average mutual funds have significantly positive timing ability. Specifically, at one-month and 12-month forecasting horizons, the average timing measures of the holdings-based tests are positive but not significant. At three-month and six-month forecasting horizons, the average timing measures are significantly positive. In addition, the proportion of funds with strongly positive timing measures is also higher than what one would expect from a sample of funds without timing ability. These results are robust to different methods used to estimate stock and fund betas, as well as variations in the testing procedure.

Time variations in fund betas can be driven by both active fund trading activities and passive portfolio weight changes. To examine the active market timing of mutual funds, we also perform the holdings-based tests using fund beta active changes. The tests provide similar results as those using fund beta levels, suggesting that mutual funds achieve market timing through active trading. The positive timing ability of mutual funds is also of economic significance. Using a contingent claim approach, we find that for the median fund, the economic value of active market timing can be as much as 0.60% per year. This result is economically significant in comparison with the documented contribution to fund performance by stock selection. For example, Wermers (2000) reports that mutual fund stock selection activities add to fund performance, on average, by 1.01% per year before transaction costs. This suggests that market timing can be an important investment strategy for mutual funds.

We identify links between several fund characteristics and market timing performance. Market timing funds tend to be those with high industry concentration in their portfolios and, to a lesser extent, those with large fund size and a tilt toward small-cap stocks. Further, fund managers adjust fund betas in response to macroeconomic variables such as aggregate dividend yield and the earnings-to-price ratio in a way that is consistent with how these variables predict market returns. However, average market timing performance remains significantly positive even after controlling for these macroeconomic variables. This evidence suggests that fund managers use not only publicly available macroeconomic information but also private or finer information to time the market.

Finally, we document that industry allocation is an important investment strategy employed by mutual funds to achieve market timing. We find a stronger market timing effect from industry shifting than from intra-industry portfolio allocations. Furthermore, mutual funds shift portfolio allocations to certain industries in response to macroeconomic conditions that are predictive of future market movement.

The remainder of the paper is organized as follows. Section 2 specifies both the return-based and holdings-based market timing tests and compares the two approaches in terms of artificial timing bias and statistical power. Section 3 describes the data and methodology used in our analysis. In Section 4, we examine mutual fund market timing ability by performing both the return-based and holdings-based tests. Section 5 conducts further analysis of mutual fund market timing activities. Section 6 concludes. Finally, details on the artificial timing of passive portfolios, statistical power of market timing tests, as well as the bootstrapping method are included in appendices.

2. Measures of market timing

2.1. Return-based measures

Let r_t and r_{mt} denote the excess fund and market returns over holding period t. The Treynor-Mazuy market timing measure (see Treynor and Mazuy (1966)) is the estimated coefficient γ from the regression

$$r_t = \alpha + \beta_0 r_{mt} + \gamma r_{mt}^2 + e_t \tag{1}$$

and the Henriksson-Merton measure (see Henriksson and Merton, 1981) is the estimated coefficient γ from

$$r_t = \alpha + \beta_0 r_{mt} + \gamma \max(r_{mt}, 0) + e_t, \tag{2}$$

where $\max(r_{mt}, 0)$ is the positive part of the market excess return. Various forms of the Treynor-Mazuy and Henriksson-Merton measures have been used in the literature to test mutual fund timing ability; see, for example, Treynor and Mazuy (1966), Kon (1983), Chang and Lewellen (1984), Henriksson (1984), Cumby and Glen (1990), Ferson and Schadt (1996), Kryzanowski, Lalancette, and To (1996), Becker, Ferson, Myers, and Schill (1999), and Jiang (2003). Most of the studies document negative but insignificant timing performance by mutual funds.

Note that both measures in (1) and (2) are based on the covariance between the time-varying fund beta β_t and the market return r_{mt},[3] that is,

$$\beta_t = \beta_0 + \gamma r_{mt} + \eta_t \quad \text{for the Treynor-Mazuy measure,} \tag{3}$$

$$\beta_t = \beta_0 + \gamma I_{r_{mt}>0} + \eta_t \quad \text{for the Henriksson-Merton measure,} \tag{4}$$

where $I_{r_{mt}>0}$ is an indicator that takes the value of one when $r_{mt}>0$ and zero otherwise. A significantly positive γ indicates positive timing ability by mutual funds.

2.2. Holdings-based measures

The return-based measures are useful when fund betas are not directly observed. However, when the portfolio holdings of a fund are observed, we can estimate the fund beta and perform tests based directly on (3) and (4). In this study, we estimate the fund beta as the weighted average of the beta estimates for stocks (and other securities) held by the fund, that is,

$$\hat{\beta}_t = \sum_{i=1}^{N} \omega_{it} \hat{b}_{it}, \tag{5}$$

where ω_{it} is the portfolio weight for stock i at the beginning of holding period $t+1$, and $\hat{b}_{it}$ is the beta for stock i estimated using data prior to period $t+1$. Various methods of estimating stock betas are detailed later in the paper.

In the spirit of (3) and (4), we measure market timing by estimating the coefficient γ directly from the regressions

$$\hat{\beta}_t = \alpha + \gamma r_{m,t+1} + \eta_{t+1}, \tag{6}$$

$$\hat{\beta}_t = \alpha + \gamma I_{r_{m,t+1}>0} + \eta_{t+1}, \tag{7}$$

where $\hat{\beta}_t$ is the fund beta estimated at the beginning of period $t+1$. The γ coefficients estimated from (6) and (7) are referred to as the holdings-based Treynor-Mazuy measure and the holdings-based Henriksson-Merton measure, respectively.

2.3. Comparison of market timing measures

In this section, we compare the potential bias and the statistical power of the return-based and holding-based market timing measures. We show that compared to the

[3]To see this, assume that the fund return follows the market model, $r_t = \alpha + \beta_t r_{mt} + \varepsilon_t$. Combining this with (3) and (4), we obtain (1) and (2), respectively, with $e_t = \eta_t r_{mt} + \varepsilon_t$.

return-based measures, the holdings-based measures do not suffer from artificial timing bias and they have better statistical power.

2.3.1. Artificial timing bias

As mentioned earlier, in addition to active market timing there are at least two other factors that can induce a convex or concave relation between fund and market returns. One is the dynamic trading effect, which arises due to certain dynamic trading strategies implemented by mutual funds. The other is the passive timing effect, which refers to the fact that the returns of a passive portfolio can also be nonlinearly related to market returns. Both effects, as illustrated below, can introduce artificial timing bias in the return-based tests.

To illustrate the effect of dynamic trading, we consider a simple case in which a fund trades in each period but returns are observed every two periods. Suppose a fund manager has no active timing ability but adjusts fund beta during the second period conditional on the realized market return in the first period. This leads to a correlation between the fund beta in the second period and the market return over the first period, which induces a seemingly contemporaneous nonlinear relation between the realized two-period fund and market returns. For example, a "positive-feedback" manager who increases market exposure after a market run-up would exhibit positive artificial timing, and a contrarian manager who reduces market exposure after a market run-up would exhibit negative artificial timing. That is, some commonly used trading strategies that are seemingly innocuous can generate nonlinear relations between fund returns and market returns.

The dynamic trading effect described above is also known in the literature as "interim trading," as it is caused by fund trading activities between the return observation dates. Note that interim trading can also cause return-based tests to underestimate true market timing ability. This is the case analyzed by Goetzmann, Ingersoll, and Ivkovich (2000). Using simulations, they show that when funds engage in daily market timing, the Henriksson-Merton measure based on monthly fund returns is biased downwards and has low power.

The holdings-based tests are robust to the dynamic trading effect. In contrast to the return-based tests, which measure the contemporaneous relation between realized fund returns and market returns, the holdings-based tests use only ex ante information on portfolio holdings. Specifically, the fund betas in (5) are computed using portfolio weights observed at the beginning of a holding period. The holdings-based measures are thus not affected by subsequent fund trading activities during the holding period. In addition, because of the use of beginning-of-period fund betas, the holdings-based tests remain unbiased in the presence of the high-frequency market timing activities described in Goetzmann, Ingersoll, and Ivkovich (2000). This, of course, is less a concern for our empirical analysis as we focus on mutual fund timing activities at relatively long horizons from one month to one year.

The passive timing effect is documented in existing studies for the return-based tests. For instance, Jagannathan and Korajczyk (1986) show that returns of certain stocks exhibit option-like features relative to market returns. As a result, returns of a passive portfolio investing in these stocks may have a convex or concave relation with market returns even when funds are not market timers. In contrast, in the holdings-based tests stock betas are estimated using data prior to the holding period, and thus are not affected by any contemporaneous nonlinear relation between stock and market returns.

To quantify the passive timing effect, we perform both return-based and holdings-based tests on passive portfolios. These portfolios are formed on the stock characteristics size, book-to-market, and momentum, and are similar to those in Daniel, Grinblatt, Titman, and Wermers (1997). The portfolio construction details are given in Appendix A. The results in Appendix A show that the passive portfolios' mean and median return-based timing measures are significantly negative. The results are consistent with Ferson and Schadt (1996), who find negative timing for a simple buy-and-hold strategy using return-based measures. In contrast, there is no significant artificial timing bias for the holdings-based tests. The mean and median of the holdings-based measures across passive portfolios are insignificantly different from zero.

To further illustrate the pattern and magnitude of the passive timing effect with respect to portfolio characteristics, we plot the return-based and holdings-based timing measures of the passive portfolios in Figs. A1 and A2. The plots show that negative biases in the return-based measures are highly related to size and are most pronounced for small-cap stocks. This pattern is consistent with the results in Jagannathan and Korajczyk (1986), who find that equal-weighted market returns exhibit a concave relation with value-weighted market returns. Stocks with a higher book-to-market ratio (value stocks) and higher momentum (past winners) also tend to exhibit more negative passive timing effect. In contrast, the holdings-based timing measures on passive portfolios are generally small in magnitude, with no systematic biases. Nevertheless, in our empirical analysis we explicitly control for the passive effect in the holdings-based tests. All individual stock beta estimates are adjusted by subtracting the average beta of the matching passive characteristics portfolio. This procedure has the added benefit of reducing cross-fund heteroskedasticity in market timing measures.

2.3.2. Statistical power

The difference in statistical power between the return-based and holdings-based tests results from different information used in each of the tests. In practice, data on fund returns are often available only at a low frequency (such as monthly or quarterly) over a limited period (a fund's life), and thus are not sufficient for the precise estimation of market timing measures. The advantage of the holdings-based measure is that with portfolio holdings information, fund betas can be estimated directly from stock betas. When estimating stock betas, we are no longer constrained by the length of a fund's life and we can take advantage of the fact that stock return data are available at a higher sampling frequency (e.g., daily). It is well documented in the literature that beta, as a variance–covariance measure, can be better estimated from data sampled at a higher frequency; see, for example, Merton (1980) and French, Schwert, and Stambaugh (1987).

To verify the advantage of the holdings-based measures, we simulate returns and portfolio holdings for mutual funds with various levels of market timing ability (γ), and perform both return-based and holdings-based Treynor-Mazuy tests on the simulated data. In the simulation, we take into account two realistic features in the data. First, the fund return data are available at a monthly frequency, while the portfolio holdings data are available at a quarterly frequency. Second, the portfolio holdings data are available for a shorter time period than the fund return data. Details of the simulation and results are provided in Appendix B. The results show that the standard errors of the holdings-based measures are several times smaller than those of the return-based measures. The simulations confirm that even when fund holdings are observed less frequently than fund returns, the holdings-based tests have better statistical power than their return-based counterparts.

3. Data and methodology

3.1. Data

We combine two mutual fund data sets in our analysis. The first is the CRSP mutual fund data set (hereafter the CRSP data set), which has information on monthly returns and fund characteristics such as total net assets, the expense ratio, loads, and the turnover ratio for all U.S.-based mutual funds. The second is the fund portfolio holdings data set from Thomson Financial (hereafter the Thomson data set). Its predecessor is the CDA/Spectrum data base used in a number of empirical studies, such as Grinblatt and Titman (1989, 1993), Daniel, Grinblatt, Titman, and Wermers (1997), and Wermers (2000, 2004). The Thomson data set contains quarterly or semiannual information on portfolio holdings for equity mutual funds investing in the U.S. market.[4] Most mutual funds in the Thomson data set are U.S.-based. There are also a number of foreign-based (mostly Canadian) funds. As Daniel, Grinblatt, Titman, and Wermers (1997) note, CDA/Spectrum does not have a survivorship problem.

We manually match the funds in the two data sets by fund names and ticker symbols. The matching procedure is similar to that in Wermers (2000). Since our focus is on actively managed U.S. domestic equity funds, we only include U.S. funds with one of the following four investment objectives defined in the Thomson data set: (a) aggressive growth, (b) growth, (c) growth and income, and (d) balanced. The investment objective codes in the Thomson data sometimes contain errors. We manually remove from our sample the index funds, foreign-based funds, U.S.-based international funds, fixed income funds, real estate funds, precious metal funds, and variable annuities that are misclassified as active U.S. domestic equity funds.

When a fund has multiple share classes, returns of these share classes are separately reported in the CRSP data set but holdings in the Thomson data set are for the entire fund. We calculate the returns of a multi-class fund as the weighted-average returns across share classes, using total net assets as the weight. Fund expense ratios, turnover ratios, and loads are similarly calculated. To ensure robust statistical inference, we require that a fund have a minimum of eight quarters of holdings data and 24 monthly return observations to be included in our sample.

