Do Mutual Funds Time the Market?
Evidence from Portfolio Holdings

George J. Jiang, Tong Yao, and Tong Yu [1]

August 28, 2004

[1]Jiang and Yao are from the Department of Finance, Eller College of Management, University of Arizona, Tucson, Arizona 85721-0108. Emails: gjiang@eller.arizona.edu and tyao@eller.arizona.edu. Yu is from College of Business Administration, University of Rhode Island. Email: tongyu@uri.edu. We would like to thank Yong Chen, Kathy Kahle, Nathan O'Connor and workshop participants in the Finance Department at the University of Arizona for helpful comments and suggestions.

Do Mutual Funds Time the Market?
Evidence From Portfolio Holdings

Abstract

Existing literature has found no evidence of market-timing ability by mutual funds using tests based on fund returns. This paper proposes alternative market-timing tests based on observed fund holdings. The holdings-based measures are shown to be more powerful than the return-based measures, and are not subject to “artificial timing” bias. Applying the holdings-based tests, we find strong evidence of mutual fund timing ability. Our findings also suggest that market-timing funds tend to have higher returns and trade more actively. Furthermore, they seem to have market-timing information beyond those common return-predictive economic variables documented in the academic studies. Finally, we quantify the potential economic value of market-timing as a contingent claim. The magnitude of the estimated values indicates that market-timing is potentially an important investment strategy deserving more academic attention.

Two strands of finance literature stand in curious contrast to each other. On the one hand, many studies have argued that aggregate stock market returns are predictable and that such predictability should have a significant impact on an investor's optimal asset allocation.[1] On the other hand, there is scant evidence that investors take advantage of such predictability in their portfolio decisions. In particular, earlier studies dating back to Treynor and Mazuy (1966) and Henriksson and Merton (1981), and recent studies such as Becker, Ferson, Myers, and Schill (1999) and Jiang (2003), all find that mutual funds on average do not exhibit significant market-timing ability.[2] Fund managers are perceived as more skillful and better informed investors. If they cannot explore market return predictability, it is unlikely that anyone else could.[3]

The market-timing measures in existing studies are mainly based on fund returns. In this paper, we argue that these measures may not be powerful enough to detect market-timing ability for the following reasons. First, existing timing tests, such as those of Henriksson and Merton (1981) and Treynor and Mazuy (1966) (hereafter "return-based measures"), are based on nonlinear relations between fund returns and market returns. Fund returns are volatile with observations available often at low frequency and over a relatively short time period. This limits the statistical power of such tests. Second, as Jagannathan and Korajczyk (1986) point out, the return-based measures may lead to incorrect inference when fund managers engage in dynamic trading or invest in securities with option-like features, a phenomenon known as the "artificial timing" bias.

To get around these problems, we use data on portfolio holdings to examine mutual fund market-timing

[1]For evidence of market return predictability, see, for example, Campbell (1987), Campbell and Schiller (1988a, 1988b), 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 how such predictability should affect an investor's asset allocation decisions, see, for example, Kendal and Stambaugh (1996), Barberis (2000), Balduzzi and Lynch (1999), Campbell and Viceira (1999), and Campbell, Chan, and Viceira (2003).

[2]In addition, Graham and Harvey (1996), Aragon (2003), Fung, Xu, and Yau (2002), and Blake, Lehmann, and Timmermann (1999) find no evidence of market-timing ability by investment newsletters, hedge funds, or pension funds. The only exception, to our knowledge, is Bollen and Busse (2001) who find positive timing ability for a sample of 230 domestic equity funds at daily frequency.

[3]One may argue that institutional constraints on mutual funds, such as no-short-sale, might prevent them from timing the market. However, market-timing can be achieved by adjusting portfolio holdings and the cash component without short-selling any securities.

activities. We estimate the beta of a fund as the weighted average of betas of individual stocks held by the fund, and then measure the relations between fund betas and market returns (hereafter "holdings-based measure"). Several studies, including Grinblatt and Titman (1989, 1994), Grinblatt, Titman, and Wermers (1995), Daniel, Grinblatt, Titman, and Wermers (1997), Wermers (2000), and Ferson and Khang (2002), have used information on portfolio holdings to evaluate mutual fund performance, in particular stock-selection ability. There are several empirical advantages when the holdings data is used to evaluate market timing. First, the tests based on portfolio holdings are more powerful than those based on fund returns. When estimating stock betas, we can use stock returns that typically are available at a higher frequency and over a longer period than for fund returns. Even though individual stock betas are still measured with errors, the estimated fund beta, as a weighted average over betas of a large number of stocks, is of much higher accuracy. Second, this approach avoids the "artificial timing" bias. "Artificial timing" refers to the scenario where fund managers change fund betas in response to past or contemporaneous market returns, while true market timing refers to the case where fund managers adjust fund betas according to information about future market returns. Since the fund beta is measured at the beginning of a holding period in our study, its covariance with the holding-period market return is not subject to "artificial timing".

We compute both return-based and holdings-based timing measures for a sample of 2,294 mutual funds over the period from 1980 to 2002. These funds have information on portfolio holdings from Thomson Financial and information on fund returns from CRSP. Consistent with the previous literature, the return-based tests indicate that on average mutual funds exhibit slightly negative, although not statistically significant, timing ability. The holdings-based tests, however, have very different results. With 1-month forecasting horizon, the average timing measures are positive but not significant. When we extend the forecasting horizons to 3-, 6-, and 12-month, the average timing measures become significantly positive. Furthermore, the proportion of funds with strongly positive timing measures is substantially higher than what one would expect from a sample of funds without timing ability. To take into account cross-fund correlations of timing measures, the statistical inference is based on a bootstrapping procedure which also addresses the final sample issue. Different methods of estimating stock betas and fund betas are employed and the results are robust.

The evidence of significant timing ability is consistent across funds with different investment objectives. However, the aggressive growth funds tend to be the most active market timers, followed by growth funds, growth and income funds, with balanced funds the least active. We also find that market timing is associated with certain fund characteristics. Active market timers tend to have higher returns. In addition, market timers trade more actively, as measured by portfolio turnovers. Funds with no loads or lower loads also tend to have better timing ability.

Given the evidence on market return predictability documented in academic studies, we explore whether the market-timing activity of mutual funds is based on several known return-predictive economic variables. We find that fund betas are positively correlated with the aggregate dividend yield and the aggregate earnings-to-price ratio, but there is no significant evidence that fund managers condition fund betas on other economic signals such as the short-term interest rate, term spread or credit spread. More importantly, even after controlling for all those economic variables, the evidence of mutual fund market-timing ability remains very significant. This suggests that mutual fund timing decisions may be based on private information beyond the documented economic variables. Identifying such conditional information remains an interesting and challenging research project.

With the strong evidence of positive timing ability for mutual funds, we further calibrate the potential economic value of the market-timing strategy. Following Glosten and Jagannathan (1994), we quantify market timing as a contingent claim on the market index. We find that the market-timing ability may contribute to annual fund returns by as high as 1.23% for the median fund. The estimated values are likely not attained by existing funds as we assume that funds continuously engage in market timing and there are no transaction costs. The magnitude of the estimated value, nevertheless, indicates the potential importance of mutual fund market-timing activities.

The remainder of the paper is organized as follows. Section I introduces both the return-based and holdings-based market-timing measures. Section II describes the data used in our analysis and the methodology for statistical inference. Section III reports the empirical results and performs robustness checks. Section IV further discusses the characteristics of market timers, the conditioning information and the potential economic value of market timing. Section V concludes.

I. Measures of Market Timing

A. Return-based Measures

Let r_t and r_{mt} denote the fund return and market return at time t. The Treynor-Mazuy market-timing measure (see Treynor and Mazuy (1966)) is the estimated coefficient γ from the following 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. These two measures are based on the dynamic relations between the time-varying fund beta β_t and the market return r_{mt}:[4]

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

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

where $I_{r_{mt}>0}$ is an indicator with value 1 when $r_{mt} > 0$ and 0 otherwise. A significantly positive γ indicates positive timing ability by mutual funds.

B. Holdings-based Measures

The return-based measures are useful when fund betas are not directly observed. However, when a fund's portfolio holdings are observed, we can estimate the fund beta and perform tests based directly on (3) and

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

(4). To be specific, we estimate the fund beta as the weighted average of the beta estimates for stocks (and other securities) held by the fund:

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

where ω_{it} is the portfolio weight for stock i at the beginning of period t, and $\hat{b}_{it}$ is the estimated beta for stock i at time t. As long as the stock beta estimates are unbiased, the fund beta is also unbiased. There are several ways to estimate stock betas, which 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 following regressions:

$$\hat{\beta}_t = \alpha + \gamma r_{mt} + \eta_t \tag{6}$$

$$\hat{\beta}_t = \alpha + \gamma I_{r_{mt}>0} + \eta_t \tag{7}$$

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.

C. Return-based Measures versus Holdings-based Measures

The main difference between the return-based measures and holdings-based measures hinges on their statistical properties and consequently the power of statistical inference. Suppose there are intra-period fund returns in each period (e.g. daily fund returns during each month), estimating the fund beta β_t each month first and then using (6) or (7) to estimate γ is equivalent to substituting the β_t equation into a fund-return generating process first and then estimating γ using (1) or (2). This result is derived in Amemiya (1978) under a general setting of random coefficient models. In practice, however, the fund returns are often observed at a low frequency, such as monthly or quarterly, and thus are not sufficient for the estimation of time-varying fund betas. The advantage associated with the holdings-based measure is that with portfolio holdings, fund betas can be estimated directly from stock betas, whereas stock betas can be estimated using higher frequency of stock returns (e.g. daily returns). This typically leads to fund beta estimates with smaller standard errors. Relative to the conventional return-based measures in (1) or (2) which are often implemented using the quarterly or monthly fund returns, the holdings-based measures using directly estimated fund betas are in general more efficient statistically.

Formally, to compare the statistical properties we start with the benchmark in an ideal case where β_t is known. Let X be the vector of r_{mt} and ι a vector of ones. The OLS estimator of γ based on model (3) is:

$$\hat{\gamma}^* = [(X - \bar{r}_m\iota)'(X - \bar{r}_m\iota)]^{-1}(X - \bar{r}_m\iota)'(B - \bar{\beta}\iota) \tag{8}$$

where B is the vector of β_t, $\bar{\beta}$ is the average of β_t, and $\bar{r}_m$ is the average of r_{mt}. The variance of the estimator is $Var(\sqrt{T}(\hat{\gamma}^* - \gamma)) = \sigma_\eta^2[(X - \bar{r}_m\iota)'(X - \bar{r}_m\iota)]^{-1}$.

Of course, β_t is unknown and has to be estimated in practice. When β_t is estimated, the holdings-based measures contain the estimation error of β_t. Suppose $\hat{\beta}_t = \beta_t + \xi_t$, where ξ_t is independent of η_t with variance σ_ξ^2. Using (6), the OLS estimator for γ (denoted by $\tilde{\gamma}$) is:

$$\tilde{\gamma} = [(X - \bar{r}_m\iota)'(X - \bar{r}_m\iota)]^{-1}(X - \bar{r}_m\iota)'(\hat{B} - \bar{\beta}_t\iota) \tag{9}$$

where $\hat{B}$ is the vector of $\hat{\beta}_t$, and $\bar{\beta}_t$ is the average of $\hat{\beta}_t$. The variance of the estimator is $Var(\sqrt{T}(\tilde{\gamma} - \gamma)) = (\sigma_\eta^2 + \sigma_\xi^2)[(X - \bar{r}_m\iota)'(X - \bar{r}_m\iota)]^{-1} > Var(\sqrt{T}(\hat{\gamma}^* - \gamma))$. The additional variance component is $[(X - \bar{r}_m\iota)'(X - \bar{r}_m\iota)]^{-1}\sigma_\xi^2$ which depends on the magnitude of the β_t estimation error.

For the return-based measures, let Y be the vector of r_{mt}^2, and Z be X augmented with a vector of ones (ι): $Z = [\iota, X]$, and finally, R the vector of r_t. The OLS estimator of γ under the Treynor-Mazuy measure in model (1) is given by:

$$\hat{\gamma} = [Y'M_1Y]^{-1}Y'M_1R \tag{10}$$

where $M_1 = I - Z(Z'Z)^{-1}Z'$, and I is an identity matrix. The variance of the estimator is $Var(\sqrt{T}(\hat{\gamma} - \gamma)) = \sigma_\epsilon^2(Y'M_1Y)^{-1} + \sigma_\eta^2(Y'M_1Y)^{-1}Y'M_1\text{diag}(X_1^2, \cdots, X_T^2)M_1Y(Y'M_1Y)^{-1}$, where σ_η^2 and σ_ϵ^2 are the variances of η_t and ϵ_t, respectively. It can be shown that $Var(\sqrt{T}(\hat{\gamma} - \gamma)) \geq \sigma_\epsilon^2[Y'M_1Y]^{-1} + \sigma_\eta^2[(X - \bar{r}_m)'(X - \bar{r}_m)]^{-1} > Var(\sqrt{T}(\hat{\gamma}^* - \gamma))$. In other words, the return-based measure $\hat{\gamma}$ is less powerful than that based on $\hat{\gamma}^*$. It also indicates that the introduction of the fund return generating process in the test is the main contributor to the power loss of $\hat{\gamma}$. The additional variance component of $\hat{\gamma}$, $[Y'M_1Y]^{-1}\sigma_\epsilon^2$, is directly determined by σ_ϵ.

Therefore, whether the return-based measures or the holdings-based measures are more powerful depends mainly on the relative magnitude of the fund return noise (ϵ_t) versus the beta estimation error (ξ_t). In practice, the magnitude of ξ_t is often much smaller than that of ϵ_t, and therefore the holdings-based measures are likely to be more efficient. The main reason is as follows. Fund returns tend to be volatile, and the observations are usually available at monthly frequency over a limited period of fund life. On the other hand, under the holdings-based measures the fund beta is estimated from individual stock betas. When estimating stock betas, we are not constrained by the length of fund life and can take advantage of the stock return data available at higher sampling frequencies (e.g. daily frequency). It is well documented that beta as a variance-covariance measure can be better estimated from data sampled at high frequency, see, for example, earlier studies by Merton (1980), and French, Schwert, and Stambaugh (1987), and recent development in Andersen, Bollerslev, Diebold, and Labys (2003).

