

Journal of Financial and Quantitative Analysis (forthcoming)

**Holdings Data, Security Returns  
and the Selection of  
Superior Mutual Funds<sup>+</sup>**

**Edwin J. Elton\***  
**(eelton@stern.nyu.edu)**

**Martin J. Gruber\***  
**(mgruber@stern.nyu.edu)**

**Christopher R. Blake\*\***  
**(cblake@fordham.edu)**

**September 14, 2009**

\* Nomura Professor of Finance, Stern School of Business, NYU

\*\* Joseph Keating, S.J., Distinguished Professor, Fordham University

<sup>+</sup> We would like to thank Ken French for supplying us with weekly data on his factors.

# **Holdings Data, Security Returns and the Selection of Superior Mutual Funds**

## **Abstract**

In this paper we show that selecting mutual funds using alpha computed from a fund's holdings and security betas produces better future alphas than selecting funds using alpha computed from a time series regression on fund returns. This is true whether future alphas are computed using holdings and security betas or a time series regression on fund returns. Furthermore, we show that the more frequently the holdings data are available, the greater the benefit. This has major implications for the SEC's recent ruling on the frequency of holdings disclosure and the information plan sponsors should collect from portfolio managers. We also explore the effect of conditioning betas on macro variables as suggested by Ferson and Schadt (1996) to identify superior-performing mutual funds as well as the alternative way of employing holdings data proposed by Grinblatt and Titman (1993).

JEL Classification G11, G12

Keywords: mutual funds, portfolios, composition, performance prediction

## I. Introduction

There have been a number of studies that use mutual fund return data to compute performance measures and which show that funds that rank high on these performance measures perform well in the future.<sup>1</sup> In this paper we examine whether computing fund alphas and betas using fund holdings data and security returns can be used to identify funds that will outperform other mutual funds, outperform passive portfolios, and outperform mutual funds selected using alpha from a time-series regression on fund returns. Two measures that we will use to rank funds are the alpha from the Fama and French three-factor model and the alpha from the four-factor Carhart model. Why might estimating alphas and betas using holdings data and security returns produce better estimates of betas and alphas than simply using the traditional method of obtaining betas and alphas by regressing past mutual fund returns on a set of indexes? Portfolio betas (mutual fund betas) are a weighted average of security betas. Thus we can compute the mutual fund betas from individual security betas using holdings data at a point in time. These are the same betas that would be computed by a time series regression on fund returns if the fund had held constant composition with the weights at the time of measurement. A big advantage of this technique compared to a time series regression of fund returns on indexes is that the fund betas will not be distorted by changes in the betas on the portfolio caused by changes in the composition of the portfolio over time.

On the other hand, it can be argued that if management attempts to hold risk constant over time and uses information not captured by historical security betas to do so, then using betas calculated from portfolio returns might forecast future portfolio betas and alphas better than estimates based on aggregating betas on individual securities.

---

<sup>1</sup> See, for example, Elton, Gruber and Blake (1996), Gruber (1996), Zheng (1999), Busse and Irving (2006), and Mamaysky, Spiegel and Zhang (2008).

Whether estimating betas from the securities in a portfolio or from mutual fund returns is better is an empirical question; one which we answer in this paper.

This paper contributes to the literature of financial economics by examining the following questions:

- (1) Does the use of holdings data and security returns to estimate betas and alphas lead to a better selection of funds than estimating alphas and betas from fund returns?
- (2) Does the three- or four-factor model lead to better ranking?
- (3) What is the effect of different frequencies of reporting holdings data, and how frequent should holdings data be reported to get most of the benefit from correctly selecting mutual funds?
- (4) Do conditional betas improve the ranking?
- (5) How does the alternative use of holdings data proposed by Grinblatt and Titman perform?

Our primary results concerning selection are that estimating portfolio alphas using betas computed from security returns and holdings data allows us to rank funds on alpha such that 1) the top funds formed from these rankings outperform index funds over the following year, 2) the performance in the next year of deciles formed on prior rankings is almost perfectly correlated with the prior ranking, 3) the performance in the top decile is considerably higher than that found by others in past research, 4) the performance is better than if funds are selected using a time series of fund returns, and 5) selection based on the three-factor model leads to higher future alphas than selections based on the four-factor model.

(6) Neither conditional betas nor the Grinblatt and Titman measure result in better rankings.