The final matched data set has 2,294 unique funds over the period from 1980 to 2002 with information on both portfolio holdings and returns.[5] Table 1 reports summary

[4]The Thomson data set is based on mandatory and voluntary fund holdings disclosures. Prior to 2004, mutual funds were required to disclose their holdings semiannually; many funds voluntarily disclosed their holdings quarterly. The SEC increased the mandatory disclosure frequency from semiannual to quarterly effective May 2004. Such a change could further improve the power of holdings-based mutual fund performance measures with more observations on fund holdings.

[5]Various existing studies also match funds from CRSP and Thomson/CDA. Wermers (2000) combines the CDA data with the CRSP data set over the period between 1975 and 1994. His sample contains 1,788 equity funds. Kacperczyk, Sialm, and Zheng (2005) match the CRSP data set with the Thomson data set over the period between 1984 and 1999 and come up with 1,971 unique equity funds. Cohen, Coval, and Pastor (2005) also match the Thomson data set with the CRSP data set over the period between 1980 and 2002Q2. They report 235 matched

Table 1
Summary statistics of fund characteristics

This table reports the characteristics of mutual funds in the sample with a breakdown according to investment objectives. Total net assets, annual return, turnover, annual load, expense ratio, percentage investment in stocks, and average number of stocks held are averaged first over time for each fund, and then across funds. A fund's annual load is calculated as the total load divided by seven. The median number of stocks held is the cross-sectional median of the time-series averages. Age is the average life of funds in the sample, where fund life is defined as the time between the first and last reported monthly returns. "NOBS—return" denotes the average number of monthly return observations across funds, and "NOBS—holdings" denotes the cross-sectional average number of quarters with portfolio holdings information.

	All funds	Aggressive growth	Growth	Growth and income	Balanced
Number of funds	2294	255	1390	441	208
Total net asset ($ millions)	482.60	545.95	397.52	773.93	360.96
Annual return (%)	9.96	11.76	9.97	9.71	8.18
Turnover (per year)	0.94	1.13	0.98	0.73	0.94
Annual load (%/year)	0.30	0.39	0.27	0.35	0.30
Expense ratio (%/year)	1.31	1.43	1.34	1.19	1.16
Investment in stocks (%)	85.89	91.77	91.81	85.06	53.49
Age (years)	13.46	17.21	11.86	16.15	13.46
Average no. of stocks held	102.15	91.57	104.70	90.94	122.39
Median no. of stocks held	61.90	66.38	61.95	59.65	63.36
NOBS—return	131.96	165.65	119.90	150.39	132.91
NOBS—holdings	27.24	36.93	24.52	30.75	26.30

statistics of our mutual fund sample during the period from 1980 to 2002. Of the 2,294 funds, more than half (1,390) are growth funds, 255 are aggressive growth funds, 411 are growth and income funds, and 208 are balanced funds. By averaging first over the time series for each fund and then across funds, we obtain the following characteristics: the average total net assets (TNA) of the funds in our sample is $482.60 million, with an annual return of 9.96%, an annual turnover ratio of 0.94, an annualized load of 0.30%, and an annual expense ratio of 1.31%. Following Sirri and Tufano (1998), the annualized load is the total load divided by seven. The funds invest 85.89% of their assets in common stocks. In addition, the average fund age, calculated as the time between the first and last monthly return observations in the CRSP data set, is 13.46 years. The mean number of stocks held in a fund is 102.15, while the median number is 61.90. For a typical fund, we have 131.96 months of return data and 27.24 quarters of portfolio holdings data during the sample period.

In our empirical test, the CRSP value-weighted index return is used as a proxy for the market return. Individual stock returns and one-month T-bill yields (our proxy of the risk-free rate) also come from CRSP. In addition, we obtain observations of several economic variables that are documented in the literature as predictive of market returns, namely, the term spread, credit spread, aggregate dividend yield on the S&P 500 index, and the

(*footnote continued*)
funds at the end of 1980 and 1,526 matched funds in 2002Q2. It appears that our mutual fund sample is at least as inclusive as those in the existing literature.

earnings-to-price ratio of the S&P 500 index. These variables are constructed using data from Global Insight.

3.2. Statistical inference: the bootstrapping approach

There are a number of issues involved in statistical inference for mutual fund timing tests. When there is only one fund, the test for timing ability can be based on the t-statistic of the timing measure, $\hat{\gamma}$. However, when there are a large number of funds, even if none of them has timing ability, by random chance there will be some funds with significant timing measures based on the t-statistics. This is analogous to the data-snooping problem discussed in Lo and MacKinlay (1990) and Sullivan, Timmermann, and White (1999). One solution is to test whether all the timing statistics are jointly zero. However, the usual joint test (e.g., the Wald test) requires inversion of the covariance matrix of the timing statistics. When the number of funds is large and the length of the time series is short, the covariance matrix is likely to be poorly estimated or even singular.

In this paper, we base the statistical inference on the cross-sectional distribution of timing measures. We focus on the following cross-sectional statistics: the mean, median, standard deviation, skewness, kurtosis, 25% and 75% quartiles, and 5%, 10%, 90%, and 95% extreme percentiles. The statistical significance at these extreme percentiles can tell us whether there are funds with strong positive timing ability or perversely negative timing ability.

There are two additional issues associated with statistical inference based on the cross-sectional distribution of the market timing measures. One is the violation of the *identically and independently distributed* (i.i.d.) assumption across funds. The other is the finite sample property of the test statistics. Consider the t-statistics of the timing measures, which are asymptotically normal under quite general conditions. However, since funds hold similar stocks, their betas are correlated. As a result, the timing measures are correlated and the t-statistics are not i.i.d. across funds. This issue is further complicated by the fact that funds may exist for only a short period of time and some funds may not overlap with each other at all. The finite sample distributions for the cross-sectional statistics, particularly for the extreme percentiles, may differ from their asymptotic counterparts. To address the above issues, we resort to a bootstrapping approach similar to that proposed by Kosowski, Timmermann, White, and Wermers (2005). In the bootstrapping procedure, we randomly resample the data under the null hypothesis of no market timing while maintaining the covariance structure across fund betas or across fund returns. Statistical inference is then based on whether the empirical distribution of market timing statistics is significantly different from what one would expect under the null hypothesis that no fund has timing ability.

4. Empirical analysis of market timing

4.1. Return-based timing measures

For comparison purposes, we first perform the return-based tests in (1) and (2) to obtain an estimate of γ for each fund, using monthly fund and market returns. The t-statistic for $\hat{\gamma}$ is computed using the Newey-West heteroskedasticity and autocorrelation-consistent variance estimator (see Newey and West (1987)) with a six-month lag. We then calculate the cross-sectional statistics for $\hat{\gamma}$ and the t-values. Statistical inference is based on the bootstrapped p-values following the procedure in Appendix C with 2,000 replications.

Fig. 1 plots the density of the cross-sectional distribution of t-statistics for all funds, together with that of the bootstrapped t-statistics. Both density functions are estimated using a Gaussian kernel and the bandwidth choice of Silverman (1986). The comparison of these two density functions provides informal but intuitive inference about the timing ability of mutual funds. For the Treynor-Mazuy t-statistics, the sample density plot is

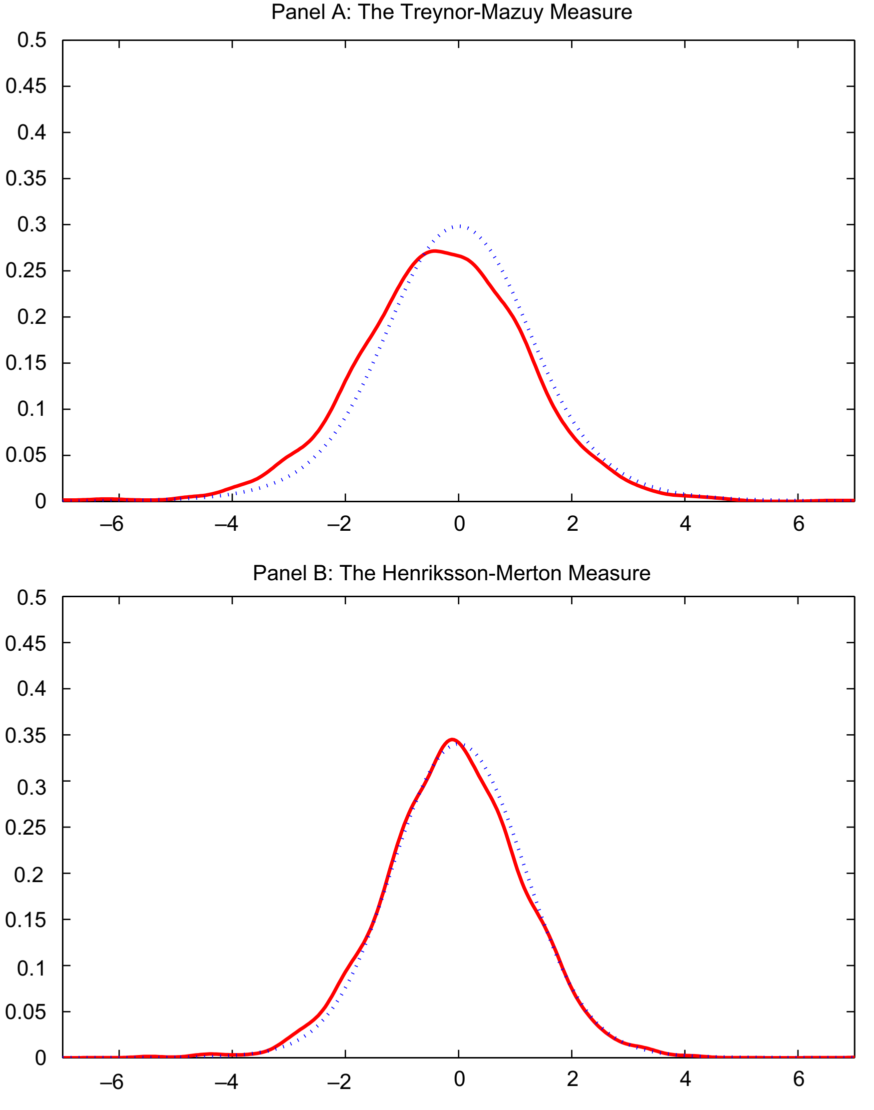

Fig. 1. The density of t-statistics for the return-based timing measures. This figure plots the density of the cross-sectional distribution of the t-statistics (the solid line) for the return-based timing measures, together with that of the bootstrapped t-statistics (the dotted line).

Table 2
Return-based timing tests
This table reports the cross-sectional distribution of the Treynor-Mazuy timing measure and the Henriksson-Merton timing measure ($\hat{\gamma}$), as well as the Newey-West t-statistics (t). The return horizon is one month. The bootstrapped p-values (p) for the timing measures and the Newey-West t-statistics are reported, respectively, in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

	5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: Treynor-Mazuy measure (1-month horizon)											
$\hat{\gamma}$	−2.10	−1.63	−0.79	−0.26	−0.15	0.30	0.79	1.25	1.09	0.32	9.10
p	(0.79)	(0.81)	(0.81)	(0.70)	(0.69)	(0.52)	(0.46)	(0.45)	(0.43)	(0.46)	(0.35)
t	−2.86	−2.14	−1.27	−0.28	−0.28	0.71	1.53	2.12	1.59	0.28	8.54
p	(0.88)	(0.84)	(0.77)	(0.67)	(0.66)	(0.51)	(0.42)	(0.37)	(0.06)	(0.32)	(0.04)
Panel B: Henriksson-Merton measure (1-month horizon)											
$\hat{\gamma}$	−0.42	−0.32	−0.15	−0.04	−0.01	0.09	0.20	0.29	0.23	−0.25	6.38
p	(0.58)	(0.62)	(0.60)	(0.59)	(0.56)	(0.60)	(0.62)	(0.63)	(0.74)	(0.57)	(0.36)
t	−2.09	−1.63	−0.86	−0.08	−0.08	0.72	1.47	1.92	1.24	−0.03	3.96
p	(0.83)	(0.78)	(0.69)	(0.57)	(0.58)	(0.44)	(0.33)	(0.28)	(0.04)	(0.54)	(0.50)

visibly shifted to the left relative to the bootstrapped one, indicating overall negative timing ability. For the Henriksson-Merton t-statistics, the density plot is also left-shifted, but to a lesser extent.

Table 2 reports the return-based measures' cross-sectional statistics, t-statistics, and bootstrapped p-values. Consistent with the previous literature, the mean and median of the timing statistics are slightly negative, suggesting generally negative timing performance. However, both the mean and median timing measures are statistically insignificant based on the bootstrapped p-values. In addition, none of the low percentiles (5th and 10th) or high percentiles (90th and 95th) is statistically significant. In summary, the return-based measures are on average negative. This pattern is consistent with those documented in previous studies. For example, using various return-based measures, Henriksson (1984), Ferson and Schadt (1996), and Jiang (2003), among others, find that on average mutual funds exhibit negative market timing. As discussed in Section 2.3, the return-based results are likely subject to the artificial timing bias.