We perform a simulation to compare the empirical properties of return-based and holdings-based timing measures. The details of the simulation design are included in Appendix A. In the simulation we take into account two realistic features of the data. First, the fund return data is available at the monthly frequency, while the portfolio holdings data is available at the quarterly frequency. Second, the portfolio holdings data exists for a shorter time period than the fund return data. The simulation results reported in Table A.I for the return-based versus holdings-based timing measures confirm our conjecture. The standard errors of the return-based measure based on (1) are several times larger than those of the holdings-based measure based on (6).

Another difference between the return-based measures and the holdings-based measures is that the former are subject to the "artificial timing" bias, as pointed out by Jagannathan and Korajczyk (1986), while the latter are not. To illustrate this, consider a fund manager without timing ability. Suppose the fund manager underwrites a call option on the market index and at the same time keeps the beta of the rest of the portfolio constant. Such a strategy may enhance fund returns when the market turns out to be less volatile than expected. The beta of the fund, however, is affected through the delta of the short call position (i.e., the sensitivity of the value of the position to changes in market index). Since the delta of a short call option decreases when the market goes up, there is a negative correlation between the fund beta and the market return. The regressions based on (1) and (2) would result in a negative γ and a misleading conclusion of "negative timing".

While we cannot infer that the findings of "negative timing" in previous empirical studies are due to the above derivatives strategy, in practice many funds do engage in various hedging strategies using either derivatives or derivatives-like portfolio insurances.[5] It is also well-known that derivative-type payoff patterns can also be generated via dynamic trading on the underlying securities. Some commonly used investment strategies, seemingly innocuous, can also generate a nonlinear relationship between fund returns and market returns. For example, a "positive-feedback" manager who increases market exposure after the market moves up has a positive γ based on (1) or (2). Similarly, a contrarian manager who "buys low and sells high" at the market level has a negative γ using the same regressions, regardless of whether he or she has any timing ability.

The holdings-based measures do not suffer from this problem. When computing the fund beta in (5), we use the beginning-of-period portfolio weights. Therefore the fund beta is measured at the beginning of a period. The regression tests whether the fund beta is correlated with the subsequent market returns instead of contemporaneous market returns. Any positive or negative covariance between fund beta and market returns clearly cannot be attributed to "artificial timing".[6]

There is also a third relevant issue. Fund returns are usually available at the monthly frequency. Goetzmann, Ingersoll, and Ivkovich (2000) show that tests based on monthly fund returns cannot detect daily or

[5] See, for example, Koski and Pontiff (1999) for the use of derivatives by mutual funds.

[6] In addition, the holdings-based measures are not biased by "window dressing" activities since "window dressing" is not based on future market returns.

weekly market-timing activities. On the other hand, for holdings-based measures, the forecasting horizon can be daily or weekly. As the market returns are available at these frequencies, it is possible to examine market-timing activities over shorter horizons, although this is not the focus of our paper.

II. Data and Methodology

A. The Data

We combine two mutual fund datasets in our analysis. The first is the fund holdings dataset from Thomson Financial (hereafter the "Thomson dataset"). Its predecessor is the CDA/Spectrum data used by a number of empirical studies, such as Daniel, Grinblatt, Titman, and Wermers (1997), Grinblatt and Titman (1989, 1993), and Wermers (2000, 2003). The Thomson dataset contains information on portfolio holdings for equity mutual funds investing in the US market starting from 1980. Most mutual funds in the Thomson dataset are US-based. However, there are also a number of foreign-based, mostly Canadian, funds. The data frequency is mostly quarterly, although for some funds and in early years only semi-annual portfolio information is available. The second dataset is the CRSP mutual fund dataset, which has information on monthly returns and total net assets for all US-based mutual funds, starting from 1962. We merge these two datasets to get information on both the stock holdings and returns for US domestic equity funds. The CRSP dataset is free of survivorship bias. We are not aware of any survivorship problem in the Thomson dataset either. As noted by Daniel, Grinblatt, Titman, and Wermers (1997), its predecessor, CDA/Spectrum, does not have a survivorship problem. However, funds are added to the Thomson dataset at a slower pace than to the CRSP dataset. Therefore, a recently started fund may be included in CRSP but not yet appears in Thomson.

In the Thomson dataset, funds are classified into nine categories according to their self-declared investment objectives. Since our focus is on US equity funds, we only include funds with one of the following four investment objectives: a) aggressive growth, b) growth, c) growth and income, and d) balanced. To make sure that the investment objective codes are accurate in the Thomson dataset, we screen fund names and check alternative sources such as Morningstar and Yahoo mutual fund websites as well as websites of fund management companies to determine the true investment objective of a fund and remove all index funds, foreign-based funds, US-based international funds, fixed-income funds, real estate funds, precious metal funds, and variable annuities from the sample.

We follow a few additional steps to further clean the Thomson dataset. Sometimes one fund has multiple fund identification numbers. This often occurs when a fund alternates using fund names with and without abbreviations when they report portfolio holdings. We manually identify such cases to combine the multiple identification numbers into one. In addition, when a fund changes its name, the identification number in the Thomson dataset can change, but the identification number in CRSP may remain the same. In this case we adjust the Thomson data according to the CRSP criteria. Finally, if a fund ceases to exist, its identification number may be re-assigned to another fund after some time. To avoid mixing two different funds together, we identify all the cases where there is a gap of over one year between two consecutive reporting dates for portfolio holdings under the same identification number, and then check whether the identification number has been recycled. To ensure the accuracy of the holdings-based timing measures, we require funds to have a minimum of 8 quarters of holdings data.

In the CRSP dataset, if a fund has multiple share classes, each share class has a unique identification number. These multiple share classes have the same portfolio compositions but may differ in sales charges or targeted investor clientele. In the Thomson dataset, they are treated as a single fund. We therefore combine multiple share classes in the CRSP data into a single fund, and calculate the weighted average returns of the fund using the total net asset values of the share classes as the weights. In addition, to ensure the accuracy of the return-based timing measures, we require funds to have at least 24 monthly return observations.

The Thomson and CRSP datasets use different fund identification number systems which cannot be matched to each other by any automatic scheme. We manually match funds between the two databases based on (1) fund names and (2) fund ticker symbols.[7] To ensure matching quality, we also compare the

[7]Fund ticker symbols become available in the Thomson dataset starting in January 1999.

investment objective, management company name, and total net assets information. The final matched database has 2,294 unique funds with information on both portfolio holdings and returns.[8]

We use the CRSP value-weighted index return as a proxy for market returns. Individual stock returns, and the 1-month T-bill yields (our proxy for the risk-free rate) are also from CRSP. We also 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 S&P 500 index, and earnings-to-price ratio on S&P 500 index. They are from DRI/CITIBASE.

Table I reports the summary statistics for our mutual fund sample during the period from 1980 to 2002. Out of a total of 2,294 funds, more than half (1,390) are growth funds. There are 255 aggressive growth funds, 411 growth and income funds, and 208 balanced funds. By averaging first over the times series data of each fund and then averaging 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%,[9] and an annual expense ratio of 1.31%. The funds invest 74.73% of 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 dataset, is 13.46 years. This number appears to be higher than that reported in the previous literature, which is likely due to the fact that we have aggregated over different share classes to a single fund. 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.

[8]We compare the number of mutual funds in our final sample with those in other studies that also use matched data from CRSP and Thomson/CDA. Wermers (2000) combines the CDA data with the CRSP dataset for the period between 1975 and 1994. His sample contains 1,788 equity funds. Kacperczyk, Sialm, and Zheng (2003) match the CRSP dataset with the Thomson dataset for the period between 1984 and 1999 and end up with 1,971 unique equity funds. Cohen, Coval, and Pastor (2004) also match the Thomson dataset with the CRSP dataset for the period between 1980 and 2002Q2. They report 235 matched 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.

[9]Following Sirri and Tufano (1998), the annualized load is the total load divided by 7.

B. Statistical Inference: The Bootstrapping Approach

We now describe the methodology used in this paper for statistical inference based on the cross-sectional distribution of the timing statistics.

Consider the holdings-based measure ($\hat{\gamma}$) from (6) as an example. When there is only one fund, one can make statistical inference on whether it has any significant timing ability based on the t-statistic of $\hat{\gamma}$. However, when there are a large number of funds, even though none of them has timing ability, just by random chance there will be some funds with significant t-statistics. Therefore we can no longer single out a fund and make inference based on its timing measure alone. This is analogous to the data-snooping problem discussed in Lo and MacKinlay (1990), Sullivan, Timmermann, and White (1999), and Kosowski, Timmermann, White, and Wermers (2001). One valid way of making inference 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 for the timing statistics. When the number of funds is large and the length of the time series data is short, the covariance matrix is likely to be singular. Such a test is therefore infeasible.

In this paper, we base the statistical inference on the cross-sectional distribution of the timing statistics. In other words, we examine whether the distribution of the timing statistics is significantly different from what one would expect under the null hypothesis that no fund has timing ability. We focus on the following cross-sectional statistics: the mean, median, standard deviation, skewness, kurtosis, the quartiles at 25% and 75%, as well as the extreme percentiles at 1%, 5%, 10%, 90%, 95%, and 99%. The statistical significance at those extreme percentiles can tell us whether there are funds with strongly positive timing ability and whether there are funds with perversely negative timing ability.

There are two additional issues associated with the statistical inference based on the cross-sectional distribution of the market-timing measures. One is the violation of the $i.i.d.$ assumption across funds. The other is the finite sample property of the test statistics. Consider the t-statistics for the timing measures, which are asymptotically normal under quite general conditions. If the t-statistics are further $i.i.d.$ across funds, then their cross-sectional properties are easy to obtain. For example, the 90th percentile for the t-

statistics of 1,000 funds should follow asymptotically the distribution of the 90th percentile of 1,000 $i.i.d.$ standard normal random variables. However, because the funds hold similar stocks, their betas are correlated. As a consequence, the timing statistics are correlated across funds as well. In this case, although the distribution for the 90th percentile statistic may still be asymptotically normal, the critical values under the $i.i.d.$ assumption are no longer valid. 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 those extreme percentiles, may differ significantly from their asymptotic ones. To address both issues, we resort to a bootstrapping approach similar to that proposed in Kosowski, Timmermann, White, and Wermers (2001). In the following, we explain in detail the bootstrapping approach used in this paper.

First consider the holdings-based measure. Let $\mathbf{b}_t$ = ($\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$, ..., $\hat{\beta}_{Kt}$) denote the estimated betas for K funds at the beginning of period t and r_{mt} the market return over the period t to $t+1$. Let $\Gamma(\mathbf{b}_t, r_{mt})$ be a cross-sectional statistic (say, the 90th percentile) for the holdings-based measure ($\hat{\gamma}$ or its t-statistic) computed using the sample fund betas $\mathbf{b}_t$ and market returns r_{mt}. In each bootstrap, we keep $\mathbf{b}_t$ unchanged but randomly resample the market returns r^*_{mt}. With $\mathbf{b}_t$ and r^*_{mt} we compute the holdings-based measure for all funds and calculate the bootstrapped cross-sectional statistic $\Gamma^*_j(\mathbf{b}_t, r^*_{mt})$. This procedure is repeated for a large number (J) of times to obtain a series of $\Gamma^*_j(\mathbf{b}_t, r^*_{mt})$, j=1, ... J. Since the market returns are randomly resampled, the fund beta $\mathbf{b}_t$ should bear no relationship with the bootstrapped market return r^*_{mt}. As the procedure preserves the covariance structure for $\mathbf{b}_t$, the distribution of $\Gamma^*_j(\mathbf{b}_t, r^*_{mt})$ approximates the distribution of $\Gamma(\mathbf{b}_t, r_{mt})$ under the null hypothesis of no timing ability with correlated fund betas. We compute the bootstrapped p-value for $\Gamma(\mathbf{b}_t, r_{mt})$ as:

$$p = \frac{1}{J} \sum_{j=1}^{J} I_{\Gamma^*_j > \Gamma} \tag{11}$$

where $I_{\Gamma^*_j > \Gamma}$ is an indicator function with value 1 if $\Gamma^*_j(\mathbf{b}_t, r^*_{mt}) > \Gamma(\mathbf{b}_t, r_{mt})$ and 0 otherwise.

For the return-based measure, we adopt a parametric bootstrap approach. First, given fund returns, we estimate the following model for each fund:

$$r_t = \alpha + \beta_0 r_{mt} + e_t \tag{12}$$

We retain the OLS estimates of parameters $\hat{\alpha}$ and $\hat{\beta}_0$, as well as the residuals $\hat{e}_t$ for all the funds. Then, we resample without replacement the market returns r^*_{mt} and construct the bootstrapped fund returns, r^*_t, as:

$$r^*_t = \hat{\alpha} + \hat{\beta}_0 r^*_{mt} + \hat{e}_t \tag{13}$$

Let $\mathbf{r}^*_t$ denote the vector of bootstrapped returns to all funds in period t. With $\mathbf{r}^*_t$ and r^*_{mt} we estimate the return-based timing measures for all funds, and calculate the bootstrapped cross-sectional statistic $\Gamma^*_j(\mathbf{r}^*_t, r^*_{mt})$. The distribution of $\Gamma^*_j(\mathbf{r}^*_t, r^*_{mt})$ is obtained after a large number of bootstraps. Since r^*_{mt} is drawn randomly, (13) ensures that $\gamma = 0$ in the bootstrapped data and at the same time preserves the covariance structure of fund returns under the null hypothesis. We then compute the return-based timing measure γ using bootstrapped fund returns and market returns, with the bootstrapped p-value computed similarly as in (11).