It is important to note that these results hold whether we evaluate performance in the ex-post period using alphas estimated from security returns and holdings or estimated from a standard time series regression of fund returns on factors.<sup>2</sup>

This brings us to the question of the appropriate frequency for reporting holdings. There has been a lot of discussion about how often funds should be required to report holdings data. In the 1970's and 1980's, funds were required to report holdings on a quarterly basis. Subsequently, funds were only required to report semi-annually, but recently the requirements were changed back to quarterly. The decision on how often funds should be required to report requires an analysis of the costs and benefits of more frequent reporting.<sup>3</sup> This article contributes to this debate by examining one benefit to investors of having holdings data at more frequent intervals and measures the size of the benefit. We show that ranking funds on the basis of alpha computed using monthly holdings data leads to better ex-post alphas than ranking on quarterly holdings and that both rankings are substantially better than ranking based on a time series regression on fund returns.<sup>4</sup> Demonstrating that monthly holdings data lead to a better ranking of funds than do quarterly holdings has important implications for investor behavior. The availability of monthly holdings data has been growing and is now available for about

---

<sup>2</sup> There is another body of literature that examines the use of holding-period data to evaluate fund performance. That literature is summarized effectively in Wermers (2006). We chose to principally study holding-period data using multi-factor models because the multi-factor approach remains the dominant approach in both the literature of financial economics and in industry practice. Thus, improving that approach has great import. We do examine one of the measures due to Grinblatt and Titman in a later section.

<sup>3</sup> Several benefits of more frequent reporting in addition to those discussed in this article are analyzed in Elton, Gruber, Krasny and Ozelge (2007).

<sup>4</sup> Quarterly reporting misses about 18% of the mutual fund trades (see Elton, Gruber, Krasny and Ozelge (2006)).

18% of all domestic mutual funds. Our results indicate that when available, such information should be used. The results of this paper mean that, at least for selecting funds, public policy should encourage the move towards monthly reporting of holdings. In addition, administrators of institutions such as pension funds (plan sponsors) should be able to obtain monthly holdings data from the managers who run these funds, and they should do so. Finally, we show that the use of quarterly holdings data captures a large part of the benefit of using monthly holdings data and that it dominates the use of semi-annual and annual holdings data.

The paper is divided into seven sections. In the second section we discuss our samples of monthly and quarterly holdings data. In the third section we discuss and present evidence on whether using holdings data can improve the selection of desirable mutual funds. This section is divided into three subsections: differences in beta using various time frames and techniques, differences in estimated alpha, and the ability to select desirable mutual funds. As part of this section, we explore how frequently holdings data need to be reported in order to be useful in selecting funds. The fourth section examines mutual fund selection using two other ranking techniques, alpha using conditional betas and the Grinblatt and Titman measure, which uses holdings data directly to judge performance. The fifth section examines for a much larger sample the use of quarterly holdings data versus mutual fund return data to select superior mutual funds. Requiring only quarterly rather than monthly holdings data allows us to repeat many aspects of the earlier analysis of this paper on a much larger sample. The sixth section discusses potential sources of bias in our study. The final section contains our conclusions.

## **II. Sample**

Data on the holdings of individual mutual funds were obtained from Morningstar. Morningstar supplied us with all of its holdings data for all domestic (U.S.) stock mutual funds they followed during the period 1994 to 2005. The data are free of survivorship bias, for once a fund enters the database for the first time it remains in the database until it ceases to exist. Note that until 2008 Morningstar was the source for the CRSP data.

The data reported by Morningstar has two limitations. Morningstar, until recently, reported only the largest 199 holdings of any mutual fund. This had little impact, since most funds that held more than 199 securities were index funds, and these were eliminated from our sample.<sup>5</sup> In addition, Morningstar does not report holdings of any security that represents less than 0.006 percent of a portfolio. This had virtually no effect on our sample, since the sum of the weights of the holdings on which we had data almost always equaled one, and, in the few cases where it was less than one, the differences were tiny.

Previous studies of holdings data have used the Thomson database as the source of holdings data. The Morningstar holdings data are much more complete. Unlike Thomson data, Morningstar data include not only holdings of traded equity, but also holdings of bonds, options, futures, preferred stock, non-traded equity and cash. Studies of mutual fund behavior from the Thomson data base ignore changes across asset categories such as the bond/stock mix and imply that the only risk parameters that matter are those estimated from traded equity securities.<sup>6</sup>

---

<sup>5</sup> We had to eliminate 18 funds from our initial sample because they held more than 199 securities and were not index funds.