4.2. Holdings-based timing measures

For the holdings-based tests, we focus on the Treynor-Mazuy measure in (6), with monthly, quarterly, semiannual, and annual forecasting horizons. In other words, market returns are measured over one, three, six, and 12 months after the portfolio holdings date. The betas of individual stocks are estimated using one-year daily stock returns prior to the portfolio holdings date. To account for the effect of nonsynchronous trading, we estimate the market model for each stock using market returns up to five daily leads and five daily lags, in addition to the contemporaneous term. That is,

$$r_{i\tau} = a_i + \sum_{q=-5}^{5} b_{iq} r_{m,\tau-q} + e_{i\tau}. \tag{8}$$

Following Dimson (1979), the stock beta is the sum of the estimated coefficients, i.e., $\tilde{b}_i = \sum_{q=-5}^{5} \tilde{b}_{iq}$. We require that a stock have at least 60 daily observations during the estimation period, otherwise we assume a value of one for the stock beta. Non-stock securities are assumed to have a beta of zero.

As discussed in Section 2.3, to eliminate potential passive timing effects and reduce cross-sectional heteroskedasticity, we further adjust the stock beta estimates by subtracting the average beta of its characteristics-matched portfolio ($\tilde{b}_i^c$), i.e.,

$$\hat{b}_i = \tilde{b}_i - \tilde{b}_i^c. \tag{9}$$

The construction of passive characteristics portfolios is detailed in Appendix A. It should be noted that our finding of positive timing ability by mutual funds is not driven by this adjustment. Using unadjusted beta $\tilde{b}_i$ to perform the timing tests, there are no material changes to the conclusions. With the estimates of fund beta based on (5), the timing coefficient γ is estimated from (6) for each fund. The t-statistic of $\hat{\gamma}$ is computed following the Newey-West method with a six-month (two-quarter) lag.[6] Finally, the bootstrapped p-values for the cross-sectional statistics are calculated using 2,000 replications following the procedure in Appendix C.

Fig. 2 plots the density of the cross-sectional distribution of the t-statistics estimated for all funds, together with that of the bootstrapped t-statistics. The density functions are estimated in the same way as in the case of return-based measures. For all four forecasting horizons (one, three, six, and 12 months), the sample density functions are right-shifted relative to the bootstrapped ones, indicating overall positive timing ability by mutual funds.

Table 3 reports the cross-sectional statistics of the holdings-based measures and their t-statistics, as well as the bootstrapped p-values. At the one-month forecasting horizon, the mean and median t-statistics are positive, but not statistically significant according to the bootstrapped p-values. At the three-month and six-month forecasting horizons, however, the bootstrapped p-values for the mean, median, and 25th, 75th, and 90th percentiles are all much smaller, below 10%. The results suggest that over these intermediate horizons, on average there is significantly positive timing ability by mutual funds. Some funds exhibit strong positive timing ability, but there is no evidence of perversely strong negative timing. The evidence is less significant at the 12-month forecasting horizon. Considering a nearly 100% average annual turnover ratio for mutual funds in our sample, the insignificant timing ability over the 12-month horizon should come as no surprise. Finally, compared to the timing measure $\hat{\gamma}$, the t-statistics have much lower kurtosis. This is consistent with statistical theory that suggests that as a pivotal statistic, the t-statistics offer more robust inference.[7] Therefore, while we report the bootstrapped p-values for both the cross-sectional $\hat{\gamma}$ and t-statistics, we focus on the latter for more reliable inference.

[6]When we examine long forecasting horizons, such as 12 months, a six-month lag may not be sufficient to capture serial correlation in the overlapping data. However, since the bootstrapping procedure is also based on overlapping data and the same lag order, the bootstrapped p-values should be robust to the choice of lag order. As an extra caution, we also set the lag order to a higher value up to 12 months as well as a data-dependent value of $0.5T^{1/3}$. There are no material changes in the results.

[7]The timing measures ($\hat{\gamma}$) are nonpivotal statistics and are heteroskedastic across funds as their distributions depend on nuisance parameters such as the variance of the error terms. The t-statistics are pivotal and have asymptotic standard normal distributions. Statistical theory (e.g., Hall, 1992) suggests that bootstrapped pivotal statistics have better convergence properties and are more homogeneous.

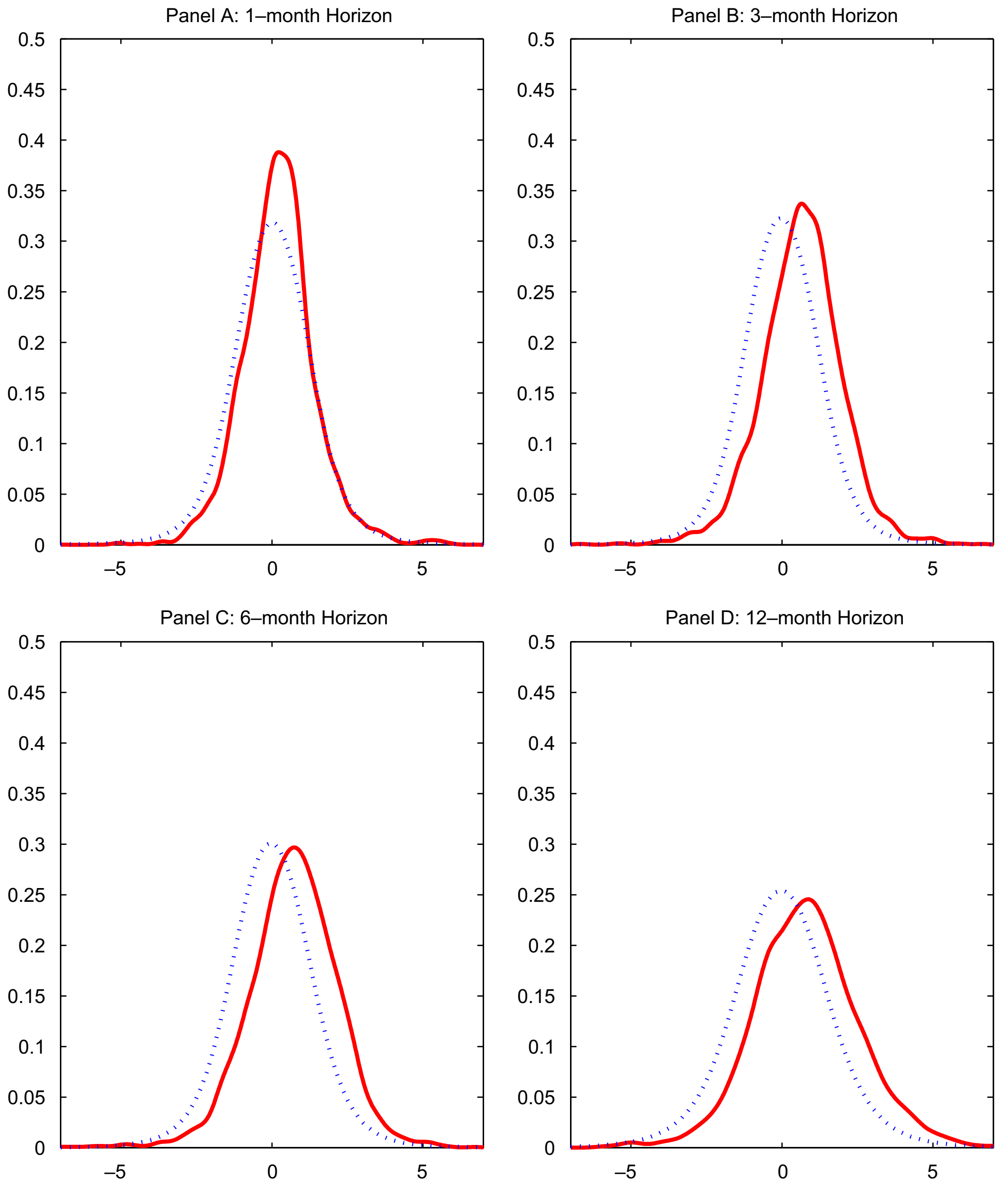

Fig. 2. The density of t-statistics for the holdings-based timing measures. This figure plots the density of the cross-sectional distribution of the t-statistics (the solid line) for the holdings-based timing measures (with one-, three-, six-, and 12-month horizons), together with that of the bootstrapped t-statistics (the dotted line).

The positive timing ability documented in Table 3 is in clear contrast with the negative evidence from Table 2. To reconcile the difference, we perform further analysis by constructing matching passive portfolios for mutual funds. Specifically, we replace

Table 3

Holdings-based tests: Treynor-Mazuy timing measure

This table reports the cross-sectional distribution of the holdings-based Treynor-Mazuy timing measure ($\hat{\gamma}$) and the Newey-West t-statistics (t) for the one, three, six, and 12 months horizons. The stock betas are estimated using the past one-year daily returns. The timing measures $\hat{\gamma}$ are pre-multiplied by 100. The bootstrapped p-values (p) for the timing measures and the Newey-West t-statistics are reported, respectively, in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

	5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: 1-Month horizon											
$\hat{\gamma}$	−1.92	−1.19	−0.44	0.33	0.19	1.01	2.10	3.04	1.71	0.78	12.74
p	(0.39)	(0.36)	(0.35)	(0.17)	(0.20)	(0.11)	(0.10)	(0.11)	(0.36)	(0.33)	(0.80)
t	−1.63	−1.21	−0.48	0.23	0.21	0.89	1.63	2.19	1.23	0.38	3.94
p	(0.13)	(0.18)	(0.19)	(0.26)	(0.29)	(0.39)	(0.38)	(0.33)	(0.76)	(0.14)	(0.32)
Panel B: 3-Month horizon											
$\hat{\gamma}$	−0.81	−0.48	−0.05	0.35	0.31	0.73	1.24	1.64	0.84	0.22	6.94
p	(0.17)	(0.14)	(0.05)	(0.05)	(0.03)	(0.03)	(0.08)	(0.14)	(0.52)	(0.45)	(0.97)
t	−1.48	−0.97	−0.14	0.65	0.67	1.44	2.23	2.69	1.33	−0.16	1.59
p	(0.11)	(0.08)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.25)	(0.72)	(0.74)
Panel C: 6-Month horizon											
$\hat{\gamma}$	−0.73	−0.43	−0.04	0.27	0.25	0.61	1.00	1.30	0.68	0.13	6.43
p	(0.38)	(0.30)	(0.09)	(0.08)	(0.03)	(0.04)	(0.08)	(0.13)	(0.32)	(0.56)	(0.92)
t	−1.64	−1.07	−0.15	0.71	0.73	1.65	2.45	2.90	1.48	−0.36	2.33
p	(0.20)	(0.16)	(0.09)	(0.08)	(0.07)	(0.05)	(0.06)	(0.10)	(0.19)	(0.89)	(0.51)
Panel D: 12-Month horizon											
$\hat{\gamma}$	−0.63	−0.37	−0.08	0.18	0.18	0.44	0.74	0.97	0.52	−0.49	6.09
p	(0.42)	(0.34)	(0.21)	(0.20)	(0.10)	(0.12)	(0.19)	(0.24)	(0.44)	(0.68)	(0.84)
t	−1.90	−1.24	−0.33	0.81	0.80	1.88	3.01	3.84	1.83	0.00	2.42
p	(0.22)	(0.19)	(0.18)	(0.12)	(0.11)	(0.09)	(0.07)	(0.07)	(0.14)	(0.53)	(0.53)

individual stocks in the fund portfolio with their corresponding characteristics-benchmark portfolios (see details in Appendix A). Note that by construction the hypothetical buy-and-hold returns of these passive portfolios mainly capture the passive timing effect in fund returns. We then perform return-based market timing tests based on the passive portfolio returns. Consistent with the evidence in Section 2.1 based on characteristics-benchmark portfolios, the results show that the average passive timing effect is negative. That is, the passive timing effect in fund returns tends to offset the effect of positive timing. As such, even in the presence of active market timing, fund returns may still exhibit a concave or insignificantly convex relation with market returns.

We also perform the holdings-based Henriksson-Merton timing tests in (7). The results, not tabulated for brevity, exhibit similar patterns as the Treynor-Mazuy tests reported in Table 3.

4.3. Holdings-based tests using alternative stock beta estimates

To ensure that the above results are robust to how stock betas are estimated, we further perform the holdings-based tests using three alternative estimates of individual stock betas, namely, (i) stock beta estimates using daily returns over the past three months, (ii) stock beta estimates using monthly returns over the past five years, and (iii) stock beta estimates conditioning on firm characteristics and macroeconomic variables and using monthly returns over the past five years. When using the past three-month daily returns to estimate stock betas, we perform the same regression as specified in (8). A minimum of 22 daily return observations are required, otherwise we assume a value of one for the stock beta. When using the past five-year monthly returns to estimate stock betas, we perform the following regression:

$$r_{i\tau} = a_i + b_{i1} r_{m\tau} + b_{i2} r_{m\tau-1} + e_{i\tau}. \tag{10}$$

The stock beta estimate is then $\tilde{b}_i = \tilde{b}_{i1} + \tilde{b}_{i2}$. We require that a stock have at least 12 monthly return observations, otherwise we assume a value of one for the stock beta. For both of the above stock beta estimates, we further use (9) to obtain the characteristics-adjusted beta $\hat{b}_i$, which is then used to compute the fund beta according to (5).