The timing measures ($\hat{\gamma}$) are non-pivotal statistics and heteroscedastic across funds as their distributions depend on nuisance parameters such as the variance of the error terms. On the other hand, their 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. This makes the cross-sectional inference based on t-statistics more efficient than that based on $\hat{\gamma}$ estimates. Therefore, while we have bootstrapped p-values of the cross-sectional statistics for both $\hat{\gamma}$ and their t-statistics, we focus more on the latter.

It is noted that in the above bootstrapping procedure, while the market returns are resampled, we have kept $\mathbf{b}_t$ and $\hat{e}_t$ fixed. As the length of the time series data varies across funds, it is difficult to jointly resample $\mathbf{b}_t$ (or $\hat{e}_t$) for all funds while keeping the length of the bootstrapped time series fixed for any given fund. However, performing the joint resampling procedure for a subsample of funds with relatively long time series observations, we find no material changes for the bootstrapped distributions, and consequently no material changes for the statistical inference.

III. Empirical Results of Market Timing

A. Return-based Measures

To obtain the return-based measures, we perform regressions (1) and (2) using monthly fund returns and market returns to obtain the estimate of $\hat{\gamma}$ for each fund. The t-statistic for $\hat{\gamma}$ is computed using the Newey-West heteroscedasticity and autocorrelation-consistent variance estimator, with a lag order of 6, which is equivalent to a 6-month time lag. We then calculate the statistics (1, 5, and 10th percentiles, 1st quartile, mean, median, 3rd quartile, 90, 95, and 99th percentiles, standard deviation, skewness, and kurtosis) on the cross-sectional distribution of $\hat{\gamma}$ as well as the t-statistics. The statistical inference is based on the bootstrapped p-values following the procedure described in Section II.B with 2,000 replications.

Figure 1 plots the density of the cross-sectional distribution of the t-statistics estimated from all funds, together with that of the bootstrapped t-statistics. Both density functions are estimated using a Gaussian kernel following the bandwidth choice of Silverman (1986). The comparison of these two density functions provides an informal but intuitive inference about the timing ability of the mutual funds. For the Treynor-Mazuy t-statistics, the sample density plot is visibly left-shifted relative to the bootstrapped one, indicating somewhat overall negative timing ability. For the Henriksson-Merton t-statistics, the density plot is also left-shifted, but to a lesser extent.

Table II reports the cross-sectional statistics for the return-based measures, the t-statistics, and their corresponding bootstrapped p-values. Consistent with the previous literature, the mean and median for the timing statistics are slightly negative, but not statistically significant based on the bootstrapped p-values. The cross-sectional standard deviations for both the Treynor-Mazuy and the Henriksson-Merton t-statistics are substantially higher than those under the null. For the Treynor-Mazuy t-statistics, the cross-sectional kurtosis is significantly higher. However, none of the extreme percentiles, the lower ones (1, 5 and 10th) and the higher ones (90, 95, and 99th), are statistically significant. In summary, the return-based measures do not provide clear evidence of positive or negative timing ability. Of course, as discussed in Section II.C, this is quite possibly due to their lack of statistical power.

B. Holdings-based Measures

To obtain the holdings-based measures, we first focus on the Treynor-Mazuy measure in (6), with monthly, quarterly, semi-annual, and annual forecasting horizons. In other words, the market return r_{mt} is measured for 1-, 3-, 6-, and 12-month after the reporting date of portfolio holdings. The betas of individual stocks are estimated on a rolling basis using the 12-month daily returns prior to the date of portfolio holdings. 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,

$$r_{i\tau} = a_i + \sum_{q=-5}^{5} b_{iq} r_{m,t-q} + e_{i\tau} \tag{14}$$

Following Dimson (1979), the stock beta is the sum of the estimated coefficients $\hat{b}_{iq}$, q=-5, ..., 5:

$$\hat{b}_i = \sum_{q=-5}^{5} \hat{b}_{iq} \tag{15}$$

We require a stock to have at least 60 daily observations during the estimation period, otherwise we assume a value of one for the stock beta.[10] Non-stock securities are assumed to have a beta of zero.

The portfolio betas are then computed according to (5), and the timing coefficient γ is estimated from (6) using OLS. To account for heteroscedasticity and autocorrelation, the t-statistic for $\hat{\gamma}$ is computed using the Newey-West covariance estimator with a lag order of 2, which is equivalent to a 6-month time lag as in the return-based regressions. Finally, the bootstrapped p-values for the cross-sectional statistics are obtained following the procedure described in Section II.B.

Figure 2 plots the density of the cross-sectional distribution of the t-statistics estimated from all funds, together with that of the bootstrapped t-statistics. The density functions are estimated the same way as in the case of return-based measures. Clearly, for all four forecasting horizons (1-, 3-, 6-, and 12-month), the sample density functions are right-shifted relative to the bootstrapped ones, indicating overall positive timing ability by mutual funds.

[10] We set the beta to one here since the unconditional expectation for stock betas is close to one. We also consider alternative beta values for stocks with insufficient data. See the discussion in Section III.C.4.

Table III reports the cross-sectional statistics for $\hat{\gamma}$ as well as their t-statistics. At the 1-month forecasting horizon, the mean and median t-statistics are positive, but not statistically significant according to the bootstrapped p-values. None of the extreme percentiles is statistically significant either. At the 3-month forecasting horizon, however, the bootstrapped p-values for the mean, median, the 25, 75, 90, 95, and 99th percentiles all become much smaller. The same is true for those cross-sectional statistics at the 6- and 12-month forecasting horizons. Therefore, at the intermediate horizons of 3 to 12 months, the holdings-based measures provide clear evidence of positive timing ability by mutual funds, with some funds exhibiting strong positive timing ability. There is no evidence that some funds have perversely strong negative timing ability based on the bootstrapped p-values for the lower extreme percentiles (1, 5, and 10th). Finally, we note that since $\hat{\gamma}$ is heteroscedastic across funds, the kurtosis for the cross-sectional distribution of $\hat{\gamma}$ is very large. In contrast, the kurtosis for the pivotal t-statistics is quite small.

C. Robustness Check

In this section, we perform robustness checks on the results reported in the previous section, with a series of variations in the estimation and testing procedure.

1. Alternative Stock Beta Estimates for Holdings-based Measures

We perform the holdings-based analysis using two alternative methods to estimate individual stock betas. First, we use monthly returns instead of daily returns, and perform the following regression:

$$r_{i\tau} = a_i + b_{i1} r_{m\tau} + b_{i2} r_{m\tau-1} + e_{i\tau} \tag{16}$$

The estimated stock beta is then $\hat{b}_i = \hat{b}_{i1} + \hat{b}_{i2}$. The regressions are performed on a rolling basis using the 60 monthly stock returns prior to the date of portfolio holdings. We require a stock to have at least 12 monthly return observations, otherwise we assume a value of one for the stock beta.

Second, we continue to use the model (14) with daily returns. However, the rolling regression is performed using the 3-month, instead of 12-month, daily returns prior to the date of portfolio holdings. A minimum of 22 daily returns are required, otherwise we assume a value of one for the stock beta.

The two alternative estimators may capture different time series features of the stock betas. Since betas are time-varying, the estimator from 5 years of monthly data may better capture the low-frequency component, while the estimator from 3 months of daily data is more likely to capture the short-term dynamic component.

Table IV reports the cross-sectional distribution of the t-statistics for the holdings-based measure with two alternative stock beta estimators. Again we test the timing ability for 1-, 3-, 6-, and 12-month forecasting horizons . When stock betas are estimated using 3 months of daily data, the results are generally consistent with those reported in Table III. At the 1-month horizon, there is no significant evidence of timing. At 3-, 6-, and 12-month horizons, there is evidence of positive timing ability by mutual funds, and also evidence of strong timing ability for certain funds. When stock betas are estimated using 5 years of monthly data, the results are consistent with those in Table III for 1-, 3-, and 6-month horizons. At the 12-month horizon, the sample mean and median t-statistics are no longer significant. But the sample statistics for the 75, 90, 95, and 99th percentiles remain significantly higher than the bootstrapped ones.

2. The Holdings-based Henriksson-Merton Measure

We also perform analysis using the holdings-based Henriksson-Merton timing measure based on (7), with 1-, 3-, and 6-month forecasting horizons. The 12-month horizon is not performed because at this horizon market returns are mostly positive, and the indicator function $I_{r_{mt}>0}$ has mostly a value of one. The individual stock betas are estimated using 1 year of daily returns, 5 years of monthly returns, and 3 months of daily returns.

The results are reported in Table V. At the 1-month horizon, the mean and median t-statistics are positive but not significant. The extreme percentiles are not significant either. At the 3-month horizon, the p-values for the mean and median t-statistics are smaller than for the 1-month horizon but still above 10%. There is some evidence of significant timing at higher extreme percentiles. For example, when the stock betas are estimated from the past 3-month daily returns, the 90 and 95th percentile statistics have p-values of 0.09. At the 6-month horizon, we find significantly positive mean and median t-statistics, except when the stock betas are estimated using the past 5-year monthly returns. Under all three stock beta estimators, the p-values for the 75, 90, 95, and 99th percentile statistics are very small.

In summary, under the holdings-based Henriksson-Merton measure, the evidence of positive timing ability is slightly weaker, but remains generally consistent with that under the holdings-based Treynor-Mazuy measure. There is a reasonable explanation for the weaker evidence using the Henriksson-Merton type measure. Model (7) assumes that fund managers condition fund betas only on the binary signal of whether or not the future market return is positive. If the true data-generating process for the fund beta is (3), the test based on (7) has relatively low power.

3. Return-based Measures: Extended Forecasting Horizons

Comparing the results in Tables II and III, one may notice that at 1-month forecasting horizon, the evidence of timing is weak for both return-based and holdings-based measures. Stronger evidence of timing in Table III appears to be obtained using longer forecasting horizons. Existing literature has implicitly focused only on the 1-month forecasting horizon when using the return-based measures. An interesting question naturally arises: if we extend the forecasting horizons, will the results for the return-based measures be consistent with the holdings-based measures?

We aggregate monthly returns into overlapping 3-, 6-, and 12-month returns and then perform regressions (1) and (2) using the aggregated returns. The results are reported in Table VI. Interestingly, we do find some evidence of positive timing ability. For example, although at the 3-month horizon the mean and median t-statistics are still negative, at the 6- and 12-month horizons they become positive. At the 6-month horizon, the p-values for the 90, 95, and 99th percentile statistics become very small for both the Treynor-Mazuy and Henriksson-Merton measures. At the 12-month horizon, we do not calculate the Henriksson-Merton measure due to the lack of variation in the value of the indication function $I_{r_{mt}>0}$. But for the Treynor-Mazuy measure, the p-values indicate that the 75, 90, and 95th percentile statistics are significantly higher than the bootstrapped ones.

Overall, due to the lack of statistical power, the evidence of positive timing for the return-based measures is weaker than that for the holdings-based measures. With extended horizons, the return-based measures are more agreeable with the holdings-based measures. By focusing only on the monthly horizon, we could have missed some important aspects of mutual fund market-timing activity.

4. Other Issues

In addition to the above robustness checks, a few other variations in the estimation and testing procedure are also considered. For brevity, we only summarize them as follows.

First, for the results reported in Table III, IV and V, we impose minimum requirement of stock returns when estimating individual stock betas. When such requirement is not met, we assume a value of one for the stock beta. In addition, when we cannot match a stock in fund holdings with the CRSP stock return data, we also assume a value of one for the stock beta. Such stocks account for a very small portion of fund holdings in value. To ensure the robustness of our results, we also experiment with the following variations in handling unmatched stocks or stocks with insufficient data: (1) setting their values to zero, and (2) estimating their betas using the post forecasting-period stock returns whenever they are available. The results are shown to be robust.

Second, when computing the Newey-West heteroscedasticity and autocorrelation-consistent t-statistics of the timing measures, we use 6 lags for the return-based measures and 2 lags for the holdings-based measures. Both are equivalent to a 6-month time lag. As the regressions involve overlapping data, the lag orders may not be sufficiently high to capture serial correlations in the data when we examine long forecasting horizons, such as 12-month. Since the bootstrapping procedure is also based on overlapping data, the bootstrapped p-values and therefore the statistical inference should be robust to the choice of lag order. As an extra caution, we also set the lag order equal to larger values with the time lag up to 12-month as well as a data-dependent value of $0.5T^{1/3}$, and the resulting p-values are indeed similar.

Finally, we also perform market-timing analysis in the presence of additional factors other than the market return, namely the size and book-to-market factors (SMB and HML), and find no material changes in our results.

IV. Fund Characteristics, Conditioning Information and Economic Value of Market Timing

A. Characteristics of Market Timers

To examine whether there is any systematic difference in market-timing activities across funds of different investment styles, we perform holdings-based timing analysis within each group with different investment objectives, namely aggressive growth, growth, growth and income, and balanced.

Table VII reports the cross-sectional distribution of the t-statistics for the holdings-based Treynor-Mazuy timing measure, where stock betas are estimated using the past 12 months of daily returns. Bootstrapped p-values are also reported. At the 1-month horizon, growth, growth and income, and balanced funds do not have significant evidence of timing ability. For the balanced funds, the mean and median t-statistics are even slightly negative. However, for aggressive growth funds, they are significantly positive. The 75, 90, 95, and 99th percentile statistics are significantly higher than the bootstrapped ones. At the 3-, 6- and 12-month forecasting horizons, there is evidence of positive average timing ability in all fund groups. The p-values for the higher extreme percentiles are mostly small, indicating that there exist some funds in each group with strong timing ability. Ranking funds according to the mean and median t-statistics, we find that aggressive growth funds are the most active market timers, followed by the growth funds, growth and income funds, and the balanced funds. The p-values reveal the same pattern.