<sup>6</sup> The Thompson database only includes data on mutual fund holdings in common stock. The difference in alpha from ignoring the portion of a portfolio which is not invested in common stock (e.g., invested in futures, options, preferred stock and bonds) is large on average and very large for individual funds. To analyze the effect of this, we use data in 2004. Over 80% of our sample had investments in securities other

From the Morningstar data we constructed two samples, one to examine the differences in results caused by the frequency of reporting, and one to examine over a much larger sample the effect of differences in beta estimation. In the first we selected all funds that reported at least two consecutive years of monthly holdings at any time after January 1998 and that reported holdings in the December prior to the start of the sample period for each fund (giving us a minimum of 25 months of data for 436 funds).<sup>7</sup> We eliminated all index, international, real estate, and specialty funds from our sample (71 funds). We then eliminated all funds that had less than 93% of their assets in cash plus stock in any month or held options or futures that represented more than 0.5% of the value of their portfolio (124 funds) in any month. This resulted in all but six funds in our sample averaging more than 99% in stock and cash and all but 16 averaging more than 99.5%. Thus our first sample consists of funds that invest almost exclusively in equity and cash and where beta estimation problems are minimal. As described below, our second sample is much larger and includes mutual funds with many more non-equity securities, such as options and futures. The first sample is interesting, for many pension plan sponsors hold funds that restrict their investments to domestic equities and cash. For individuals, this is a feasible constraint to implement. Finally, we eliminated funds that existed for less than a year prior to our first month of holdings data (26 funds). This was

---

than common stock and cash. The difference in the absolute value of alpha for the average fund holding other securities is 33 basis points. Examining the 10% of the funds with the largest difference from ignoring other securities, the average difference in alpha is 1.09% per year.

<sup>7</sup> The data included monthly holdings data for only a very small number of funds before 1998, so we started our sample in that year. In 1998, 2.5% of the common stock funds reporting holdings to Morningstar reported those holdings for every month in that year. By 2004, the percentage had grown to 18%.

necessary in order to estimate betas using the funds' returns. Our first sample consists of 215 funds and 317 pairs of years.

Our analysis of the first sample indicates that bottom-up beta leads to better selection of funds and that the use of quarterly holdings data provides much of the benefit of using monthly holdings data. Quarterly holdings allow us to use a much larger sample and examine whether our conclusions hold with a reporting interval required presently of all funds. To see if our analysis applied to a wide cross-section of funds, we limited the filters.

We did apply the following filters to our second sample:

1. Quarterly data reported for at least two years.
2. Eliminated specialty, international, index and real estate funds.
3. The aggregate asset value of the reported holdings data was within 0.1% of the aggregate asset value reported by the fund.<sup>8</sup> After applying this filter we are left with 1,255 funds.

In addition, we use CRSP and Morningstar data and data from the International Center for Finance at Yale University for weekly and monthly return data on funds, along with the returns on factors (described later in this paper) as compiled by Ken French and available on his web site.<sup>9</sup>

### **III. Measuring Performance (Sample 1)**

In this section and the following section we compare selecting mutual funds using alpha computed with betas estimated both from holdings data and from mutual fund

---

<sup>8</sup> Twenty-six out of 1,578 funds were eliminated using this filter. This captures all funds that hold more than 199 securities and are not index funds, as well as gross errors in reporting. At the end of our sample, Morningstar increased the number of securities reported.

<sup>9</sup> Weekly return series were constructed by compounding daily returns.

returns. We refer to the first type as “bottom-up” betas and the second as “top-down” betas. The literature of financial economics contains extensive discussion of mutual funds performance based on comparing the return on a mutual fund to the returns on a set of indexes or factors that spans the types of securities the mutual fund holds. The most frequently used multi-factor measure of portfolio performance is the three-factor model developed by Fama and French. We use the Fama and French model, as follows:<sup>10</sup>

$$(1) \quad R_{it} - R_{ft} = \alpha_i + \beta_{iM} I_{Mt} + \beta_{iSMB} I_{SMBt} + \beta_{iHML} I_{HMLt} + \varepsilon_{it}$$

where

$R_{it}$  is the return on mutual fund  $i$  in period  $t$

$R_{ft}$  is the risk-free rate in period  $t$

$I_{Mt}$  is the excess return on the market (above the risk-free rate) in period  $t$

$I_{SMBt}$  is the return on the “small minus big” (SMB) factor in period  $t$

$I_{HMLt}$  is the return on the “high minus low” (HML) book-to-market factor in period  $t$

$\beta_{ik}$  is the sensitivity of fund  $i$  to the  $k$ th factor

$\alpha_i$  is the excess return on portfolio  $i$  above that which can be earned on a portfolio of the three factors that has the same risk.