To obtain stock betas conditional on firm characteristics and macroeconomic variables, we follow Avramov and Chordia (2006) and estimate the following model using monthly observations:

$$\begin{aligned} r_{i\tau} = {} & a_i + b_{i1} r_{m\tau} + b_{i2} z_{\tau-1} r_{m\tau} + b_{i3} \mathrm{Size}_{i\tau-1} r_{m\tau} + b_{i4} \mathrm{Size}_{i\tau-1} z_{\tau-1} r_{m\tau} \\ & + b_{i5} \mathrm{BM}_{i\tau-1} r_{m\tau} + b_{i6} \mathrm{BM}_{i\tau-1} z_{\tau-1} r_{m\tau} + e_{i\tau}, \end{aligned} \tag{11}$$

where $z_{\tau-1}$ is the credit premium (a macroeconomic variable), $\mathrm{Size}_{i\tau-1}$ is the log of market capitalization, and $\mathrm{BM}_{i\tau-1}$ is the book-to-market ratio. The credit premium is the average yield spread between Moody's Baa-rated corporate bonds and Aaa-rated corporate bonds, where we obtain the data from Global Insight. Data on market capitalization come from CRSP. The book value of equity used in computing the book-to-market ratio is based on the end of the most recently reported fiscal year from Compustat. All three conditioning variables are lagged one month. The model is estimated on a rolling basis using monthly data over the past 60 months. The stock beta estimate is based on the estimated coefficients, i.e., $\hat{b}_{it} = \tilde{b}_{i1} + \tilde{b}_{i2} z_{t-1} + \tilde{b}_{i3} \mathrm{Size}_{it-1} + \tilde{b}_{i4} \mathrm{Size}_{it-1} z_{t-1} + \tilde{b}_{i5} \mathrm{BM}_{it-1} + \tilde{b}_{i6} \mathrm{BM}_{it-1} z_{t-1}$. Since there are more independent variables involved in model (11), we require that a stock have at least 36 monthly observations. Otherwise, we replace the conditional beta with the unconditional beta estimated using the past five-year monthly returns. For the conditional beta estimates, we do not make characteristics-based adjustments.

These three alternative estimators may capture different time-varying components of the stock beta. The estimator based on five years of monthly returns may capture the long-run component, the estimator based on three months of daily returns is more likely to capture the short-term dynamic component, and the third estimator explicitly takes into account the beta variation associated with firm characteristics and macroeconomic conditions.

Table 4 reports the cross-sectional distribution of the t-statistics for the holdings-based measure with alternative stock beta estimates. The results are generally consistent with those reported in Table 3, where stock betas are estimated using 12 months of daily returns. At the one- and 12-month horizons there is positive but insignificant evidence of timing. At the three- and six-month horizons the mean and median of the t-statistics are significantly positive. The only noticeable difference in comparison with Table 3 is that when stock betas are estimated using five years of monthly returns, the results are slightly weaker at the six-month horizon, where the bootstrapped p-values for the mean and

Table 4
Holdings-based tests: alternative stock beta estimates

This table reports the cross-sectional distribution of the Newey-West t-statistics (t) for the holdings-based Treynor-Mazuy timing measure with alternative estimates of stock betas. The stock betas are estimated from (i) the past three-month daily returns (3m), (ii) the past five-year monthly returns (5y), and (iii) the past five-year monthly returns conditional on firm characteristics and macroeconomic variables (5yc). The bootstrapped p-values (p) for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

$\hat{\beta}$		5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: 1-Month horizon												
3m	t	−1.64	−1.17	−0.40	0.32	0.28	1.05	1.79	2.34	1.23	0.13	1.45
	p	(0.11)	(0.09)	(0.09)	(0.10)	(0.12)	(0.16)	(0.23)	(0.23)	(0.86)	(0.31)	(0.84)
5y	t	−1.62	−1.24	−0.64	0.18	0.15	0.91	1.60	2.18	1.24	0.59	3.63
	p	(0.14)	(0.23)	(0.32)	(0.30)	(0.36)	(0.37)	(0.35)	(0.27)	(0.68)	(0.09)	(0.39)
5yc	t	−1.62	−1.17	−0.55	0.28	0.26	1.05	1.80	2.34	1.24	0.12	1.47
	p	(0.08)	(0.06)	(0.14)	(0.13)	(0.17)	(0.18)	(0.17)	(0.33)	(0.71)	(0.33)	(0.80)
Panel B: 3-Month horizon												
3m	t	−1.66	−1.15	−0.39	0.44	0.41	1.26	2.04	2.61	1.37	0.12	2.46
	p	(0.09)	(0.07)	(0.08)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.26)	(0.33)	(0.44)
5y	t	−1.56	−1.17	−0.42	0.53	0.57	1.38	2.16	2.72	1.44	0.53	4.58
	p	(0.19)	(0.21)	(0.14)	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.17)
5yc	t	−1.67	−1.21	−0.45	0.50	0.52	1.39	2.19	2.72	1.38	0.09	0.71
	p	(0.20)	(0.16)	(0.09)	(0.03)	(0.03)	(0.01)	(0.02)	(0.06)	(0.14)	(0.37)	(0.94)
Panel C: 6-Month horizon												
3m	t	−1.73	−1.21	−0.48	0.43	0.42	1.34	2.11	2.65	1.41	0.23	1.72
	p	(0.14)	(0.11)	(0.11)	(0.07)	(0.07)	(0.03)	(0.07)	(0.08)	(0.22)	(0.83)	(0.57)
5y	t	−1.76	−1.29	−0.56	0.52	0.55	1.43	2.37	3.01	1.55	0.15	1.64
	p	(0.35)	(0.33)	(0.28)	(0.12)	(0.11)	(0.09)	(0.06)	(0.06)	(0.09)	(0.32)	(0.73)
5yc	t	−1.91	−1.38	−0.59	0.53	0.50	1.53	2.56	3.26	1.61	0.16	0.72
	p	(0.38)	(0.32)	(0.24)	(0.05)	(0.06)	(0.02)	(0.01)	(0.03)	(0.03)	(0.29)	(0.94)
Panel D: 12-Month horizon												
3m	t	−2.06	−1.43	−0.61	0.30	0.27	1.18	2.06	2.76	1.52	0.13	2.47
	p	(0.28)	(0.20)	(0.19)	(0.18)	(0.18)	(0.18)	(0.18)	(0.13)	(0.28)	(0.33)	(0.34)
5y	t	−2.19	−1.52	−0.77	0.31	0.27	1.29	2.33	3.01	1.70	0.08	3.06
	p	(0.44)	(0.38)	(0.39)	(0.32)	(0.34)	(0.27)	(0.21)	(0.21)	(0.25)	(0.45)	(0.39)
5yc	t	−2.16	−1.59	−0.73	0.49	0.40	1.50	2.78	3.76	1.88	0.30	2.15
	p	(0.41)	(0.38)	(0.32)	(0.14)	(0.18)	(0.10)	(0.03)	(0.02)	(0.02)	(0.19)	(0.44)

median t-statistics are slightly above the 10% level. Overall, using different methods of estimating stock betas, the holdings-based tests provide robust evidence of positive market timing performance by mutual funds.

4.4. Holdings-based tests using active changes of fund beta

The holdings-based tests essentially measure the covariance between fund beta levels at the beginning of a holding period and the holding period market returns. Note that time variation in fund betas can be driven by both active fund trading activities and passive portfolio weight changes due to nonproportional changes in stock prices. Changes in fund beta due to active trading are more relevant when measuring active market timing. To examine the effect of active trading on market timing, we further decompose a fund's beta into two components,

$$\hat{\beta}_t = \hat{\beta}_t^{t-h} + \Delta\hat{\beta}_t, \tag{12}$$

where $\hat{\beta}_t^{t-h}$ is what the portfolio beta of a fund would be at month t if it passively holds all the positions in its portfolio from h months ago (at month $t-h$), and $\Delta\hat{\beta}_t$ is the change in beta due to fund trading activities over the past h months. In this section, we use fund beta changes ($\Delta\hat{\beta}_t$), instead of fund beta levels ($\hat{\beta}_t$), in performing the market timing tests.

Specifically, fund beta changes due to active trading from month $t-h$ to month t are constructed as follows:

$$\Delta\hat{\beta}_t = \hat{\beta}_t - \hat{\beta}\hat{\beta}_t^{t-h} = \sum_{i=1}^{N_t} \omega_{it}\hat{b}_{it} - \sum_{i=1}^{N_t^P} \omega_{it}^{t-h}\hat{b}_{it}, \tag{13}$$

where ω_{it} is the fund portfolio weight of asset i at month t, $\hat{b}_{it}$ is the beta estimate of asset i at month t, and N_t is the number of assets held in the fund portfolio at month t. In the second term, ω_{it}^{t-h} is the passive portfolio weight of asset i at month t inferred from fund portfolio holdings at month $t-h$, and N_t^P is the number of assets held by the fund at month $t-h$. The general expression of ω_{it}^{t-h} is given by

$$\omega_{it}^{t-h} = \frac{n_{it}^{t-h} p_{it}}{\sum_{i=1}^{N_t^P} n_{it}^{t-h} p_{it}}, \tag{14}$$

where p_{it} is the price of asset i at month t, and n_{it}^{t-h} is the number of shares of asset i held by the passive portfolio at month t, based on the number of shares of asset i held by the fund at month $t-h$ (n_{it-h}). In calculating n_{it}^{t-h}, we adjust n_{it-h} with the share adjustment factor from CRSP, which takes into account stock splits and stock dividends, and we take into account fund portfolio changes that are not the result of active fund trading. For instance, even by passively holding all positions in a fund portfolio, a new stock can be added to the portfolio due to spinoffs or an existing stock can be dropped from the portfolio due to delisting. In the case of spinoffs, we maintain the assumption that new shares received by a fund are passively held by the fund. In the case of mergers and acquisitions, shares of the target firms are exchanged into shares of the acquiring firms. CRSP does not provide the exact number of new shares distributed to each existing share in spinoffs or share exchange ratios for mergers and acquisitions. Thus, in the case of spinoffs we calculate the number of new shares received per share of the parent firm using the cash values of spinoffs and the ex date prices of the new issues in CRSP. For mergers and acquisitions, share exchange ratios are calculated from the delisting values of target firms and the post-merger prices of the acquiring firms in CRSP. In addition, stocks delisted for performance-related reasons during the period between months $t-h$ and t are assumed to have zero values in passive fund holdings at month t. Finally, all cash dividends received by a fund are kept in cash positions. For cash and other non-equity holdings, we assume a monthly rate of return equal to the one-month T-bill yield.

We perform the Treynor-Mazuy market timing tests in (6) using fund betas' active changes, $\Delta\hat{\beta}_t$. Again, we focus on monthly, quarterly, semiannual, and annual forecasting horizons. The betas of individual stocks ($\hat{b}_{it}$) are estimated using one-year daily returns prior to the date of portfolio holdings, without adjusting for stock characteristics. The time lag h is set to 12 months. That is, we measure fund betas' active changes over the past 12 months. The results based on fund beta changes over the past six months (i.e., $h = 6$ months) are consistent but are slightly weaker, with the estimates of timing coefficient in general smaller in magnitude. Similar to the tests using fund beta levels, we require that a fund have at least eight observations of $\Delta\hat{\beta}_t$ to ensure the robustness of inference for trading-based measures.

Table 5 reports the cross-sectional statistics of the holdings-based measures using active changes in fund beta, as well as the t-statistics and bootstrapped p-values. The t-statistic of $\hat{\gamma}$ is computed following the Newey-West method with a six-month (two-quarter) lag. Similar to Table 3, the mean and median estimates of the timing coefficient ($\hat{\gamma}$) are positive at all horizons. At both one-month and 12-month forecasting horizons, the mean and median timing measures are statistically insignificant, as indicated by the bootstrapped p-values. At three-month and six-month horizons, however, the mean and median timing measures are all significantly positive. The 75th, 90th, and 95th percentiles of the t-statistics are also significantly higher than the bootstrapped ones, indicating strong timing ability by

Table 5
Holdings-based tests using active changes of fund beta

This table reports the cross-sectional distribution of the holdings-based Treynor-Mazuy timing measure ($\hat{\gamma}$) and the Newey-West t-statistics (t) using fund betas' active changes, which are defined in (13). The timing measures $\hat{\gamma}$ are pre-multiplied by 100. The bootstrapped p-values (p) for the timing measures and the Newey-West t-statistics are reported, respectively, in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

	5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: 1-Month horizon											
$\hat{\gamma}$	−2.42	−1.44	−0.40	0.45	0.28	1.22	2.39	3.17	2.08	2.64	29.88
p	(0.53)	(0.42)	(0.22)	(0.12)	(0.11)	(0.05)	(0.10)	(0.18)	(0.17)	(0.05)	(0.14)
t	−1.73	−1.22	−0.49	0.30	0.30	1.04	1.79	2.37	1.30	0.07	2.43
p	(0.11)	(0.12)	(0.16)	(0.19)	(0.20)	(0.25)	(0.28)	(0.28)	(0.84)	(0.42)	(0.76)
Panel B: 3-Month horizon											
$\hat{\gamma}$	−1.01	−0.58	−0.13	0.33	0.23	0.72	1.44	1.95	1.02	0.47	8.40
p	(0.26)	(0.16)	(0.09)	(0.08)	(0.06)	(0.05)	(0.05)	(0.09)	(0.25)	(0.39)	(0.85)
t	−1.60	−1.14	−0.31	0.48	0.49	1.27	2.01	2.58	1.34	0.02	2.19
p	(0.10)	(0.13)	(0.08)	(0.09)	(0.07)	(0.10)	(0.13)	(0.13)	(0.55)	(0.48)	(0.77)
Panel C: 6-Month horizon											
$\hat{\gamma}$	−0.83	−0.48	−0.07	0.28	0.24	0.62	1.08	1.51	0.76	0.35	4.94
p	(0.41)	(0.29)	(0.10)	(0.08)	(0.04)	(0.04)	(0.07)	(0.09)	(0.23)	(0.43)	(0.97)
t	−1.65	−1.12	−0.23	0.61	0.60	1.41	2.33	2.92	1.48	0.06	3.40
p	(0.14)	(0.13)	(0.09)	(0.09)	(0.08)	(0.11)	(0.09)	(0.11)	(0.32)	(0.46)	(0.45)
Panel D: 12-Month horizon											
$\hat{\gamma}$	−0.60	−0.38	−0.10	0.13	0.13	0.37	0.66	0.89	0.49	−0.64	7.35
p	(0.33)	(0.31)	(0.22)	(0.25)	(0.15)	(0.19)	(0.26)	(0.32)	(0.60)	(0.64)	(0.75)
t	−2.02	−1.27	−0.41	0.51	0.56	1.43	2.37	2.91	1.61	−0.30	4.10
p	(0.22)	(0.16)	(0.18)	(0.20)	(0.16)	(0.20)	(0.21)	(0.27)	(0.54)	(0.79)	(0.36)

some funds. These results suggest that fund betas' active changes are significantly correlated with future market returns, providing further evidence of positive market timing ability by mutual funds.