Next, we examine whether the market-timing activities are associated with a few other fund characteristics. Our hypotheses are as follows. First, smaller funds may be more flexible in adjusting fund betas and therefore may be more active market timers. Second, a successful market timer may have higher total returns. Third, funds with strong timing ability may be able to charge higher fees, leading to higher expense ratios. Fourth, active timers trade more aggressively, leading to higher turnover ratios. Fifth, some funds impose higher loads to discourage investors from getting in and out of the funds frequently. Those funds are more likely to invest in illiquid securities and less likely to time the market.

We test these hypotheses using the following panel data regression that intends to gauge the effect of various fund characteristics on the fund timing measure:

$$\hat{\beta}_{it} = a_i + (b_0 + b_1\text{Turnover} + b_2\text{Load} + b_3\text{Expense} + b_4\text{RET} + b_5 \ln(\text{TNA}))r_{mt} + e_{it} \tag{17}$$

where $\hat{\beta}_{it}$ is the beta estimate of fund i with stock betas estimated using the past 12-month daily returns, Turnover is the annual turnover ratio, Load is the annualized load of a fund, Expense is the annual expense ratio, RET is the fund return, and ln(TNA) is the natural log of the total net assets value. The panel data model is specified with fixed effects a_i to capture the individual fund characteristics. Turnover, Load, Expense, RET, and TNA are the averages over their respective time series for each fund during the sample period. The panel data in our sample is unbalanced as not all funds co-exist during the entire sample period.

Table VIII reports the coefficient estimates with t-statistics in the parentheses next to the coefficient estimates. The error term e_{it} in the panel data model is assumed to have a flexible structure with cross-sectional heteroscedasticity and correlation as well as autocorrelation. The standard errors of the coefficient estimates are estimated following the robust procedure in Chamberlain (1985). For all four forecasting horizons, two variables emerge with consistently high t-statistics: Load and RET. The t-statistics for Turnover are high in three out of four horizons. Expense and TNA occasionally have high t-statistics but not consistent across horizons. Thus, the results seem to support our Hypotheses 2, 4, and 5. That is, market timing ability does contribute to higher returns of a fund, active timers tend to trade more actively, and they also tend to be no-load or low-load funds.

B. Conditioning Information of Mutual Fund Market Timing

Academic studies have long suggested that aggregate market returns are predictable and that such predictability may significantly affect an investor's asset allocation decisions. Empirical studies have also identified many economic variables that are predictive of market returns. We therefore turn to the following two questions. First, do fund managers adjust fund betas in response to those economic variables? Second, do fund managers rely on such return-predictive economic variables in their market timing?

We focus on several return-predictive economic variables documented in the finance and economics literature. They include the short-term interest rate, term premium, credit premium, aggregate dividend yield, and aggregate earnings-to-price ratio. Following the literature, we use the 1-month T-bill yield as the short-term interest rate. The term premium is the yield spread between the 10-year T-bond and the 1-month T-bill. The credit premium is the average yield spread between Moody's Baa-rated corporate bonds and Moody's Aaa-rated corporate bonds. The aggregate dividend yield and earnings-to-price ratio are based on the S&P 500 stocks.

To address the first question, we use the following panel data model to examine whether fund managers adjust fund betas based on those variables:

$$\hat{\beta}_{it} = a_i + bM_{t-1} + e_{it} \tag{18}$$

where M_{t-1} is the vector of the economic variables and b a vector of coefficients. Similar to (17), the error term e_{it} is assumed to have a flexible structure and the standard errors are estimated following the same robust procedure. The regression has an R^2 of 0.03 with coefficient estimates -7.81 (-1.01) for the short rate, -18.30 (-1.63) for the term premium, -12.02 (-0.43) for the credit premium, 50.91 (2.45) for the aggregate dividend yield, and 4.38 (2.05) for the aggregate earnings-to-price ratio, respectively, where the numbers in parentheses are the t-statistics of the coefficient estimates. The results indicate that the aggregate dividend yield and the aggregate earnings-to-price ratio have a significantly positive impact on fund betas, while the short-term interest rate, term spread, and credit spread do not.

To address the second question of whether or not fund managers rely on economic variables to time the market, we perform the following regression for each fund:

$$\hat{\beta}_t = a + bM_{t-1} + \gamma r_{mt} + e_t \tag{19}$$

where the fund subscript is dropped, M_{t-1} is a vector of five economic variables and b is a vector of coefficients. Our interest is whether the timing measure γ remains significant after controlling for the return-predictive economic variables. The same intuition is also behind the conditional return-based market-timing test of Ferson and Schadt (1996). We first estimate γ and its Newey-West t-statistics (with a lag order of 2) for each fund, and then compute their cross-sectional statistics.

The statistical inference of the timing measure in (19) is based on the following parametric bootstrap procedure extended from that in Section II.B. First, we estimate the following models using OLS for the monthly series of macroeconomic variables and market returns:

$$M_{jt} = a_j + b_j M_{jt-1} + e_{jt}, \quad j = 1, ..., 5 \tag{20}$$

$$r_{mt} = a_r + \sum_{j=1}^{5} b_{rj} M_{jt-1} + e_{rt} \tag{21}$$

where M_{jt} denotes the $j-th$ macroeconomic variable. We retain the parameter estimates $\hat{a}_j$, $\hat{b}_j$, $\hat{a}_r$, and $\hat{b}_{rj}$, as well as the estimated residuals $\hat{e}_{jt}$ and $\hat{e}_{rt}$. Then, in each bootstrap, we jointly resample the estimated residuals for $\hat{e}_{rt}$ and $\hat{e}_{jt-1}$, with j=1, ..., 5. Denote the resampled series by e^*_{rt} and e^*_{jt-1}, we then construct the monthly series of bootstrapped economic variables M^*_{jt} and the bootstrapped market return r^*_{mt} as:

$$M^*_{jt} = \hat{a}_j + \hat{b}_j M_{jt-1} + e^*_{jt}, \quad j = 1, ..., 5 \tag{22}$$

$$r^*_{mt} = \hat{a}_r + \sum_{j=1}^{5} \hat{b}_{rj} M^*_{jt-1} + e^*_{rt} \tag{23}$$

We perform the market-timing test of (19) for each fund, using the bootstrapped M^*_{jt} and r^*_{mt} as regressors, and compute the bootstrapped p-values for the Newey-West t-statistics based on 2,000 replications. Since the conditional timing test (19) has more regressors than the original test (6), we increase the minimum requirement of holdings data from 8 to 10 quarters. This reduces the total number of funds in the sample by 113.

The cross-sectional distribution of the t-statistics for γ are reported in Table IX with the bootstrapped p-values. The results suggest that adding macroeconomic control variables does not substantially change the test statistics of the timing measures. For example, at the 1-month forecasting horizon, the p-values for the mean and median t-statistics are above 0.1. At 3-, 6-, and 12-month forecasting horizons, those p-values are below 0.1. Many other statistics, such as those for the 90, 95 and 99th percentiles, remain significant. In other words, the market-timing ability of mutual funds does not disappear after controlling for the return-predictive economic variables.

Combining the results from (18) and (19), while there is evidence that fund managers do adjust fund betas in response to certain economic variables, they do not solely rely on those variables in timing the market. There are two possible explanations. First, we have not exhausted all the conditional return-predictive signals documented in the academic studies. It is possible that mutual fund market-timing activities are related to other conditional signals not examined here. Second and perhaps more likely, fund managers are using a set of conditional information beyond the documented economic variables. Identifying such conditional information is an interesting and yet challenging task. We leave this for future research.

C. Potential Economic Value of Market Timing

The above findings indicate that fund managers not only have positive timing ability but also likely use private signals to time the market. Such ability is of economic value. Merton (1981) points out that the value of market timing can be quantified as that of a contingent claim on the market index. Glosten and Jagannathan (1994) provide a general continuous-time framework for implementing this intuition. Here we follow their approach to assess the potential economic value of market timing. To be specific, suppose the fund return process is:

$$r_t - r_f = \beta_0(r_{mt} - r_f) + \gamma(r_{mt} - r_f)^2 + e_t \tag{24}$$

where r_f is the constant risk free rate, $\gamma = 0$ indicates no timing ability, and $\gamma > 0$ indicates positive timing ability. Therefore, market timing generates additional terminal payoff $\gamma(r_{mt} - r_f)^2$ to a fund, with maturity equal to the forecasting horizon (1-, 3-, 6-, or 12-month). Further assume that the market index P_m follows a geometric Brownian motion,

$$dP_m/P_m = u_m dt + \sigma_m dW \tag{25}$$

Then, the value of market timing V can be derived as:

$$V = \gamma E^Q[(r_{mt} - r_f)^2] = \gamma(1 + r_f)(e^{\sigma_m^2} - 1) \tag{26}$$

where $E^Q[.]$ denotes the expectation under the risk-neutral measure.

We compute $\hat{V} = \hat{\gamma}(e^{\hat{\sigma}_m^2} - 1)$ for each fund, where $\hat{\gamma}$ is estimated using regression (6) with stock betas estimated from the past 1-year daily returns, and $\hat{\sigma}_m^2$ is estimated from monthly returns of the CRSP value-weighted index between 1964 and 2002. We then examine the properties of the cross-sectional distribution for $\hat{V}$. Similar to $\hat{\gamma}$, $\hat{V}$ is heteroscedastic across funds. We therefore also base our inference on the re-scaled variable, $\hat{V}/\sigma_v$, where σ_v is the standard deviation of $\hat{V}$ and obtained via bootstrapping. The bootstrapping procedure is a slight modification of that used for γ. The difference is that while the resampled market returns for the subperiod of 1980-2002 are used to compute γ, the resampled returns for the whole period 1964-2002 are used to compute σ_m^2. Denote the estimated value of timing for a fund in the j-th bootstrap by $\hat{V}_j$. Its standard deviation σ_v is estimated across all bootstraps. Using the same bootstrapping procedure, we also obtain bootstrapped p-values for the cross-sectional statistics for $\hat{V}$ and $\hat{V}/\sigma_v$.

The results are reported in Table X. The value of $\hat{V}$ is in annualized percentage terms. The mean and median values of $\hat{V}$ are quite high. For example, at the 3-month horizon $\hat{V}$ has a median of 1.22%. Suppose the annualized risk free rate is 4%, equation (26) implies that the annualized economic value of timing is about 1.23% for the median fund. The economic value is statistically significant based on the bootstrapped p-values. Further, the estimates of $\hat{V}$ for different horizons are of similar size.

We caution that the results reported in Table X are potential economic value of market timing. In the Glosten and Jagannathan (1994) framework, fund managers are assumed to be able to adjust fund betas continuously, and there are no transaction costs. With discrete-time trading and realistic transaction costs, the full value of market timing may not be attainable. In addition, fund betas are assumed to follow the data-generating process in (3). A more realistic beta process is probably somewhere between (3) and (4). The latter would imply a lower value of timing. Finally, the market index may not follow a simple geometric Brownian motion. Realistic features, such as stochastic volatility and jumps, may further affect the value of timing as a contingent claim. Nonetheless, the magnitude of the potential economic values suggests the importance of market-timing strategy by mutual funds.

V. Conclusion

In this paper, we examine mutual fund timing ability using both fund returns data and portfolio holdings data. Consistent with the previous literature, the return-based measures indicate that on average mutual funds do not have positive timing ability. On the other hand, the holdings-based measures provide evidence of significantly positive timing ability by mutual funds. The results also suggest that there are some funds with strong timing ability; the proportion of those funds in our sample far exceeds what one would expect if no fund has any timing ability.

We offer two explanations for the different results between the return-based measures and the holdings-based measures. First, the holdings-based measures are more accurate and statistically more powerful in detecting market timing ability. Second, the return-based measures may suffer from "artificial timing" bias, while the holdings-based measures do not. For these reasons, the positive evidence based on the holdings-based measures should weigh more heavily in shaping our belief about mutual fund timing ability.

Classifying the funds according to investment objectives, we find that aggressive growth funds are the most active market timers, followed by growth funds, growth and income funds, with balanced funds the least active. There is also evidence that market timers trade more actively, have higher returns, and tend to be of no loads or low loads. However, the market timing funds seem to have market-timing information beyond those common return-predictive economic variables documented in the academic studies. What conditional information is used by fund managers to time the market remains an interesting and open question.

Finally, we quantify the potential economic value of timing as a contingent claim on the market index. We find that market timing may contribute to annual fund performance by as high as 1.23% for the median fund. By making simplifying assumptions and omitting transaction costs, our estimate is probably higher than the attainable value by mutual funds. Nonetheless, the magnitude of the estimated value indicates that market-timing is an important investment strategy deserving more academic attention.

References

Amemiya, T., 1978, A note on a random coefficients model, *International Economic Review* 19(3), 793-796.

Andersen, Torben, Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2003, Modeling and forecasting realized volatility, *Econometrica* 71, 529-626.

Aragon, G., 2002, Timing multiple markets: Theory and evidence, *Working Paper*, Boston College.

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

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

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

Blake, David, Bruce Lehmann, and Alan Timmermann, 1999, Asset allocation dynamics and pension fund performance, *Journal of Business* 72, 429-461.

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

Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373-399.

Campbell, John Y., Yeung Lewis Chan, and Luis M. Viceira, 1999, A multivariate model of strategic asset allocation, *Working paper*, Harvard University.

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

Campbell, John Y., and Robert J. Shiller, 1988b, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661-676.

Campbell, John Y. and Luis M. Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433-495.

Chamberlain, Gary, 1985, Panel data; in Griliches and Intriligator (eds.) *The Handbook of Econometrics* Vol. II (North Holland, Amsterdam).

Cochrane, J., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 207-234.

Coggin, T. Daniel, Frank J. Fabozzi, and Shafiqur Rahman, 1993, The investment performance of us equity pension fund managers, *Journal of Finance* 48, 1039-1056.

Cohen, Randolph, Joshua Coval, and Lubos Pastor, 2004, Judging fund managers by the company they keep, *Journal of Finance* forthcoming.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.