4.5. Economic value of market timing

A few studies, such as Henriksson and Merton (1981), Merton (1981), and Glosten and Jagannathan (1994), point out that market timing is equivalent to a contingent claim on the market index. We follow this intuition to quantify the economic value of the market timing ability of mutual funds in our sample. Specifically, given the fund return process in (1), market timing generates an additional terminal payoff γr_{mt}^2 to a fund with the maturity of the payoff equal to the forecasting horizon, where r_{mt} is the market return in excess of the constant risk-free rate r_f. Assume that the market index P_m follows a geometric Brownian motion, $\mathrm{d}P_m/P_m = u_m\,\mathrm{d}t + \sigma_m\,\mathrm{d}W$. Then the value of market timing V can be derived as the expected present value of γr_{mt}^2 under the risk-neutral measure (Q):

$$V = \frac{1}{1+r_f}\mathrm{E}^Q[\gamma r_{mt}^2] = (1+r_f)\gamma(e^{\sigma_m^2} - 1). \tag{15}$$

To illustrate, note that at the six-month horizon, the median timing measure is 0.25% in Table 3 (based on the level of fund beta) and 0.24% in Table 5 (based on active changes in fund beta). We further assume an annual risk-free rate of 3% and monthly market return variance of 0.2% based on the CRSP value-weighted index. Based on the above timing measures, the economic values of market timing from (15) are, respectively, 0.63% and 0.60%/year. These values are economically large, indicating that market timing can be an important investment strategy for mutual funds. We note, however, that the numbers cannot be directly interpreted as the realized contribution to fund performance since they are derived under the assumption of continuous trading and zero transaction costs. Also, the timing coefficients used in computing the economic values are the holdings-based measures that capture only active timing, without adjusting for the negative passive timing effect in fund returns.

5. Further analysis

Having documented statistically and economically significant mutual fund market timing performance, we conduct further analysis to investigate the following issues. First, which types of funds are more likely to be successful market timers? Second, do fund managers use public information, or any additional private information, in predicting and timing market returns? And third, motivated by a recent study by Avramov and Wermers (2006), is industry shifting an important element of mutual funds' market timing strategy?

5.1. Characteristics of market timers

While the results in the previous section indicate positive timing ability for the average fund, they also show a wide variation of timing performance across funds. It is therefore tempting to identify funds or fund managers that possess superior market timing ability. Unfortunately, as noted earlier, out of a large sample of funds, it is difficult to pinpoint whether the timing ability of a specific fund is due to true ability or pure luck. However, we can investigate whether there is any commonality among market timers. For example, are small funds or large funds better market timers? Do successful market timing funds charge high fees to extract rents from investors? Does active market timing result in frequent trading and thus high turnover? Furthermore, we are interested in whether market timing funds are associated with certain investment styles, and whether they make concentrated sector bets in their portfolios.

To examine the above questions, we consider seven fund characteristics. The first three are total net assets, the expense ratio, and turnover. In each year, we obtain the percentile rank of each variable across funds. The annual percentile ranks are then averaged over time for each fund. We use the percentile ranks, instead of the underlying variable, to remove any potential time trend in the variable. The next three characteristics are related to the stocks held by a fund, namely, market capitalization or size, book-to-market, and momentum as measured by the past 12-month returns. In each quarter, we first obtain the percentile ranks of size, book-to-market, and momentum for all stocks in CRSP. We then compute the weighted average of the percentile ranks for each fund, where the weights are proportional to the values of stock positions in the portfolio. Finally, these quarterly measures are averaged over time for each fund. The last fund characteristic we include is the industry concentration index (ICI). This measure is computed for each fund following Kacperczyk, Sialm, and Zheng (2005) and using the Fama-French 12-industry SIC classification. ICI is also computed quarterly and then averaged over time.

We perform the following panel data regression to gauge the relation between various fund characteristics and the timing measure:

$$\hat{\beta}_{it} = a_i + (b_0 + b\mathrm{C}_i)r_{m,t+1} + e_{it}, \tag{16}$$

$$\Delta\hat{\beta}_{it} = a_i + (b_0 + b\mathrm{C}_i)r_{m,t+1} + e_{it}, \tag{17}$$

where C_i denotes the vector of fund characteristics, and $\hat{\beta}_{it}$ is the fund beta calculated from stock beta estimates using the past one-year daily returns. The models are estimated with both fixed fund and time effects. Since the fixed effects may not completely capture the covariance of the error terms, we further rely on the bootstrapped p-values for statistical inference.

The results are reported in Table 6. A noted pattern is that active market timing funds tend to have high industry concentration in their portfolios, as the coefficient of industry concentration index (ICI) is significantly positive in both regressions (16) and (17). This suggests a possible link between concentrated sector bets and market timing. To a lesser extent, active market timers also tend to have large fund size and a tilt toward small-cap stocks. Coefficients of these fund characteristics are significant in only one of the above regressions. Note that by the definition of the p-values in Eq. (C.1) of Appendix C, the high p-value (close to one) for the coefficient on stock size (SIZE) suggests that the coefficient estimate is significantly below the corresponding bootstrapped values. As noted in Chen, Hong, Huang, and Kubik (2004), fund size is detrimental to stock selection performance due to liquidity issues. As a result, large funds often resort to strategies involving sector bets or industry shifts. We perform further analysis on industry allocation in Section 5.3 to shed more light on this issue. Finally, small-cap stocks have a wide dispersion of betas, which makes it easier for funds to shift portfolio beta by changing the weights of such stocks. Kacperczyk, Sialm, and Zheng (2005) also document that funds with high industry concentration tend to overweigh small-cap stocks.

Table 6
Characteristics of market timers

This table reports the coefficient estimates ($\hat{b}$) of fund characteristics in regressions (16) and (17). The regression is performed jointly for all funds with a vector of fund characteristics. Fund betas are computed from stock betas estimated using the past one-year daily returns. TNA, EXPENSE, and TURNOVER denote the average cross-fund percentile ranks of total net assets (TNA), the expense ratio, and turnover for each fund. SIZE, B/M RATIO, and MOMENTUM denote the average cross-sectional percentile ranks of stock characteristics in fund portfolios based on market capitalization, book-to-market, and momentum. ICI is the average industry concentration index for each fund computed following Kacperczyk, Sialm, and Zheng (2005). The bootstrapped p-values (p) for the coefficient estimates and t-statistics are reported in the parentheses underneath.

	TNA	EXPENSE	TURNOVER	SIZE	B/M RATIO	MOMENTUM	ICI
Panel A: Regression of fund beta levels							
3-Month horizon							
$\hat{b}$	1.15	−0.92	0.37	−0.34	−0.51	0.95	0.36
p	(0.12)	(0.82)	(0.60)	(0.65)	(0.66)	(0.15)	(0.07)
t	2.06	−1.43	0.61	−2.31	−1.84	1.93	3.44
p	(0.08)	(0.86)	(0.60)	(0.68)	(0.69)	(0.14)	(0.04)
6-Month horizon							
$\hat{b}$	0.76	−0.74	−0.38	−0.27	−0.17	0.77	0.25
p	(0.14)	(0.81)	(0.74)	(0.68)	(0.49)	(0.20)	(0.08)
t	1.94	−1.50	−0.84	−2.52	−1.57	2.08	3.33
p	(0.11)	(0.83)	(0.78)	(0.69)	(0.49)	(0.18)	(0.06)
Panel B: Regression of fund betas' active changes							
3-Month horizon							
$\hat{b}$	0.98	−0.52	1.13	−0.82	−0.61	0.29	0.41
p	(0.20)	(0.72)	(0.10)	(0.91)	(0.68)	(0.39)	(0.06)
t	1.42	−0.51	1.98	−4.32	−1.66	0.93	3.06
p	(0.17)	(0.75)	(0.12)	(0.93)	(0.71)	(0.37)	(0.04)
6-Month horizon							
$\hat{b}$	0.54	−0.75	0.89	−0.71	−0.17	0.17	0.22
p	(0.27)	(0.82)	(0.29)	(0.90)	(0.48)	(0.56)	(0.10)
t	1.09	−1.09	1.74	−4.93	−1.39	0.82	2.41
p	(0.25)	(0.83)	(0.28)	(0.91)	(0.48)	(0.55)	(0.08)

5.2. Public information and market timing

The literature documents that macroeconomic variables such as the interest rate and aggregate dividend yield can predict stock market returns. Note that the horizons at which these variables predict market returns are very similar to the horizons at which we find active market timing for mutual funds. It is thus natural to ask whether mutual funds rely on these macroeconomic variables to time the market. Here, we focus on five return-predictive economic variables well documented in the finance and economics literature, namely, the short-term interest rate, term premium, credit premium, aggregate dividend yield, and aggregate earnings-to-price ratio. We use the one-month T-bill yield as the short-term interest rate. The term premium is the yield spread between the 10-year T-bond and one-month T-bill. The credit premium is the average yield spread between Moody's Baa-rated and Aaa-rated corporate bonds. The aggregate dividend yield and earnings-to-price ratio are based on the S&P 500 index.

To see if these variables are part of fund managers' information set, we perform the following regressions for each fund:

$$\hat{\beta}_t = a + bM_{t-1} + e_t, \quad (18)$$

$$\Delta\hat{\beta}_t = a + bM_{t-1} + e_t, \quad (19)$$

where M_{t-1} is the vector of five economic variables (lagged one month) and b is a vector of coefficients. Since the above models involve more regressors than the holdings-based tests in (6), we require at least 10 quarterly observations on $\hat{\beta}_t$ (or $\Delta\hat{\beta}_t$) for a fund to be included in the analysis. We report the mean and median t-statistics of the estimated coefficients, as well as their bootstrapped p-values, in Panel A of Table 7. Overall, the results based on fund beta levels and fund betas' active changes are consistent. The mean and median coefficient estimates on the short-term interest rate, term premium, and credit premium are all statistically insignificant, but the mean and median coefficient estimates of the aggregate dividend yield and the aggregate earnings-to-price ratio are significantly positive in both regressions. Thus, fund managers do take advantage of some of the publicly available macroeconomic information when they adjust the market exposure of fund portfolios.

Also of interest is whether fund managers possess any private information about market returns that is not captured by these macroeconomic variables. To address this question, we perform the following holdings-based tests for each fund in the spirit of Ferson and Schadt (1996):

$$\hat{\beta}_t = a + bM_{t-1} + \gamma r_{m,t+1} + e_t, \quad (20)$$

$$\Delta\hat{\beta}_t = a + bM_{t-1} + \gamma r_{m,t+1} + e_t. \quad (21)$$

We are interested in whether the timing measure γ remains significant after controlling for the return-predictive economic variables. The results are reported in Panel B of Table 7, where the bootstrapped p-values are computed following the procedure detailed in Appendix C. For both models (20) and (21), adding macroeconomic variables does not substantially change the test statistics of the timing measures. At three- and six-month forecasting horizons, the bootstrapped p-values for the mean and median t-statistics are all below 10%.

In summary, fund managers adjust the market exposure of fund portfolios in response to public information, especially to variables such as the aggregate dividend yield and the aggregate earnings-to-price ratio. But they do not rely solely on those variables to time the market. That is, fund managers may possess information about market returns beyond what is contained in the five macroeconomic variables. Identifying such information remains an interesting task.

5.3. *Industry allocation and market timing*

The above results suggest that fund managers use macroeconomic information related to the business cycle in market timing. In a recent study, Avramov and Wermers (2006) document that optimal portfolios of mutual funds formed on macroeconomic information exhibit large variations in industry tilt over the business cycle. They find that shifting industry weights is an important mechanism for achieving superior portfolio returns. The results in Table 6 also suggest that funds with high industry concentration in their portfolios are more likely to be active market timers. In the following

Table 7
Public information and market timing

Panel A reports the Newey-West t-statistics (t) of the mean and median coefficient estimates of the macroeconomic variables in regressions (18) and (19). The regressions are performed for each fund with a vector of five economic variables—short-term interest rate (Short Rate), term premium (Term Premium), credit premium (Credit Premium), aggregate dividend yield (Dividend Yield), and aggregate earnings-to-price ratio (Earnings/Price). Panel B reports the cross-sectional distribution of the Newey-West t-statistics (t) for the γ estimates of the regressions (20) and (21) at three-month (3-M) and six-month (6-M) horizons. Fund betas are computed from stock betas estimated using the past one-year daily returns. The bootstrapped p-values (p) for t-statistics are reported in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

Panel A: Fund beta and macroeconomic variables

		Short rate	Term premium	Credit premium	Dividend yield	Earnings/price
Regression of fund beta levels						
Mean	t	−0.26	−0.23	−0.13	1.10	1.23
	p	(0.65)	(0.61)	(0.55)	(0.04)	(0.02)
Median	t	−0.23	−0.32	−0.03	0.98	1.14
	p	(0.63)	(0.66)	(0.50)	(0.06)	(0.03)
Regression of fund betas' active changes						
Mean	t	−0.19	0.06	−0.21	1.06	1.15
	p	(0.64)	(0.46)	(0.65)	(0.00)	(0.01)
Median	t	−0.21	0.08	−0.12	0.91	1.07
	p	(0.67)	(0.45)	(0.59)	(0.01)	(0.01)

Panel B: Market timing controlling for public information

		5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Holdings-based tests using fund beta levels												
3-M	t	−1.43	−0.90	−0.06	0.70	0.72	1.47	2.16	2.73	1.29	0.04	1.94
	p	(0.09)	(0.05)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.20)	(0.42)	(0.46)
6-M	t	−1.64	−1.06	−0.19	0.66	0.71	1.53	2.31	2.90	1.54	−1.31	19.26
	p	(0.16)	(0.11)	(0.06)	(0.06)	(0.04)	(0.05)	(0.07)	(0.07)	(0.10)	(0.98)	(0.03)
Holdings-based tests using fund betas' active changes												
3-M	t	−1.71	−1.09	−0.34	0.46	0.45	1.22	1.96	2.57	1.40	0.47	5.85
	p	(0.27)	(0.13)	(0.09)	(0.08)	(0.08)	(0.08)	(0.07)	(0.05)	(0.09)	(0.12)	(.16)
6-M	t	−1.72	−1.09	−0.21	0.76	0.69	1.62	2.60	3.40	1.60	0.52	2.83
	p	(0.18)	(0.10)	(0.05)	(0.03)	(0.04)	(0.02)	(0.02)	(0.01)	(0.06)	(0.15)	(.51)

we examine a related issue, i.e., whether mutual funds actively shift industry weights to time the market.

To perform the analysis, we construct industry beta and industry-adjusted beta for the funds in our sample. The industry beta of a fund is the weighted average of industry portfolio betas:

$$\beta_t^I = \sum_{k=1}^{N_I} \omega_{kt}^I \hat{b}_{kt}^I, \tag{22}$$

where ω_{kt}^I is the fund portfolio weight in industry k at the beginning of period $t+1$, $\hat{b}_{kt}^I$ is the beta of industry k, calculated as the value-weighted average of individual stock betas (estimated using the past one-year daily returns), and N_I is the number of industries. We classify stocks into 12 industries following the Fama-French industry SIC classification. The industry-adjusted beta of a fund is the difference between the beta of the fund portfolio and the industry beta, $\hat{\beta}_t - \beta_t^I$, where $\hat{\beta}_t$ is the fund beta computed in (5), based on individual stock betas estimated using the past one-year daily returns.

We perform holdings-based market timing tests using both the industry beta and the industry-adjusted beta, with results reported in Table 8. The measures based on industry beta represent market timing achieved by industry rotation, whereas the measures based on industry-adjusted beta represent market timing achieved through intra-industry allocation. The results are consistent when the tests are based on fund beta levels (Panel A) and when they are based on fund beta changes (Panel B). According to the bootstrapped p-values, at three-month and six-month horizons the mean and median timing measures based on both industry betas and industry-adjusted betas are significantly positive. However, the magnitude of the mean and median timing measures based on industry betas is substantially higher than those based on industry-adjusted betas.

The above results suggest that, on average, mutual funds shift portfolio weights to high beta industries in anticipation of market upswings, and to low beta industries when expecting market downturns. Industry rotation is an important element of the mutual fund market timing strategy.

To further link these findings with those in the previous section, we investigate whether mutual funds shift industry weights in response to business cycle-related macroeconomic information. We perform the analysis by regressing quarterly changes of fund industry beta $\Delta\beta_t^I$ against the five macroeconomic variables considered in the previous section, namely, the short-term interest rate, term premium, credit premium, aggregate dividend yield, and aggregate earnings-to-price ratio. The regression results show that changes in fund industry beta are negatively correlated with the level of the short rate, and positively correlated with both the aggregate dividend yield and the aggregate earnings-to-price ratio. The median coefficients of these variables are, respectively, −0.06 (0.97), 0.45 (0.00), and 0.02 (0.02), where the numbers in parentheses are the bootstrapped p-values. The results suggest that funds tend to shift to high (low) beta industries when the short rate is low (high), the aggregate dividend yield is high (low), or the aggregate earnings-to-price ratio is high (low). In other words, fund managers respond to macroeconomic information when shifting industry weights of fund portfolios.

Finally, we note that our findings are consistent with some anecdotal evidence regarding how fund managers achieve market timing. For example, in a letter to fund shareholders

Table 8
Industry allocation and market timing

This table reports the cross-sectional distribution of the holdings-based Treynor-Mazuy timing measure ($\hat{\gamma}$) based on fund industry beta in (22) and industry-adjusted beta, respectively. The tests are also performed using beta changes due to active trading. The timing coefficient estimates capture the effects of industry rotation and intra-industry allocation in market timing. The bootstrapped p-values (p) for the timing coefficient estimates ($\hat{\gamma}$) are reported in the parentheses underneath. "Stdev," "Skew," and "Kurto" denote the cross-sectional standard deviation, skewness, and excess kurtosis, respectively.

	5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: Holdings-based tests using find beta levels											
Effect of industry rotation at 3-month horizon											
$\hat{\gamma}$	−0.39	−0.24	−0.04	0.22	0.19	0.45	0.79	1.09	0.56	0.56	7.77
p	(0.19)	(0.16)	(0.09)	(0.06)	(0.04)	(0.10)	(0.11)	(0.14)	(0.61)	(0.32)	(0.91)
Effect of intra-industry allocation at 3-month horizon											
$\hat{\gamma}$	−0.56	−0.34	−0.07	0.07	0.08	0.26	0.50	0.68	0.42	−0.74	8.47
p	(0.19)	(0.13)	(0.06)	(0.08)	(0.05)	(0.12)	(0.14)	(0.19)	(0.89)	(0.68)	(0.75)
Effect of industry rotation at 6-month horizon											
$\hat{\gamma}$	−0.25	−0.18	−0.03	0.19	0.16	0.36	0.62	0.81	0.44	0.34	8.00
p	(0.24)	(0.25)	(0.16)	(0.08)	(0.06)	(0.08)	(0.13)	(0.17)	(0.42)	(0.36)	(0.80)
Effect of intra-industry allocation at 6-month horizon											
$\hat{\gamma}$	−0.56	−0.32	−0.08	0.06	0.08	0.24	0.44	0.58	0.37	−0.38	9.31
p	(0.40)	(0.31)	(0.10)	(0.10)	(0.05)	(0.11)	(0.21)	(0.26)	(0.62)	(0.84)	(0.51)
Panel B: Holdings-based tests using fund betas' active changes											
Effect of industry rotation at 3-month horizon											
$\hat{\gamma}$	−0.42	−0.24	−0.05	0.25	0.18	0.45	0.82	1.02	0.61	0.37	2.52
p	(0.23)	(0.17)	(0.11)	(0.08)	(0.03)	(0.09)	(0.11)	(0.35)	(0.43)	(0.58)	(0.35)
Effect of intra-industry allocation at 3-month horizon											
$\hat{\gamma}$	−0.67	−0.37	−0.11	0.10	0.07	0.28	0.64	0.98	0.63	0.96	28.58
p	(0.28)	(0.22)	(0.10)	(0.07)	(0.04)	(0.15)	(0.26)	(0.31)	(0.55)	(0.27)	(0.20)
Effect of industry rotation at 6-month horizon											
$\hat{\gamma}$	−0.36	−0.21	−0.03	0.17	0.18	0.35	0.58	0.82	0.44	0.27	1.89
p	(0.27)	(0.19)	(0.12)	(0.04)	(0.02)	(0.07)	(0.12)	(0.18)	(0.49)	(0.80)	(0.48)
Effect of intra-industry allocation at 6-month horizon											
$\hat{\gamma}$	−0.54	−0.32	−0.09	0.10	0.06	0.27	0.55	0.88	0.50	1.08	12.89
p	(0.29)	(0.21)	(0.11)	(0.06)	(0.03)	(0.07)	(0.15)	(0.18)	(0.40)	(0.25)	(0.61)

the manager of Fidelity Contrafund provided the following remarks on the outlook of the market and his investment strategy[8]:

> I believe the economy will begin to slow. It's hard to predict when, but with higher oil prices, the Federal Reserve Board raising interest rates, high consumer debt and a possible slowing in the appreciation of real estate, economic growth should slip … [A] sluggish economic backdrop often favors less-cyclical, steadier blue-chip growth stocks … I have confidence in our ability to find those companies.

Less cyclical, steadier blue chip stocks typically have lower betas, and tilting toward these stocks due to pessimistic market views is a textbook example of market timing.

[8]"Fund Talk: The Manager's Overview," Shareholder Update and Semiannual Report, Fidelity Contrafund, June 30, 2005.

Interestingly, many fund managers seem to be more explicit when discussing stock selection strategies, but less so when discussing market timing strategies. The above example shows that certain strategies pursued by mutual funds, such as industry rotation, may have the actual effect of market timing, while not explicitly referred to as such.

6. Conclusion

Using traditional return-based timing measures, existing studies find insignificant and sometimes even negative market timing performance by mutual funds. These results seem to suggest that mutual fund managers do not exploit the predictability of market returns documented in the economics and finance literature. In this paper, we propose new measures of market timing performance for mutual funds using information on fund portfolio holdings. Compared to traditional timing measures based on realized fund returns, our proposed market timing measures are not subject to artificial timing bias and have better statistical power.

Using the holdings-based tests, we find that, on average, actively managed U.S. domestic equity funds possess positive timing ability. There are also some funds with strong timing skills, and the proportion of these funds in our sample exceeds what one would expect if no fund has any timing ability. Furthermore, the tests based on fund beta changes show that mutual funds time the market through active trading.

Additional evidence suggests that fund managers not only adjust the market exposure of fund portfolios in response to macroeconomic conditions, but also use private information to time the market. Active market timers tend to have high industry concentration in their portfolios and, to a lesser extent, large fund size and an investment style tilting toward small-cap stocks. We also document that industry allocation plays an important role in mutual fund market timing activities, and that mutual funds shift industry weights in response to macroeconomic information.

The findings of positive market ability by mutual funds have potentially important implications. We show that the positive market ability of mutual funds is of both statistical and economic significance. Market timing as an investment strategy may deserve more attention when evaluating the active ability of mutual funds.

Appendix A. Market timing tests on passive portfolios

The passive characteristics portfolios in our analysis are constructed following Daniel, Grinblatt, Titman, and Wermers (1997). First, at the end of every month during our sample period, we sort all stocks into five groups according to market capitalization using the NYSE breakpoints. Second, within each of the five size-sorted groups, we sort stocks into five subgroups according to the book-to-market ratio (BM). The market value of a firm is measured at the end of each month, and the book value of equity is based on the most recently reported fiscal year. Finally, within each of the 25 size-and-BM-sorted subgroups, we further sort stocks into five subgroups based on momentum, which is measured by a stock's total return during the previous six months. Within each subgroup, we form equally weighted portfolios that are rebalanced in the end of each month. This results in a total of 125 passive characteristics portfolios.

We perform both the return-based and holdings-based market timing tests on these 125 passive portfolios using the value-weighted CRSP index return as the market return.

Table A1
Artificial timing bias of return-based versus holdings-based timing tests

		5%	10%	25%	Mean	Median	75%	90%	95%	St dev	Skew	Kurto
Panel A: Return-based tests (1-month horizon)												
Treynor	$\hat{\gamma}$	−3.26	−3.07	−2.33	−1.23	−1.14	−0.16	0.61	0.72	1.32	0.14	−0.77
-Mazuy	p	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.98)	(0.18)	(0.21)	(0.00)	(0.46)	(1.00)
	t	−6.50	−5.39	−3.97	−2.24	−2.15	−0.31	1.16	1.36	2.55	0.03	−0.38
	p	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.93)	(0.53)	(0.62)	(0.00)	(0.42)	(0.84)
Panel B: Holdings-based tests (3-month horizon)												
Treynor	$\hat{\gamma}$	−0.29	−0.22	−0.06	0.13	0.16	0.32	0.45	0.58	0.28	−0.01	0.15
-Mazuy	p	(0.16)	(0.20)	(0.17)	(0.18)	(0.12)	(0.20)	(0.31)	(0.28)	(0.57)	(0.50)	(0.48)
	t	−0.68	−0.50	−0.14	0.48	0.42	1.03	1.62	2.15	0.86	0.42	0.09
	p	(0.05)	(0.08)	(0.14)	(0.16)	(0.18)	(0.22)	(0.21)	(0.13)	(0.60)	(0.07)	(0.33)

The return-based tests are based on monthly returns, and the holdings-based tests are based on portfolio betas measured at the beginning of each quarter and market returns over the quarter. Stock betas used for calculating portfolio betas are estimated using the past one-year daily returns (see Section 4.2). Statistical inference is based on the bootstrapped p-values of the timing coefficient γ and its Newey-West t-statistics, following the procedure described in Appendix C with 2,000 replications. The results are reported in Table A1. The high p-values (close or equal to one) for the return-based timing coefficients, according to the definition in Eq. (C.1) of Appendix C, suggest that the empirical estimate of market timing is consistently smaller than the bootstrapped timing coefficients, thus evidence of significant negative timing.