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

Fama, Eugene F., and Kenneth R. French, 1987, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.

Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3-25.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23-49.

Fama, Eugene F., and W. Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115-146.

Ferson, Wayne E., 1989, Changes in expected security returns, risk and the level of interest rates, *Journal of Finance* 44, 1191-1217.

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

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

French, K.R., G.W. Schwert, and R.F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3-30.

Fung, Hung-Gay, Eleanor Xu, and Jot Yau, 2002, Global hedge funds: risk, return, and market timing, *Financial Analyst Journal* 58, 19-30.

Glosten, Lawrence, and Ravi Jagannathan, 1994, A contingent claims approach to performance evaluation, *Journal of Empirical Finance* 1, 133-166.

Graham, John, and Campbell Harvey, 1996, Market timing ability and volatility implied in investment newsletters' asset allocation recommendations, *Journal of Financial Economics* 42, 397-421.

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62, 393-416.

Grinblatt, Mark, and Sheridan Titman, 1994, A study of monthly mutual fund returns and performance evaluation techniques, *Journal of Financial and Quantitative Analysis* 29, 419-444.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088-1105.

Goetzmann, William N., Jonathan Ingersoll Jr., and Zoran Ivkovich, 2000, Monthly measurement of daily timers, *Journal of Financial and Quantitative Analysis* 35, 257-290.

Hall, P., 1992, *The Bootstrap And Edgeworth Expansion* (Springer Verlag).

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

Jagannathan, R., and R. Korajczyk, 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, Marcin, Clemens Sialm and Lu Zheng, 2003, On the industry concentration of actively managed equity mutual funds, *Working Paper*, University of Michigan.

Keim, Donald, and Robert Stambaugh, 1986, Predicting returns in stock and bond markets, *Journal of Financial Economics* 17, 357-390.

Kendal, S., and Robert Stambaugh, 1996, On the predictability of stock returns: An asset-allocation perspective, *Journal of Finance* 51, 385-424.

Koski, Jennifer Lynch, and Jeffrey Pontiff, 1999, How are derivatives used? Evidence from the mutual fund industry, *Journal of Finance* 54, 791-816.

Kosowski, R., A. Timmermann, H. White, and R. Wermers, 2001, Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis, *Working Paper*, University of Maryland.

Lamont, O., 1998, Earnings and expected returns, *Journal of Finance* 53, 1563-1587.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815-849.

Lewellen, J., 1999, The time-series relations among expected return, risk, and book-to-market, *Journal of Financial Economics* 54, 5-43.

Lo, Andrew, and Craig MacKinlay, 1990, Data snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

Merton, Robert C., 1980, On estimating the expected return on the market: an exploratory investigation, *Journal of Financial Economics* 8, 323-361.

Merton, Robert C., 1981, on market timing and investment performance part I: An equilibrium theory of value for market forecasts, *Journal of Business* 54, 363-406.

Pesaran, Hashem, and Allan Timmermann, 1994, A generalization of the non-parametric henriksson-merton test of market timing, *Economics Letters* 44, 1-7.

Pontiff, J., and L. Schall, 1998, Book-to-market ratios as predictors of market returns, *Journal of Financial Economics* 49, 141-160.

Silverman, Bernie W., 1986, *Density Estimation For Statistics And Data Analysis* (Chapman and Hall: London).

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

Sullivan, Ryan, Alan Timmermann, and Halbert White, 1999, Data-snooping, technical trading rules and the bootstrap, *Journal of Finance* 54, 1647-1692.

Treynor, Jack, and Kay Mazuy, 1966, Can mutual funds outguess the market?, *Harvard Business Review* 44, 131-136.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *Journal of Finance* 54, 581-622.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses, *Journal of Finance* 55, 1655-1703.

Wermers, Russ, 2003, Is money really “smart”? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, *Working Paper*, University of Maryland.

Appendix A. Simulation of Return-based versus Holdings-based Measures

We simulate daily returns for N stocks with the following return process:

$$r_{it} - r_f = \beta_i(r_{mt} - r_f) + \epsilon_{it} \tag{A.1}$$

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

Assume that a fund manager knows perfectly a stock's beta, β_i. She ranks all stocks according to their betas, and divides them into a high-beta group (H), and a low-beta group (L), each with N/2 stocks. She calculates the average betas for the two groups: $\beta^H = \sum_{i \in H} \beta_i/(2N)$ and $\beta^L = \sum_{i \in L} \beta_i/(2N)$. At the beginning of each quarter, the fund manager receives a signal about market returns for the next quarter:

$$S_t = r_{mt} - r_f + e_{st} \tag{A.2}$$

where $e_{st} \sim N(0, \sigma_s^2)$.

To take advantage of the timing signal S_t, the fund manager adjusts portfolio weights ω_{it} in the following way: the weight for each stock in the high-beta group is $\omega_{Ht} = 1/N + 2\gamma(S_t - u_m)/(\beta^H - \beta^L)/N$, and the weight for each stock in the low-beta group is $\omega_{Lt} = 1/N - 2\gamma(S_t - u_m)/(\beta^H - \beta^L)/N$. She holds the portfolio weights constant for the next quarter. Therefore the fund beta for the next quarter is

$$\beta_t = \gamma(S_t - u_m) + \bar{\beta}$$

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, γ measures the aggressiveness of her timing strategy. The fund return is:

$$\begin{aligned} r_t - r_f &= \beta_t(r_{mt} - r_f) + \sum_{i \in H} \omega_{Ht}\epsilon_{it} + \sum_{i \in L} \omega_{Lt}\epsilon_{it} = [\gamma(S_t - u_m) + \bar{\beta}](r_{mt} - r_f) + e_t \\ &= (\bar{\beta} - \gamma u_m)(r_{mt} - r_f) + \gamma(r_{mt} - r_f)^2 + \gamma e_{st}(r_{mt} - r_f) + e_t \end{aligned} \tag{A.3}$$

where $e_t = \sum_{i \in H} \omega_{Ht}\epsilon_{it} + \sum_{i \in L} \omega_{Lt}\epsilon_{it}$.

We simulate the stock returns r_{it} and the fund manager's private information S_t according to (A.1) and (A.2), respectively. Subsequently we obtain the fund manager's portfolio weights ω_{Ht} and ω_{Lt}, the fund beta process β_t, and the fund return process r_t. For simplicity, we set $r_f = 0$, $u_m = 0.5\%/22$. We set the number of stocks N=60, approximately the median number of stocks held by the mutual funds in our sample (from Table I). The standard deviation of the idiosyncratic stock return ϵ_{it} is set as $\sigma_e = 0.17/\sqrt{22}$, where

0.17 is approximately the cross-sectional average of the standard deviation of monthly return residuals using a market model, based on the CRSP stock return data. The standard deviation of the market return is set as $\sigma_m = 0.045/\sqrt{22}$, close to that of the CRSP value-weighted index returns. The total number of months of stock returns and fund returns T = 132 + 60, where 132 is the mean number of monthly stock return observations for a typical fund in our sample, and an additional 60 months of stock returns are simulated for the purpose of estimating the stock betas. We vary two parameters in the simulation. First, the noise-signal ratio for S_t, σ_s/σ_m, takes values of 0.25, 0.5, 1, 2, and 4. Second, the timing measure γ is set to 0.1, 0.36, 1.35, and 3.11, with 0.36, and 3.11 being the median and 90th percentile of the estimated γ from our data (Table III) at the 3-month forecasting horizon. We also consider γ = 0.10 and 1.57 in order to cover wide range degrees of timing activities.

In each simulation, we estimate the return-based measure (1) and the holdings-based measure (6). To estimate the return-based measure (RB), we use 132 monthly fund returns (from month 61 to month 182). For the holdings-based measure, we assume that the portfolio holdings are observed only at quarter-end. We estimate the stock betas on a rolling basis, using three methods: (1) with 264 daily returns (1-year) prior to a quarter-end (HB1), (2) with 60 monthly returns (5 year) prior to a quarter-end (HB2), and (3) with 66 daily returns (3 months) prior to a quarter-end (HB3). We then estimate the fund betas using the quarterly portfolio weights and estimated stock betas. Finally, we compute the holdings-based measure using 27 quarters of fund betas, from quarter 21 to quarter 47. Note that the return-based and holdings-based timing measures employ time series data of different lengths. This is consistent with our data sample.

The simulation is replicated 10,000 times to obtain a series of return-based and holdings-based timing measures. Table A.I reports the means and standard deviations of various measures. The results confirm our conjecture in Section I.C. The means of both types of measures are quite close to the true values, except that there is a slight downward bias in the return-based measure. There are significant differences in the standard deviations. For example, when $\gamma = 0.10$ and $\sigma_s/\sigma_m = 0.25$, the standard deviation for the return-based measure is 0.72, while those for the three holdings-based measures are 0.07, 0.16, and 0.12, respectively. The two types of measures are closest to each other in standard deviation for the case where $\gamma = 3.11$ and $\sigma_s/\sigma_m = 4$. Even in this case, the standard deviation of the return-based measure is about 50% larger than those for the holdings-based measures.

We also perform simulations with k=0. In this case, the standard deviation for the return-based measure is 0.72, versus 0.08, 0.16, and 0.12 for the three holdings-based measures. For brevity, these results are not reported.

Table I
Summary Statistics of Fund Characteristics

Table I 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 first averaged over the time series data of each fund, and then averaged across funds. A fund's annual load is calculated as the total load divided by 7. 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 & Income	Balanced
Number of funds	2294	255	1390	441	208
Total net assets ($ millions)	482.60	545.95	397.52	773.93	360.96
Annual return (%)	9.96	11.76	9.97	9.71	8.18
Turnover (%/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 (%)	74.73	70.37	81.36	66.44	46.71
Age (years)	13.46	17.21	11.86	16.15	13.46
Average number of stocks held	102.15	91.57	104.70	90.94	122.39
Median number 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

Table II

Return-based Timing Measures

Table II reports the cross-sectional distribution of the estimated Henriksson-Merton measure and Treynor-Mazuy measure and the Newey-West t-statistics. The forecasting horizon is 1-month. The bootstrapped p-values for both the γ estimates and the Newey-West t-statistics are reported, respectively, in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

	1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
Panel A: Treynor-Mazuy Measure (1-Month Forecasting Horizon)													
$\hat{\gamma}$	-3.13	-2.10	-1.63	-0.79	-0.26	-0.15	0.30	0.79	1.25	2.67	1.09	0.32	9.10
p	(0.65)	(0.79)	(0.81)	(0.81)	(0.70)	(0.69)	(0.52)	(0.46)	(0.45)	(0.47)	(0.43)	(0.46)	(0.35)
t	-4.08	-2.86	-2.14	-1.27	-0.28	-0.28	0.71	1.53	2.12	3.42	1.59	0.28	8.54
p	(0.89)	(0.88)	(0.84)	(0.77)	(0.67)	(0.66)	(0.51)	(0.42)	(0.37)	(0.28)	(0.06)	(0.32)	(0.04)
Panel B: Henriksson-Merton Measure (1-Month Forecasting Horizon)													
$\hat{\gamma}$	-0.66	-0.42	-0.32	-0.15	-0.04	-0.01	0.09	0.20	0.29	0.53	0.23	-0.25	6.38
p	(0.47)	(0.58)	(0.62)	(0.60)	(0.59)	(0.56)	(0.60)	(0.62)	(0.63)	(0.71)	(0.74)	(0.57)	(0.36)
t	-3.05	-2.09	-1.63	-0.86	-0.08	-0.08	0.72	1.47	1.92	2.90	1.24	-0.03	3.96
p	(0.89)	(0.83)	(0.78)	(0.69)	(0.57)	(0.58)	(0.44)	(0.33)	(0.28)	(0.20)	(0.04)	(0.54)	(0.50)

Table III

Holdings-based Treynor-Mazuy Measure

Table III reports the cross-sectional distribution of the estimated holdings-based Treynor-Mazuy measure and the Newey-West t-statistics with 1-, 3-, 6-, and 12-month forecasting horizons. The stock betas are estimated using the past 1-year daily returns. The timing measures $\hat{\gamma}$ are pre-multiplied by 100. The bootstrapped p-values for both the γ estimates and the Newey-West t-statistics are reported, respectively, in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

	1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
						Panel A: 1-Month Forecasting Horizon							
$\hat{\gamma}$	-14.96	-5.33	-3.22	-0.90	1.35	0.34	2.09	5.00	7.48	24.27	30.54	40.16	1870
p	(0.31)	(0.40)	(0.44)	(0.29)	(0.12)	(0.14)	(0.11)	(0.10)	(0.14)	(0.10)	(0.39)	(0.03)	(0.05)
t	-3.00	-1.73	-1.22	-0.53	0.19	0.23	0.95	1.58	2.02	3.10	1.21	-0.41	5.75
p	(0.28)	(0.18)	(0.14)	(0.18)	(0.29)	(0.24)	(0.33)	(0.47)	(0.52)	(0.62)	(0.86)	(0.85)	(0.53)
						Panel B: 3-Month Forecasting Horizon							
$\hat{\gamma}$	-7.61	-2.78	-1.48	-0.21	1.05	0.36	1.57	3.11	4.55	13.81	12.20	22.49	662.4
p	(0.26)	(0.33)	(0.21)	(0.04)	(0.04)	(0.01)	(0.02)	(0.04)	(0.08)	(0.07)	(0.72)	(0.21)	(0.63)
t	-3.09	-1.79	-1.15	-0.25	0.60	0.64	1.51	2.30	2.77	3.88	1.42	-0.27	4.60
p	(0.50)	(0.32)	(0.13)	(0.03)	(0.02)	(0.01)	(0.01)	(0.03)	(0.05)	(0.12)	(0.14)	(0.80)	(0.73)
						Panel C: 6-Month Forecasting Horizon							
$\hat{\gamma}$	-6.40	-2.38	-1.24	-0.16	1.20	0.46	1.44	2.85	3.97	15.81	11.50	22.73	633.9
p	(0.39)	(0.45)	(0.29)	(0.07)	(0.01)	(0.01)	(0.01)	(0.03)	(0.06)	(0.00)	(0.50)	(0.20)	(0.64)
t	-4.04	-2.09	-1.39	-0.26	0.72	0.78	1.79	2.73	3.35	4.65	1.76	-0.55	7.77
p	(0.79)	(0.47)	(0.31)	(0.06)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.09)	(0.04)	(0.92)	(0.15)
						Panel D: 12-Month Forecasting Horizon							
$\hat{\gamma}$	-5.41	-2.00	-1.20	-0.17	0.85	0.36	1.12	2.30	3.20	10.37	8.47	19.27	546.1
p	(0.33)	(0.36)	(0.34)	(0.11)	(0.05)	(0.04)	(0.07)	(0.09)	(0.16)	(0.06)	(0.54)	(0.23)	(0.67)
t	-5.36	-2.61	-1.59	-0.37	0.85	0.87	2.07	3.44	4.40	6.43	2.24	-0.41	7.90
p	(0.83)	(0.48)	(0.26)	(0.10)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)	(0.86)	(0.14)

Table IV

Holdings-based Treynor-Mazuy Measure: Alternative Stock Betas

Table IV reports the cross-sectional distribution of the Newey-West t-statistics for the holdings-based Treynor-Mazuy measure with 1-, 3-, 6-, and 12-month forecasting horizons. The holdings-based measures are based on alternative stock betas estimated from the past five-year monthly returns (5y) and the past 3-month daily returns (3m), respectively. The bootstrapped p-values for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

		1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
							Panel A: 1-Month Forecasting Horizon							
5y	t	-2.84	-1.64	-1.23	-0.52	0.20	0.24	0.97	1.63	1.99	2.92	1.16	-0.19	4.34
	p	(0.15)	(0.07)	(0.14)	(0.14)	(0.25)	(0.21)	(0.31)	(0.39)	(0.57)	(0.81)	(0.94)	(0.69)	(0.90)
3m	t	-2.81	-1.73	-1.15	-0.46	0.29	0.30	1.04	1.72	2.23	3.25	1.22	0.04	5.32
	p	(0.09)	(0.11)	(0.03)	(0.04)	(0.10)	(0.09)	(0.15)	(0.27)	(0.30)	(0.55)	(0.87)	(0.41)	(0.65)
							Panel B: 3-Month Forecasting Horizon							
5y	t	-3.39	-1.87	-1.19	-0.30	0.59	0.69	1.50	2.31	2.91	4.20	1.49	-0.17	4.92
	p	(0.69)	(0.39)	(0.13)	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.02)	(0.06)	(0.10)	(0.72)	(0.63)
3m	t	-3.02	-1.86	-2.23	-0.40	0.44	0.46	1.27	2.11	2.73	4.01	1.43	-0.14	5.15
	p	(0.50)	(0.32)	(0.13)	(0.03)	(0.02)	(0.01)	(0.01)	(0.03)	(0.05)	(0.12)	(0.12)	(0.71)	(0.56)
							Panel C: 6-Month Forecasting Horizon							
5y	t	-4.70	-2.43	-1.56	-0.44	0.63	0.70	1.75	2.86	3.50	5.17	1.90	-0.34	5.50
	p	(0.88)	(0.71)	(0.47)	(0.13)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.87)	(0.45)
3m	t	-3.96	-2.23	-1.44	-0.45	0.48	0.57	1.47	2.35	2.90	4.12	1.62	-0.64	6.64
	p	(0.86)	(0.65)	(0.31)	(0.07)	(0.03)	(0.01)	(0.01)	(0.02)	(0.03)	(0.08)	(0.05)	(0.97)	(0.19)
							Panel D: 12-Month Forecasting Horizon							
5y	t	-6.80	-3.51	-2.26	-0.80	0.38	0.44	1.67	3.02	3.99	6.23	2.38	-0.33	6.05
	p	(0.96)	(0.83)	(0.64)	(0.33)	(0.23)	(0.20)	(0.12)	(0.08)	(0.07)	(0.08)	(0.03)	(0.83)	(0.33)
3m	t	-4.86	-2.36	-1.64	-0.56	0.41	0.42	1.45	2.58	3.21	4.64	1.82	-0.61	7.07
	p	(0.93)	(0.42)	(0.31)	(0.10)	(0.09)	(0.08)	(0.05)	(0.02)	(0.05)	(0.13)	(0.06)	(0.98)	(0.11)

Table V

Holdings-based Henriksson-Merton Measure

Table V reports the cross-sectional distribution of the Newey-West t-statistics for the holdings-based Henriksson-Merton measure with 1-, 3-, 6-, and 12-month forecasting horizons. The stock betas are estimated using the past 1-year daily returns (1y), the past five-year monthly returns (5y), and the past 3-month daily returns (3m), respectively. The bootstrapped p-values for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

		1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
							Panel A: 1-Month Forecasting Horizon							
1y	t	-2.62	-1.83	-1.34	-0.70	0.09	0.15	0.90	1.56	1.90	2.54	1.15	-0.15	3.20
	p	(0.61)	(0.67)	(0.52)	(0.47)	(0.39)	(0.32)	(0.32)	(0.28)	(0.36)	(0.62)	(0.34)	(0.79)	(0.90)
5y	t	-2.48	-1.77	-1.31	-0.61	0.23	0.29	1.04	1.70	2.10	2.80	1.19	-0.05	3.28
	p	(0.42)	(0.57)	(0.46)	(0.32)	(0.19)	(0.13)	(0.14)	(0.14)	(0.15)	(0.29)	(0.22)	(0.61)	(0.84)
3m	t	-2.54	-1.74	-1.33	-0.70	0.06	0.06	0.80	1.49	1.84	2.65	1.11	0.04	3.43
	p	(0.36)	(0.47)	(0.45)	(0.46)	(0.48)	(0.44)	(0.51)	(0.45)	(0.59)	(0.68)	(0.70)	(0.44)	(0.86)
							Panel B: 3-Month Forecasting Horizon							
1y	t	-2.58	-1.62	-1.21	-0.55	0.32	0.33	1.17	1.85	2.22	3.13	1.23	-0.19	3.69
	p	(0.52)	(0.39)	(0.35)	(0.28)	(0.16)	(0.16)	(0.09)	(0.09)	(0.13)	(0.15)	(0.15)	(0.83)	(0.61)
5y	t	-3.02	-1.88	-1.41	-0.71	0.17	0.20	1.05	1.79	2.16	2.93	1.26	-0.17	3.15
	p	(0.85)	(0.70)	(0.63)	(0.51)	(0.31)	(0.27)	(0.14)	(0.10)	(0.14)	(0.24)	(0.15)	(0.82)	(0.92)
3m	t	-2.81	-1.64	-1.13	-0.47	0.29	0.29	1.05	1.78	2.25	3.25	1.20	-0.01	3.79
	p	(0.67)	(0.33)	(0.15)	(0.12)	(0.12)	(0.10)	(0.12)	(0.09)	(0.09)	(0.13)	(0.30)	(0.56)	(0.65)
							Panel C: 6-Month Forecasting Horizon							
1y	t	-3.88	-2.40	-1.66	-0.60	0.51	0.59	1.63	2.62	3.28	4.66	1.81	-0.07	6.03
	p	(0.90)	(0.87)	(0.72)	(0.34)	(0.09)	(0.07)	(0.02)	(0.01)	(0.01)	(0.02)	(0.03)	(0.58)	(0.18)
5y	t	-4.70	-2.82	-2.07	-0.79	0.35	0.44	1.58	2.65	3.22	4.80	2.01	-0.43	11.11
	p	(0.98)	(0.96)	(0.92)	(0.54)	(0.16)	(0.11)	(0.02)	(0.01)	(0.01)	(0.02)	(0.03)	(0.87)	(0.08)
3m	t	-3.53	-2.25	-1.52	-0.52	0.45	0.49	1.46	2.42	2.97	4.47	1.63	-0.06	4.55
	p	(0.88)	(0.86)	(0.64)	(0.17)	(0.04)	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)	(0.04)	(0.57)	(0.46)

Table VI

Return-based Timing Measures: Extended Forecasting Horizons

Table VI reports the cross-sectional distribution of the Newey-West t-statistics for the return-based Treynor-Mazuy measure (TM) and Henriksson-Merton measure (HM) with 3-, 6-, and 12-month forecasting horizons. The bootstrapped p-values for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

		1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
						Panel A: 3-Month Forecasting Horizon								
TM	t	-4.02	-2.61	-1.95	-1.12	-0.15	-0.17	0.77	1.84	2.52	3.77	1.55	0.04	3.85
	p	(0.90)	(0.83)	(0.77)	(0.70)	(0.60)	(0.59)	(0.50)	(0.31)	(0.22)	(0.12)	(0.06)	(0.45)	(0.55)
HM	t	-3.43	-2.17	-1.66	-0.92	-0.10	-0.16	0.67	1.65	2.26	3.21	1.34	0.07	3.71
	p	(0.85)	(0.75)	(0.70)	(0.64)	(0.56)	(0.60)	(0.50)	(0.33)	(0.24)	(0.20)	(0.09)	(0.39)	(0.63)
						Panel B: 6-Month Forecasting Horizon								
TM	t	-6.17	-4.08	-3.06	-1.42	0.37	0.42	2.14	3.90	4.79	6.55	2.74	-0.06	3.52
	p	(0.87)	(0.85)	(0.79)	(0.61)	(0.28)	(0.28)	(0.13)	(0.02)	(0.02)	(0.01)	(0.01)	(0.41)	(0.84)
HM	t	-6.14	-4.29	-3.37	-1.63	0.14	0.32	1.91	3.36	4.35	6.49	2.64	-0.10	3.19
	p	(0.88)	(0.86)	(0.85)	(0.69)	(0.39)	(0.36)	(0.18)	(0.08)	(0.05)	(0.03)	(0.05)	(0.47)	(0.85)
						Panel C: 12-Month Forecasting Horizon								
TM	t	-22.49	-12.80	-7.39	-2.53	0.17	0.44	3.49	7.44	9.98	17.49	7.279	0.38	12.67
	p	(0.84)	(0.82)	(0.68)	(0.34)	(0.23)	(0.22)	(0.08)	(0.04)	(0.06)	(0.18)	(0.14)	(0.41)	(0.36)

Table VII

Investment Objectives and Market Timing

Table VII reports the cross-sectional distribution of the Newey-West t-statistics for the holdings-based Treynor-Mazuy measure for four fund groups. "AGG" refers to the aggressive growth funds, "GRO" the growth funds, "G&I" the growth and income funds, and "BAL" the balanced funds. The stock betas are estimated from the past 1-year daily returns. The bootstrapped p-values for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

		1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
						Panel A: 1-Month Forecasting Horizon								
AGG	t	-2.39	-1.16	-0.82	-0.17	0.54	0.54	1.23	1.87	2.48	3.65	1.13	0.05	4.45
	p	(0.12)	(0.00)	(0.00)	(0.01)	(0.02)	(0.02)	(0.05)	(0.13)	(0.10)	(0.20)	(0.79)	(0.46)	(0.37)
GRO	t	-2.97	-1.70	-1.22	-0.54	0.22	0.26	1.01	1.63	2.08	3.21	1.22	-0.35	5.76
	p	(0.24)	(0.15)	(0.15)	(0.22)	(0.29)	(0.25)	(0.29)	(0.41)	(0.49)	(0.60)	(0.86)	(0.80)	(0.52)
G&I	t	-3.03	-1.83	-1.30	-0.60	0.05	0.14	0.76	1.27	1.65	2.84	1.10	-0.33	3.98
	p	(0.35)	(0.30)	(0.24)	(0.25)	(0.47)	(0.33)	(0.60)	(0.83)	(0.90)	(0.77)	(0.96)	(0.78)	(0.72)
BAL	t	-4.46	-2.45	-1.67	-0.81	-0.16	-0.06	0.54	1.36	1.88	2.41	1.35	-1.08	7.61
	p	(0.92)	(0.83)	(0.68)	(0.52)	(0.73)	(0.61)	(0.85)	(0.75)	(0.68)	(0.90)	(0.42)	(0.96)	(0.14)
						Panel B: 3-Month Forecasting Horizon								
AGG	t	-3.52	-1.56	-0.90	-0.11	0.68	0.90	1.53	2.22	2.73	3.33	1.36	-0.76	4.87
	p	(0.92)	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.21)	(0.05)	(0.97)	(0.28)
GRO	t	-3.16	-1.81	-1.10	-0.25	0.66	0.70	1.60	2.36	2.82	3.87	1.43	-0.25	4.27
	p	(0.37)	(0.18)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.14)	(0.15)	(0.81)	(0.96)
G&I	t	-2.96	-1.76	-1.16	-0.25	0.56	0.54	1.43	2.17	2.64	4.34	1.37	0.19	4.17
	p	(0.34)	(0.14)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.16)	(0.27)	(0.77)
BAL	t	-3.56	-2.01	-1.59	-0.59	0.20	0.26	1.10	1.85	2.25	3.51	1.47	-0.70	7.18
	p	(0.78)	(0.49)	(0.61)	(0.09)	(0.13)	(0.07)	(0.05)	(0.07)	(0.22)	(0.27)	(0.06)	(0.95)	(0.17)
						Panel C: 6-Month Forecasting Horizon								
AGG	t	-3.09	-1.33	-0.89	-0.15	0.79	0.94	1.74	2.43	2.92	3.81	1.44	-0.69	4.86
	p	(0.44)	(0.05)	(0.05)	(0.04)	(0.02)	(0.01)	(0.02)	(0.03)	(0.06)	(0.21)	(0.22)	(0.95)	(0.27)
GRO	t	-3.58	-2.10	-1.40	-0.20	0.79	0.82	1.84	2.88	3.53	4.86	1.77	-0.20	6.74
	p	(0.59)	(0.47)	(0.35)	(0.06)	(0.04)	(0.03)	(0.02)	(0.03)	(0.04)	(0.09)	(0.07)	(0.73)	(0.22)
G&I	t	-5.15	-2.01	-1.45	-0.32	0.62	0.72	1.68	2.67	3.21	4.43	1.76	-0.68	6.00
	p	(0.96)	(0.40)	(0.36)	(0.07)	(0.04)	(0.02)	(0.01)	(0.01)	(0.04)	(0.10)	(0.03)	(0.94)	(0.19)
BAL	t	-5.62	-2.49	-1.76	-0.61	0.33	0.51	1.48	2.57	2.94	4.17	1.99	-1.91	13.95
	p	(0.99)	(0.75)	(0.64)	(0.25)	(0.17)	(0.07)	(0.05)	(0.04)	(0.10)	(0.16)	(0.01)	(0.99)	(0.03)
						Panel D: 12-Month Forecasting Horizon								
AGG	t	-4.35	-2.31	-1.46	-0.12	0.71	0.76	1.77	2.73	3.43	4.25	1.78	-0.92	6.73
	p	(0.62)	(0.40)	(0.23)	(0.03)	(0.08)	(0.06)	(0.07)	(0.10)	(0.10)	(0.41)	(0.23)	(0.97)	(0.11)
GRO	t	-5.00	-2.50	-1.51	-0.38	0.95	0.91	2.18	3.67	4.81	6.84	2.29	-0.38	8.37
	p	(0.71)	(0.42)	(0.25)	(0.14)	(0.07)	(0.05)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)	(0.83)	(0.11)
G&I	t	-6.79	-2.63	-1.71	-0.41	0.71	0.86	1.95	3.10	3.96	6.28	2.26	-0.42	7.90
	p	(0.97)	(0.50)	(0.32)	(0.11)	(0.07)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.03)	(0.84)	(0.12)
BAL	t	-6.33	-3.21	-1.90	-0.66	0.65	0.68	1.92	3.48	4.39	5.90	2.29	-0.36	4.29
	p	(0.95)	(0.75)	(0.45)	(0.21)	(0.08)	(0.06)	(0.04)	(0.02)	(0.04)	(0.09)	(0.05)	(0.78)	(0.54)