We further plot the return-based and holdings-based Treynor-Mazuy timing coefficients of the 125 passive portfolios against the portfolio characteristics in Figs. A1 and A2. The 25 panels represent size-and-BM-sorted portfolios indexed by (i,j), where i denotes the size ranking from top (small) to bottom (large), and j the BM ranking from left (low) to right (high). Within each panel, or the size-and-BM-sorted portfolio, we plot the timing coefficients of five momentum portfolios (sorted from low to high on momentum). The x axis in each panel denotes the momentum ranking. The results are summarized in Section 2.3.1.

Appendix B. Statistical power of market timing tests

To compare the statistical power of the return-based and holdings-based timing measures, we perform the following simulations. Daily returns of N stocks are simulated using the process

$$r_{it} - r_f = \beta_i(r_{mt} - r_f) + \varepsilon_{it}, \tag{B.1}$$

where $\beta_i \sim \mathrm{U}(0,2)$, $e_{it} \sim \mathrm{N}(0,\sigma_\varepsilon^2)$, $r_{mt} - r_f$ (the market excess return) $\sim \mathrm{N}(u_m,\sigma_m^2)$, and ε_{it} is i.i.d. over time and across stocks. Assuming 22 trading days in a month, we aggregate stock and market returns at both monthly and quarterly frequencies.

We assume that a fund manager has perfect knowledge of a stock's beta and divides stocks into a high-beta group (H) and a low-beta group (L), each with $N/2$ stocks. She calculates the average betas of the two groups: $\beta^H = \sum_{i\in H}\beta_i/(2N)$ and $\beta^L = \sum_{i\in L}\beta_i/(2N)$.

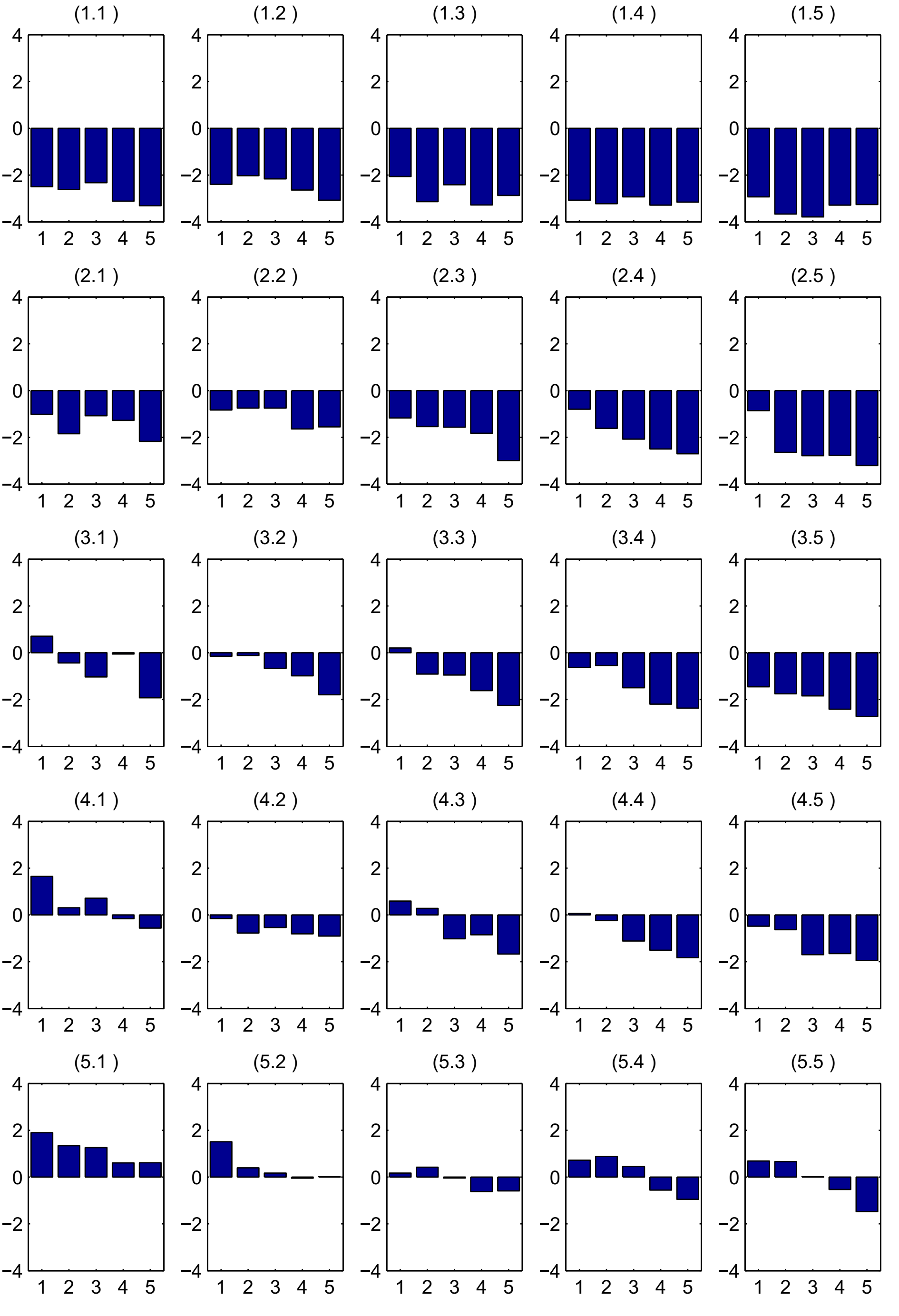

Fig. A1. Return-based timing coefficients of passive portfolios. This figure plots the return-based timing coefficients of the 125 passive portfolios. The panels represent 25 size-and-BM-sorted portfolios, where the five rows are sorted from small to large according to size and the five columns are further sorted from low to high according to the book-to-market ratio. Within each panel, the *x* axis represents five portfolios sorted from low to high according to momentum.

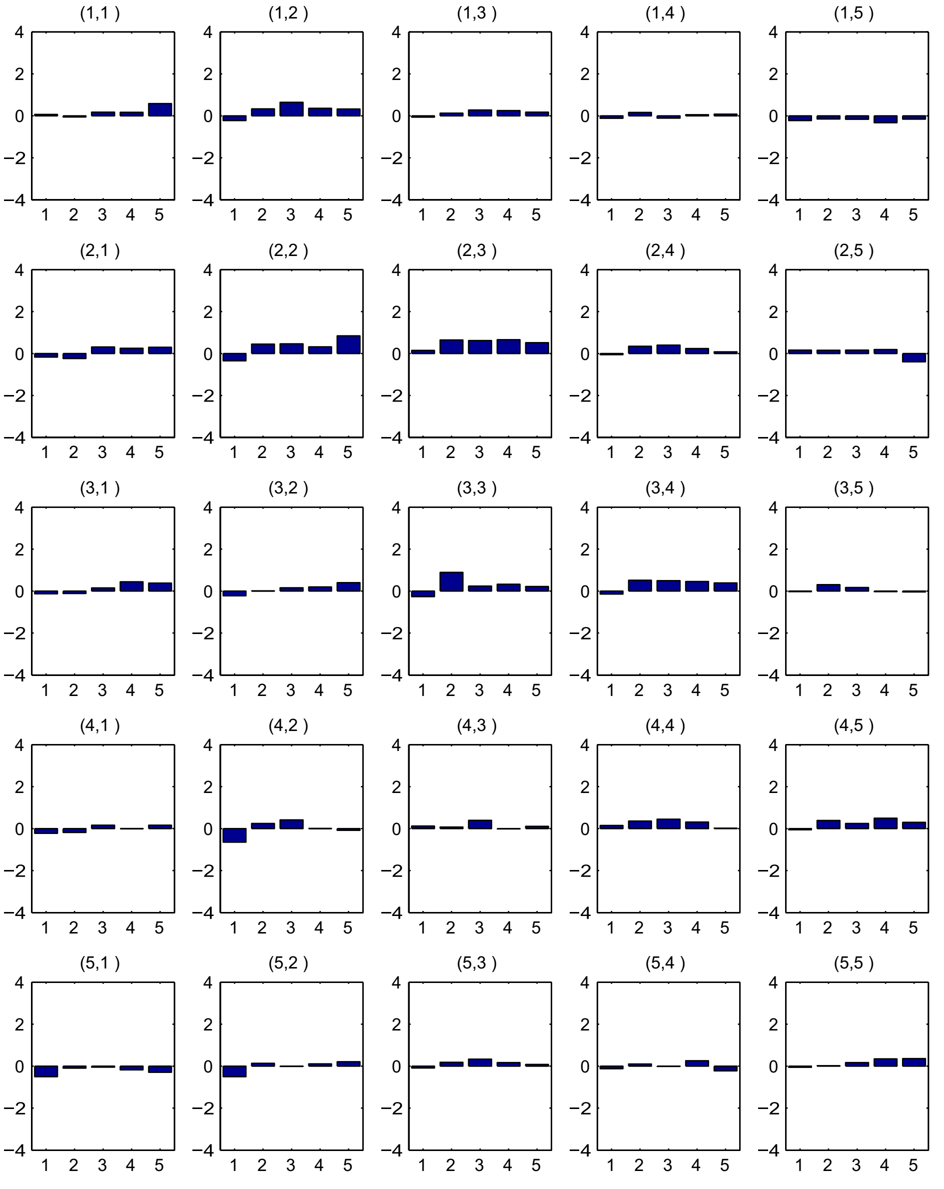

Fig. A2. Holdings-based timing coefficients of passive portfolios. This figure plots the holdings-based timing coefficients of the 125 passive portfolios. The panels represent 25 size-and-BM-sorted portfolios, where the five rows are sorted from small to large according to size and the five columns are further sorted from low to high according to the book-to-market ratio. Within each panel, the x axis represents five portfolios sorted from low to high according to momentum.

At the beginning of each quarter, the fund manager receives a signal about the market returns in the next quarter:

$$S_t = r_{m,t+1} - r_f + e_{s,t+1}, \tag{B.2}$$

where $e_{s,t+1} \sim \mathrm{N}(0, \sigma_s^2)$. To take advantage of the timing signal S_t, the fund manager adjusts the portfolio weights ω_{it} in the following way: the weight for each stock in the high-beta group is set to $\omega_{Ht} = 1/N + 2\gamma(S_t - u_m)/(\beta^H - \beta^L)/N$, while the weight for each stock in the low-beta group is set to $\omega_{Lt} = 1/N - 2\gamma(S_t - u_m)/(\beta^H - \beta^L)/N$. There is no active trading during the quarter. The fund beta at the beginning of the quarter is $\beta_t = \bar{\beta} + \gamma(S_t - u_m)$, where $\bar{\beta} = \sum_{i=1}^{N} \beta_i / N$ is the average beta of the N stocks.

In the above setup, σ_s measures the accuracy of the fund manager's market timing information, and γ measures the aggressiveness of the fund manager's timing strategy. The fund return is $r_{t+1} - r_f = \beta_t(r_{m,t+1} - r_f) + \sum_{i \in H} \omega_{Ht} \varepsilon_{i,t+1} + \sum_{i \in L} \omega_{Lt} \varepsilon_{i,t+1} = (\bar{\beta} - \gamma u_m)(r_{m,t+1} - r_f) + \gamma(r_{m,t+1} - r_f)^2 + \gamma e_{s,t+1}(r_{mt} - r_f) + e_{t+1}$, where $e_{t+1} = \sum_{i \in H} \omega_{Ht} \varepsilon_{i,t+1} + \sum_{i \in L} \omega_{Lt} \varepsilon_{i,t+1}$.

For simplicity, we set the risk-free rate, r_f, to zero, and the expected market return, u_m, to 0.5%/22. We set the number of stocks, N, to 60, approximately the median number of stocks held by funds in our sample (from Table 1). The daily standard deviation of the idiosyncratic stock return ε_{it} is $\sigma_e = 0.17/\sqrt{22}$, where 0.17 is approximately the average standard deviation of the residuals of the monthly market model for all CRSP stocks. The daily standard deviation of the market return is $\sigma_m = 0.045/\sqrt{22}$, close to that of the CRSP value-weighted index returns. The total number of monthly stock return and fund return observations is set to $T = 132 + 60$, where 132 is the mean number of monthly return observations for a typical fund in our sample, and the additional 60 months of stock returns are for the purpose of estimating stock betas. We vary two parameters in the simulation. First, the noise-to-signal ratio of S_t, i.e., σ_s/σ_m, takes values of 0.5, 1, 2, and 4. Second, the timing measure γ is set variously to 0.30, 0.70, and 1.60, which are close to the median, 75th, and 95th percentiles of the holdings-based Treynor-Mazuy timing measures at the three-month forecasting horizon (see Table 3).

In each simulation, we perform both the return-based and holdings-based Treynor-Mazuy tests. To be consistent with our data structure, the return-based tests (RB) are based on 132 monthly fund returns (from month 61 to month 182), while the holdings-based tests are based on 27 quarterly fund betas (from quarter 21 to quarter 47). Stock betas are estimated using the past one-year daily returns (HB1), the past three-month daily returns (HB2), and the past five-year monthly returns (HB3), respectively (see Sections 4.2 and 4.3). Table B1 reports the means and standard deviations of both measures with

Table B1
Statistical power of return-based versus holdings-based measures

$\frac{\sigma_s}{\sigma_m}$	$\gamma = 0.30$				$\gamma = 0.70$				$\gamma = 1.60$			
	RB	HB1	HB2	HB3	RB	HB1	HB2	HB3	RB	HB1	HB2	HB3
0.50	0.300	0.300	0.300	0.302	0.691	0.700	0.700	0.701	1.575	1.600	1.602	1.601
	(0.71)	(0.08)	(0.12)	(0.16)	(0.73)	(0.09)	(0.14)	(0.17)	(0.83)	(0.13)	(0.22)	(0.20)
1.00	0.300	0.301	0.301	0.302	0.690	0.700	0.700	0.701	1.577	1.600	1.602	1.601
	(0.71)	(0.08)	(0.12)	(0.16)	(0.74)	(0.12)	(0.16)	(0.18)	(0.86)	(0.21)	(0.28)	(0.26)
2.00	0.300	0.301	0.301	0.303	0.695	0.701	0.701	0.702	1.578	1.597	1.600	1.598
	(0.71)	(0.10)	(0.14)	(0.17)	(0.75)	(0.18)	(0.21)	(0.23)	(0.95)	(0.39)	(0.43)	(0.42)
4.00	0.302	0.301	0.301	0.303	0.692	0.699	0.700	0.701	1.582	1.608	1.612	1.610
	(0.73)	(0.16)	(0.19)	(0.22)	(0.83)	(0.33)	(0.35)	(0.36)	(1.25)	(0.76)	(0.79)	(0.78)

10,000 replications. The results confirm that the holdings-based tests have more statistical power than their return-based counterparts. For example, when $\gamma = 0.30$ and $\sigma_s/\sigma_m = 0.25$, the standard deviation of the return-based measure is 0.71, while those of the three holdings-based measures are 0.08, 0.12, and 0.16, respectively.

Appendix C. Bootstrapping procedures

C.1. Bootstrapping procedure for holdings-based tests

This procedure is used for holdings-based market timing tests, including the analysis of fund characteristics in Section 5.1 and industry allocation in Section 5.3. Let $\mathbf{b}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \ldots, \hat{\beta}_{Kt})$ denote the beta estimates for K funds at the beginning of period $t+1$ and $r_{m,t+1}$ the market return over the period $t+1$. Let $\Gamma(\mathbf{b}_t, r_{m,t+1})$ be a cross-sectional statistic (say, the 90th percentile) of the holdings-based measure ($\hat{\gamma}$ or its t-statistic). The bootstrapping procedure is to obtain the distribution of a particular cross-sectional statistic $\Gamma^*_j(\mathbf{b}_t, r^*_{m,t+1})$ under the null hypothesis that there is no timing ability, and then compare it to $\Gamma(\mathbf{b}_t, r_{m,t+1})$. To ensure no timing ability, in each bootstrap we keep $\mathbf{b}_t$ unchanged but randomly sample the market returns $r^*_{m,t+1}$. With $\mathbf{b}_t$ and $r^*_{m,t+1}$ we estimate the holdings-based measure for all funds and calculate the bootstrapped cross-sectional statistic $\Gamma^*_j(\mathbf{b}_t, r^*_{m,t+1})$. This procedure is repeated a large number (J) of times to obtain a series $\Gamma^*_j(\mathbf{b}_t, r^*_{m,t+1})$, $j = 1, \ldots, J$. Since the procedure preserves the covariance structure of $\mathbf{b}_t$ with correlated fund betas, the distribution of $\Gamma^*_j(\mathbf{b}_t, r^*_{m,t+1})$ approximates the distribution of $\Gamma(\mathbf{b}_t, r_{m,t+1})$ under the null hypothesis of no timing ability. The bootstrapped p-value of $\Gamma(\mathbf{b}_t, r_{m,t+1})$ is computed as

$$p = \frac{1}{J}\sum_{j=1}^{J} I_{\Gamma^*_j > \Gamma}, \tag{C.1}$$

where $I_{\Gamma^*_j > \Gamma}$ is an indicator that takes the value one if $\Gamma^*_j(\mathbf{b}_t, r^*_{m,t+1}) > \Gamma(\mathbf{b}_t, r_{m,t+1})$ and zero otherwise. J is set to 2,000 in our empirical analysis. A low value of p (close to zero) implies that the estimated timing measure is consistently higher than its bootstrapped values and thus evidence of positive timing ability, while a high value of p (close to one) implies that the estimated timing measure is consistently lower than its bootstrapped values and thus evidence of negative timing ability.

C.2. Bootstrapping procedure for return-based tests

For the return-based timing tests, we adopt a parametric bootstrapping procedure. First, we estimate the following market return model for each fund:

$$r_t = \alpha + \beta_0 r_{mt} + e_t, \tag{C.2}$$

We retain the OLS estimates of the parameters $\hat{\alpha}$ and $\hat{\beta}_0$, as well as the residuals $\hat{e}_t$ for all funds. To ensure no timing ability, the bootstrapped fund returns $\mathbf{r}^*_t$ are generated from the randomly sampled market returns r^*_{mt} based on the estimated model (C.2) using $\hat{\alpha}$ and $\hat{\beta}$ while keeping $\hat{e}_t$ fixed. With bootstrapped fund returns, we estimate the return-based timing measures for all funds and calculate the cross-sectional statistic $\Gamma^*_j(\mathbf{r}^*_t, r^*_{mt})$. This procedure preserves the covariance structure of fund returns under the null hypothesis of no timing ability. The bootstrapped p-values of the cross-sectional statistics are computed following Eq. (C.1).

C.3. Bootstrapping procedure for holdings-based tests controlling for public information

This procedure is used for holdings-based market timing tests in (20) and (21) with control of public information, of which the tests in (18) and (19) are special cases. We first estimate the following models using OLS based on monthly series of market returns and macroeconomic variables:

$$M_{jt} = a_j + b_j M_{jt-1} + e_{jt}, \quad j = 1, \ldots, 5, \tag{C.3}$$

$$r_{mt} = a_r + \sum_{j=1}^{5} b_{rj} M_{jt-1} + e_{rt}, \tag{C.4}$$

where M_{jt} denotes the jth macroeconomic variable. We retain the parameter estimates $\hat{a}_j$, $\hat{b}_j$, $\hat{a}_r$, and $\hat{b}_{rj}$, as well as the residuals $\hat{e}_{jt}$ and $\hat{e}_{rt}$. In each bootstrap, the economic variables M^*_{jt} and market return r^*_{mt} are generated based on the estimated models (C.3) and (C.4) by jointly resampling from the residuals $\hat{e}_{rt}$ and $\hat{e}_{jt-1}$, with $j = 1, \ldots, 5$. With the bootstrapped macroeconomic variables and market returns, we perform holdings-based tests for each fund. The bootstrapped p-values are obtained via Eq. (C.1).

References

Avramov, D., Chordia, T., 2006. Asset pricing models and financial market anomalies. Review of Financial Studies 19 (3), 1001–1040.

Avramov, D., Wermers, R., 2006. Investing in mutual funds when returns are predictable. Journal of Financial Economics 81, 339–377.

Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: some utility cost calculations. Journal of Financial Economics 52, 47–78.

Barberis, N.C., 2000. Investing for the long run when returns are predictable. Journal of Finance 55, 225–264.

Becker, C., Ferson, W., Myers, D.H., Schill, M.J., 1999. Conditional market timing with benchmark investors. Journal of Financial Economics 52, 119–148.

Blake, D., Lehmann, B., Timmermann, A., 1999. Asset allocation dynamics and pension fund performance. Journal of Business 72, 429–461.

Bollen, N., Busse, J.A., 2001. On the timing ability of mutual fund managers. Journal of Finance 56, 1075–1094.

Busse, J., Irvine, P., 2006. Bayesian alphas and mutual fund performance. Journal of Finance 61 (5), 2251–2288.

Campbell, J.Y., 1987. Stock returns and the term structure. Journal of Financial Economics 18, 373–399.

Campbell, J.Y., Chan, Y.L., Viceira, L.M., 2003. A multivariate model of strategic asset allocation. Journal of Financial Economics 67, 41–80.

Campbell, J.Y., Shiller, R.J., 1988a. The dividend–price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Campbell, J.Y., Shiller, R.J., 1988b. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.

Campbell, J.Y., Viceira, L.M., 1999. Consumption and portfolio decisions when expected returns are time varying. Quarterly Journal of Economics 114, 433–495.

Chang, E.C., Lewellen, W.G., 1984. Market timing and mutual fund investment performance. Journal of Business 57, 57–72.

Chen, J., Hong, H., Huang, M., Kubik, J., 2004. Does fund size erode mutual fund performance? the role of liquidity and organization. American Economic Review 94, 1276–1302.

Cochrane, J., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance 46, 207–234.
Coggin, T.D., Fabozzi, F.J., Rahman, S., 1993. The investment performance of U.S. equity pension fund managers. Journal of Finance 48, 1039–1056.
Cohen, R., Coval, J., Pastor, L., 2005. Judging fund managers by the company they keep. Journal of Finance 60, 1057–1096.
Cumby, R., Glen, J., 1990. Evaluating the performance of international mutual funds. Journal of Finance 45, 497–521.
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.
Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7, 197–226.
Edelen, R.M., 1999. Investor flows and the assessed performance of open-end mutual funds. Journal of Financial Economics 53, 439–466.
Fama, E.F., French, K.R., 1987. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.
Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.
Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–49.
Fama, E.F., Schwert, W., 1977. Asset returns and inflation. Journal of Financial Economics 5, 115–146.
Ferson, W.E., 1989. Changes in expected security returns, risk and the level of interest rates. Journal of Finance 44, 1191–1217.
Ferson, W.E., Khang, K., 2002. Conditional performance measurement using portfolio weights: evidence for pension funds. Journal of Financial Economics 65, 249–282.
Ferson, W.E., Schadt, R., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51, 425–462.
French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.
Fung, H., Xu, E., Yau, J., 2002. Global hedge funds: risk, return, and market timing. Financial Analyst Journal 58, 19–30.
Glosten, L., Jagannathan, R., 1994. A contingent claims approach to performance evaluation. Journal of Empirical Finance 1, 133–166.
Graham, J., Harvey, C., 1996. Market timing ability and volatility implied in investment newsletters asset allocation recommendations. Journal of Financial Economics 42, 397–421.
Grinblatt, M., Titman, S., 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 393–416.
Grinblatt, M., Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. Journal of Business 66, 47–68.
Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies portfolio performance and herding: a study of mutual fund behavior. American Economic Review 85, 1088–1105.
Goetzmann, W.N., Ingersoll Jr., J., Ivkovich, Z., 2000. Monthly measurement of daily timers. Journal of Financial and Quantitative Analysis 35, 257–290.
Hall, P., 1992. The Bootstrap and Edgeworth Expansion. Springer, New York.
Henriksson, R.D., 1984. Market timing and mutual fund performance: an empirical investigation. Journal of Business 57, 73–96.
Henriksson, R.D., Merton, R.C., 1981. On market timing and investment performance II: statistical procedures for evaluating forecasting skills. Journal of Business 54, 513–534.
Jagannathan, R., Korajczyk, R., 1986. Assessing the market timing performance of managed portfolios. Journal of Business 59, 217–236.
Jiang, W., 2003. A nonparametric test of market timing. Journal of Empirical Finance 10, 399–425.
Kacperczyk, M., Sialm, C., Zheng, L., 2005. On the industry concentration of actively managed equity mutual funds. Journal of Finance 60, 1983–2011.
Kandel, S., Stambaugh, R., 1996. On the predictability of stock returns: an asset-allocation perspective. Journal of Finance 51, 385–424.
Keim, D., Stambaugh, R., 1986. Predicting returns in stock and bond markets. Journal of Financial Economics 17, 357–390.

Kon, S.J., 1983. The market timing performance of mutual fund managers. Journal of Business 56, 323–347.
Kosowski, R., Timmermann, A., White, H., Wermers, R., 2005. Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis. Journal of Finance, forthcoming.
Kryzanowski, L., Lalancette, S., To, M., 1996. Performance attribution using an APT with prespecified factors, Unpublished working paper. Concordia University and University of British Columbia.
Lamont, O., 1998. Earnings and expected returns. Journal of Finance 53, 1563–1587.
Lewellen, J., 1999. The time-series relations among expected return, risk, and book-to-market. Journal of Financial Economics 54, 5–43.
Lo, A., MacKinlay, C., 1990. Data snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–468.
Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. Journal of Financial Economics 8, 323–361.
Merton, R.C., 1981. On market timing and investment performance part I: an equilibrium theory of value for market forecasts. Journal of Business 54, 363–406.
Newey, W., West, K., 1987. A simple positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Pontiff, J., Schall, L., 1998. Book-to-market ratios as predictors of market returns. Journal of Financial Economics 49, 141–160.
Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.
Sirri, E., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.
Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping technical trading rules and the bootstrap. Journal of Finance 54, 1647–1692.
Treynor, J., Mazuy, K., 1966. Can mutual funds outguess the market? Harvard Business Review 44, 131–136.
Warther, V., 1995. Aggregate mutual fund flows and security returns. Journal of Financial Economics 39, 209–235.
Wermers, R., 1999. Mutual fund herding and the impact on stock prices. Journal of Finance 54, 581–622.
Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transaction costs, and expenses. Journal of Finance 55, 1655–1703.
Wermers, R., 2004. Is money really ''smart''? New evidence on the relation between mutual fund flows manager behavior, and performance persistence. Working Paper, University of Maryland.