Table VIII

Fund Characteristics and Market Timing

Table VIII reports the estimated coefficients of the following panel data model:

$$\hat{\beta}_{it} = a_i + (b_0 + b_1\text{Turnover} + b_2\text{Load} + b_3\text{Expense} + b_4\text{RET} + b_5\ln(\text{TNA}))r_{mt} + e_{it}$$

where $\hat{\beta}_{it}$ is the beta estimate of fund i with stock betas estimated using the past 12-month daily returns. The forecasting horizons are 1-, 3-, 6-, and 12-month. Turnover is the annual fund turnover. Load is the annual fund load. Expense is the annual expense ratio, RET is the average annual return. Ln(TNA) is the natural log of total net assets. Turnover, Load, Expense, RET, and TNA used in the panel data model are the time series averages for each fund. The t-statistics are computed following a robust procedure and reported in the parentheses next to the coefficient estimates.

Forecasting Horizon	Constant	Turnover	Load	Expense	RET	ln(TNA)
1-Month	0.31 (0.46)	0.76 (1.37)	-1.71 (-2.46)	0.70 (1.23)	2.69 (2.18)	-0.33 (-0.41)
3-Month	0.06 (0.38)	0.42 (1.90)	-1.46 (-3.01)	0.10 (0.97)	1.01 (2.08)	0.11 (1.25)
6-Month	-0.07 (-0.30)	0.25 (1.74)	-1.20 (-2.42)	0.34 (2.22)	1.40 (1.94)	0.20 (1.28)
12-Month	0.17 (0.27)	0.10 (1.80)	-1.19 (-2.26)	0.21 (2.25)	1.77 (2.26)	0.20 (1.86)

Table IX

Conditional Information and Market Timing

Table IX reports the cross-sectional distribution of the Newey-West t-statistics for the γ estimates from the following regression:

$$\hat{\beta}_t = a + bM_{t-1} + \gamma r_{mt} + e_t$$

The regression is performed for each fund, where M_{t-1} is a vector of five economic variables – Short, Term, Credit, D/P and E/P. Fund betas are computed using stock beta estimates based on the past 1-year daily returns. The bootstrapped p-values for the Newey-West t-statistics are reported in the parentheses underneath. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

	1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
					Panel A: 1-Month Forecasting Horizon								
t	-3.14	-1.88	-1.36	-0.61	0.20	0.20	1.00	1.71	2.27	3.55	1.35	-0.02	8.61
p	(0.35)	(0.32)	(0.27)	(0.23)	(0.22)	(0.21)	(0.18)	(0.24)	(0.20)	(0.27)	(0.43)	(0.50)	(0.44)
					Panel B: 3-Month Forecasting Horizon								
t	-2.73	-1.68	-1.20	-0.46	0.41	0.47	1.24	1.96	2.45	3.50	1.33	-0.27	7.54
p	(0.06)	(0.10)	(0.09)	(0.07)	(0.04)	(0.02)	(0.03)	(0.06)	(0.09)	(0.29)	(0.53)	(0.72)	(0.55)
					Panel C: 6-Month Forecasting Horizon								
t	-3.34	-2.11	-1.45	-0.56	0.41	0.46	1.36	2.23	2.84	4.16	1.54	-0.10	4.69
p	(0.65)	(0.70)	(0.46)	(0.17)	(0.04)	(0.02)	(0.01)	(0.00)	(0.01)	(0.03)	(0.03)	(0.61)	(0.94)
					Panel D: 12-Month Forecasting Horizon								
t	-4.57	-2.86	-2.08	-0.80	0.35	0.32	1.48	2.71	3.48	5.69	2.05	0.70	10.28
p	(0.90)	(0.89)	(0.91)	(0.57)	(0.07)	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.12)	(0.33)

Table X

Potential Economic Value of Market Timing

Table X reports the cross-sectional distribution of the economic value of market timing and the scaled value of timing $\hat{V} = \hat{\gamma}(e^{\sigma_m^2} - 1)$, where $\hat{\gamma}$ is the holdings-based Treynor-Mazuy measure with stock betas estimated using the past 1-year daily returns, and σ_m^2 is the variance estimate of the continuously-compounded market returns. $\hat{V}$ is annualized and in percentage terms. $\hat{V}/\sigma_v$ is the scaled value of timing, where σ_v is the standard deviation of bootstrapped $\hat{V}$s. The numbers in the parentheses underneath are the bootstrapped p-values. "Stdev", "Skew", and "Kurto" denote the standard deviation, skewness, and kurtosis, respectively.

	1%	5%	10%	25%	Mean	Median	75%	90%	95%	99%	Stdev	Skew	Kurto
Panel A: 1-Month Forecasting Horizon													
$\hat{V}(\%)$	-37.22	-13.26	-8.01	-2.24	3.37	0.84	5.21	12.44	18.61	60.40	76.01	40.16	1870
p	(0.17)	(0.20)	(0.23)	(0.18)	(0.14)	(0.18)	(0.18)	(0.20)	(0.28)	(0.17)	(0.57)	(0.03)	(0.05)
$\hat{V}/\sigma_v$	-2.53	-1.63	-1.12	-0.46	0.14	0.19	0.79	1.34	1.66	2.36	0.99	-0.27	3.50
p	(0.79)	(0.60)	(0.37)	(0.23)	(0.26)	(0.20)	(0.26)	(0.29)	(0.34)	(0.34)	(0.41)	(0.95)	(0.17)
Panel B: 3-Month Forecasting Horizon													
$\hat{V}(\%)$	-18.99	-6.94	-3.69	-0.513	2.85	1.22	3.93	7.78	11.34	34.45	30.43	21.66	661.9
p	(0.10)	(0.14)	(0.08)	(0.02)	(0.06)	(0.03)	(0.06)	(0.14)	(0.22)	(0.17)	(0.88)	(0.22)	(0.65)
$\hat{V}/\sigma_v$	-2.10	-1.27	-0.86	-0.19	0.37	0.45	1.00	1.41	1.71	2.30	0.92	-0.46	3.51
p	(0.37)	(0.13)	(0.06)	(0.02)	(0.05)	(0.04)	(0.10)	(0.23)	(0.28)	(0.35)	(0.59)	(1.00)	(0.10)
Panel C: 6-Month Forecasting Horizon													
$\hat{V}(\%)$	-16.00	-5.95	-3.11	-0.41	3.01	1.14	3.59	7.12	9.94	39.55	28.76	22.73	633.9
p	(0.17)	(0.23)	(0.14)	(0.04)	(0.03)	(0.03)	(0.05)	(0.09)	(0.17)	(0.01)	(0.70)	(0.20)	(0.63)
$\hat{V}/\sigma_v$	-2.22	-1.43	-1.00	-0.20	0.41	0.55	1.10	1.53	1.79	2.59	1.02	-0.52	3.93
p	(0.59)	(0.37)	(0.23)	(0.05)	(0.06)	(0.04)	(0.08)	(0.16)	(0.21)	(0.12)	(0.30)	(1.00)	(0.02)
Panel D: 12-Month Forecasting Horizon													
$\hat{V}(\%)$	-13.61	-5.04	-3.02	-0.44	2.13	0.90	2.81	5.78	8.06	26.10	21.33	19.27	546.0
p	(0.16)	(0.19)	(0.20)	(0.08)	(0.08)	(0.07)	(0.15)	(0.21)	(0.31)	(0.13)	(0.73)	(0.23)	(0.67)
$\hat{V}/\sigma_v$	-2.07	-1.35	-0.91	-0.22	0.33	0.48	1.00	1.33	1.54	2.00	0.90	-0.67	3.66
p	(0.54)	(0.33)	(0.17)	(0.09)	(0.13)	(0.08)	(0.16)	(0.34)	(0.44)	(0.51)	(0.55)	(1.00)	(0.04)

Table A.I

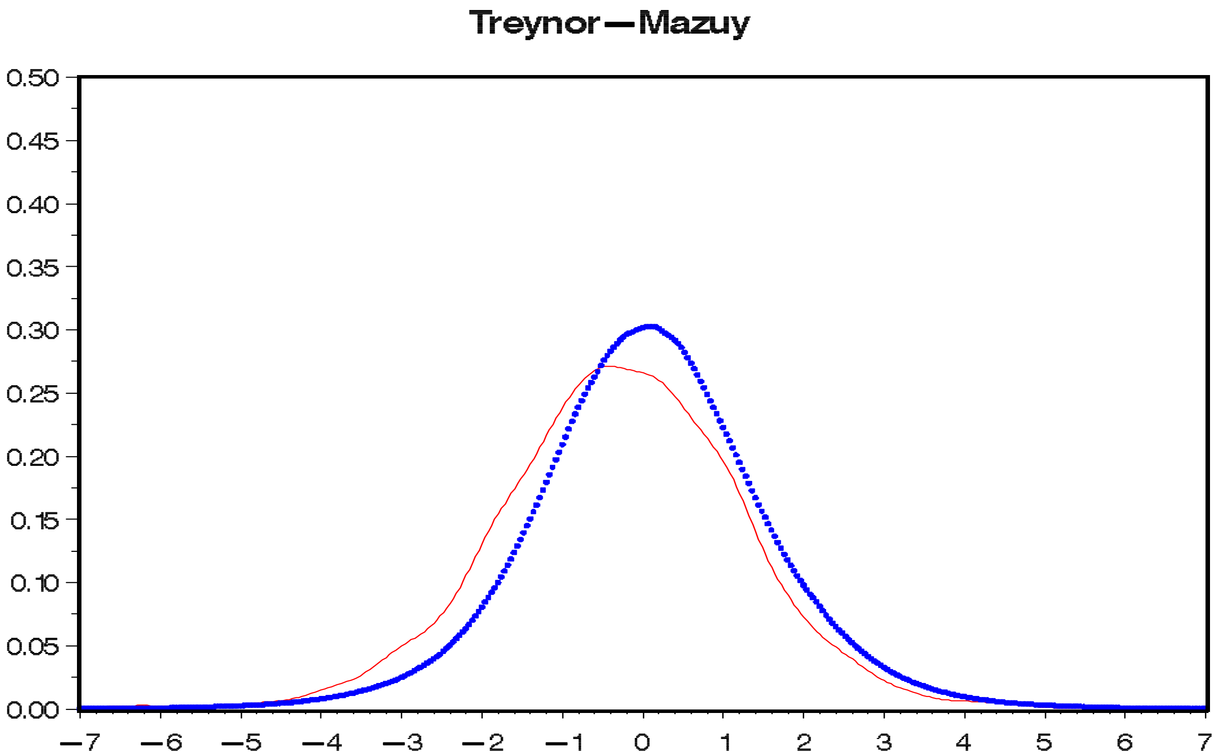
Simulation Results of Return-based versus Holdings-based Measures

Table A.I reports the results of the simulation detailed in Appendix A for both return-based and holdings-based timing measures. We simulate daily returns according to (A.1), the market timing signals according to (A.2), and obtain corresponding fund betas and fund returns. We then compute the return-based Treynor-Mazuy timing measure (RB), and three holdings-based measures with stock betas estimated using (1) the past 1-year daily returns (HB1), (2) the past 5-year monthly returns (HB2), and (3) the past 3-month daily returns (HB3). The simulation is performed with different values for the timing measure γ and the noise-signal ratio σ_s/σ_m. The mean and standard deviation (in the parentheses below the mean) of various timing measures with 10,000 replications are reported.

γ	0.10				0.36				1.35				3.11			
$\frac{\sigma_s}{\sigma_m}$	RB	HB1	HB2	HB3	RB	HB1	HB2	HB3	RB	HB1	HB2	HB3	RB	HB1	HB2	HB3
0.25	0.092	0.100	0.099	0.100	0.347	0.359	0.359	0.360	1.316	1.349	1.349	1.350	3.040	3.108	3.108	3.109
	(0.72)	(0.07)	(0.16)	(0.12)	(0.71)	(0.08)	(0.16)	(0.12)	(0.82)	(0.09)	(0.17)	(0.18)	(1.11)	(0.15)	(0.23)	(0.35)
0.50	0.092	0.100	0.099	0.100	0.353	0.360	0.359	0.360	1.334	1.350	1.349	1.351	3.068	3.111	3.108	3.112
	(0.72)	(0.08)	(0.16)	(0.12)	(0.71)	(0.08)	(0.16)	(0.12)	(0.82)	(0.12)	(0.19)	(0.20)	(1.14)	(0.22)	(0.28)	(0.39)
1.00	0.094	0.100	0.099	0.099	0.358	0.359	0.360	0.361	1.330	1.350	1.350	1.350	3.051	3.108	3.107	3.108
	(0.72)	(0.08)	(0.16)	(0.12)	(0.72)	(0.10)	(0.17)	(0.13)	(0.82)	(0.18)	(0.23)	(0.24)	(1.20)	(0.38)	(0.42)	(0.50)
2.00	0.097	0.099	0.099	0.099	0.348	0.361	0.360	0.360	1.331	1.354	1.357	1.356	3.056	3.101	3.102	3.102
	(0.72)	(0.08)	(0.16)	(0.12)	(0.73)	(0.11)	(0.18)	(0.15)	(0.90)	(0.33)	(0.36)	(0.37)	(1.45)	(0.74)	(0.77)	(0.81)
4.00	0.098	0.100	0.100	0.100	0.357	0.360	0.361	0.360	1.312	1.351	1.350	1.348	3.053	3.112	3.110	3.112
	(0.73)	(0.09)	(0.17)	(0.13)	(0.76)	(0.19)	(0.24)	(0.21)	(1.12)	(0.65)	(0.67)	(0.68)	(2.17)	(1.49)	(1.50)	(1.52)

Figure 1: The Density of t-statistics for the Return-based Timing Measures

Figure 1 plots the density of the cross-sectional distribution of the t-statistics for the return-based timing measures estimated from all funds, together with that of the bootstrapped t-statistics (the dotted line).



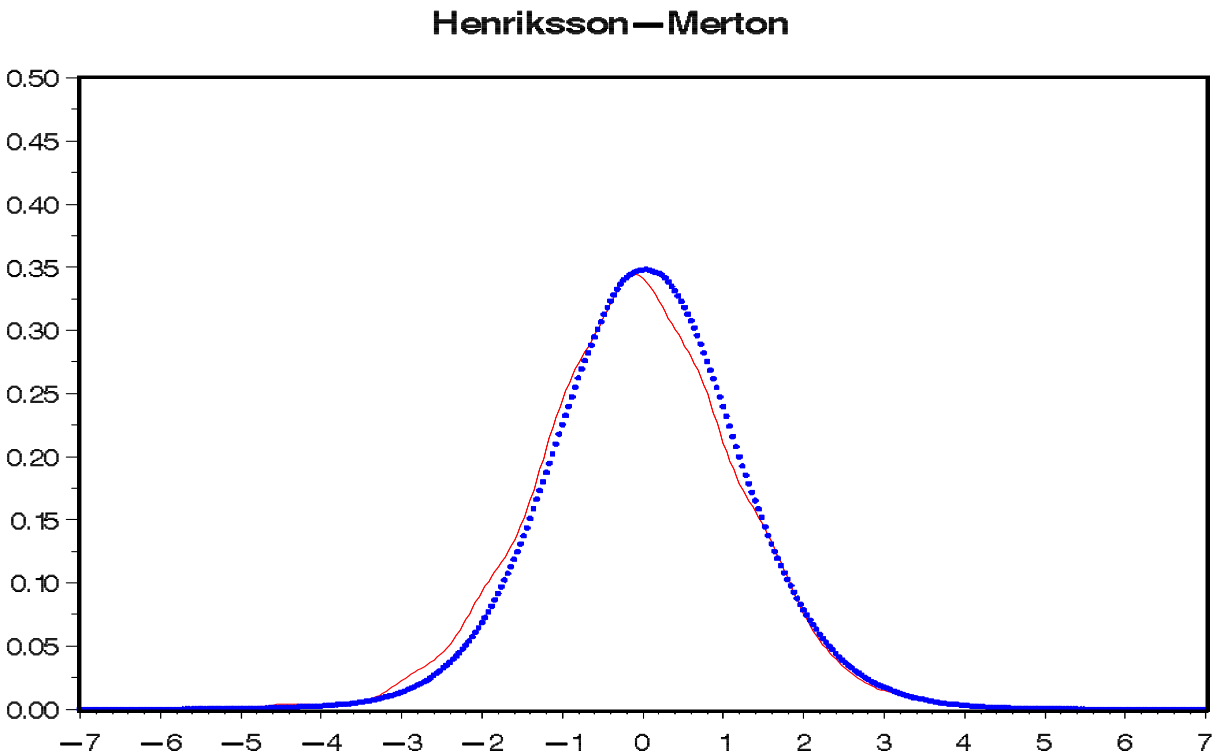
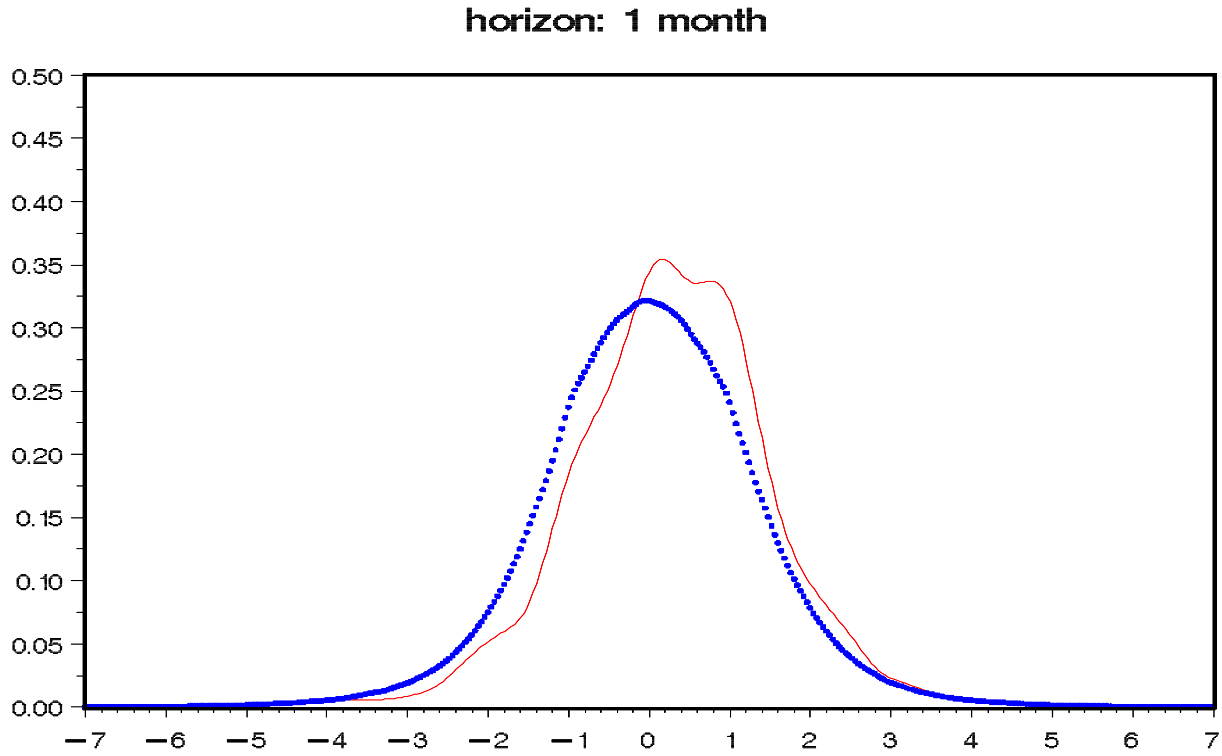
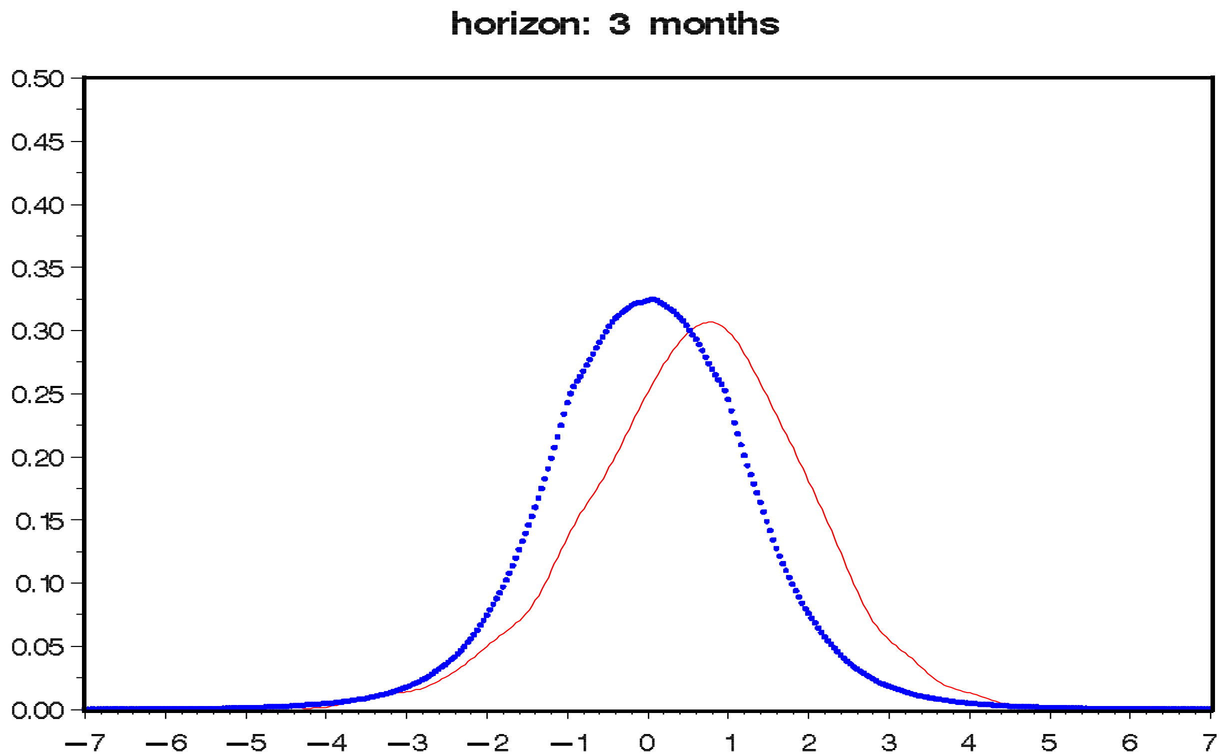


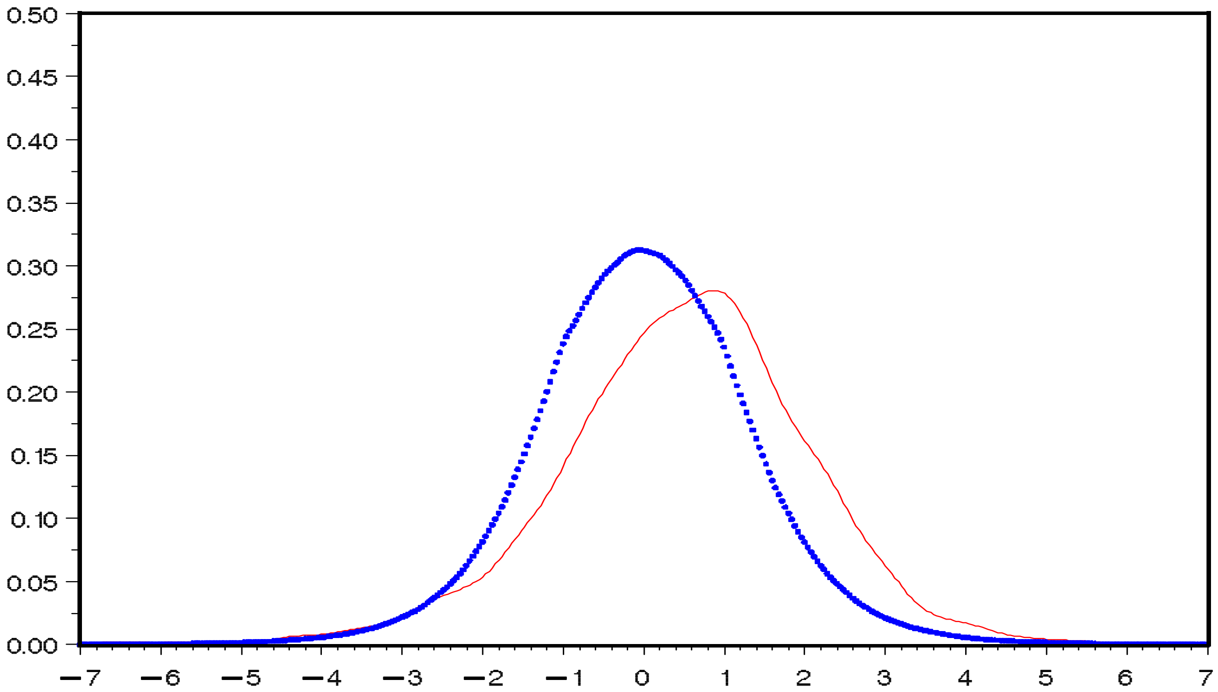
Figure 2: The Density of t-statistics for the Holdings-based Timing Measures

Figure 2 plots the density of the cross-sectional distribution of the t-statistics for the holdings-based timing measures estimated from all funds with 1-, 3-, 6-, and 12-month forecasting horizons, together with that of the bootstrapped t-statistics (the dotted line).





horizon: 6 months



horizon: 12 months