To make this study comparable to research using Carhart’s (1997) model, we also included a momentum factor obtained from Ken French’s web site. The results were often similar and are discussed where interesting in footnotes and text.

The standard way to implement a model of this type is to use regression analysis on the time series of returns for a mutual fund. We refer to this estimation technique as “top down.”

---

<sup>10</sup> The composition of the factors is described on Ken French’s web site.

A possible problem with this approach is that to the extent that management changes the sensitivity of a portfolio to any of the model's factors, e.g., changes market sector, industry, or security exposure, the sensitivities (betas) can be seriously mis-estimated. Given a change in risk exposure, the betas that are produced can be quite different from the betas that exist at a moment in time or the average beta over time.<sup>11</sup>

One way to avoid this problem is to use the actual composition of a mutual fund to estimate its beta at a point in time by aggregating the estimated betas on each of the fund's underlying assets (bottom-up betas). This is appropriate, since we know from portfolio theory that the betas on the portfolio (mutual fund) are a weighted average of the betas on the securities that comprise it. Furthermore, using this approach, the estimated portfolio betas are exactly the same as those that would be obtained from a time series regression on fund returns if the composition of the portfolio was held constant over time. However, there is a potential disadvantage to the bottom-up approach. Suppose that management is attempting to hold portfolio betas constant over time. If they do so using estimated betas obtained by using a time series regression, then the two approaches we examine (top-down and bottom-up) would produce identical results. However, if management attempts to hold the betas on the portfolio constant and has a better way of estimating the betas on stocks than looking at historical regression, then the top-down approach could be the best way to forecast beta. Thus, which approach is best is an empirical question.

Before examining forecasting abilities it is worthwhile to examine what the differences are in beta using the different approaches and different intervals.

#### **A. Estimating Beta**

---

<sup>11</sup> See Dybvig and Ross (1985) for a general discussion, and Elton, Gruber, Brown and Goetzmann (2007) for an example of this phenomenon in the single-index case.

In this section we describe in detail alternative techniques for estimating fund betas as well as tests for comparing the results of these techniques.

## **1. Bottom-Up Estimation**

Our sample allows us to estimate the mutual fund betas from holdings data as frequently as monthly. To do this at any point in time, we estimate a time series regression (equation (1)) using 36 months of past return data for each security in the fund.<sup>12</sup> There are two problems. First, if less than 36 months of data are available, we use as many months of data as are available unless there are less than 12 months available. If we have less than 12 months of data available, we set the beta for the stock equal to the average beta for all other stocks in the portfolio. On average this had to be done for less than 1.4% of the securities in any portfolio. The second problem involves the estimation of equation (1) for securities other than common stock.

For T-bills and bonds with less than one year to maturity we set all betas to zero. For each of the following categories of investments – long-term bonds, preferreds and convertibles – we used an index of that category and obtained estimated betas by running a regression of the category index's returns against the three- or four-factor model. Each bond, convertible or preferred was assumed to have the same beta as the relevant index. Finally, for options and futures we used the beta on the underlying instrument adjusted for the leverage inherent in options and futures.<sup>13</sup>

## **2. Top-Down Fund Returns Estimation**

---

<sup>12</sup> We capped all computed betas at two standard deviations from the mean. We also tried the Vasicek adjustment. The capped betas and the Vasicek betas produced almost identical results.

<sup>13</sup> For example, the adjusted beta for an option is the security beta times  $N(d_1)$  times the market price of the security divided by the market price of the option.  $N(d_1)$  was estimated assuming the option was fairly valued and using the Black Scholes formula.

Mutual fund betas can be approximated by simply running a time series regression of the return for any mutual fund on the factors employed in equation (1). We use the standard 36 months of data in our estimates.<sup>14</sup> To get more frequent estimates we perform the regression each month. We also examine the accuracy of the betas if the regressions are estimated quarterly, semi-annually, or annually.

### **3. Empirical Results**

We are concerned with the usefulness of different ranking techniques on alpha to lead to the selection of superior funds. Before examining the fundamental question, it is useful to examine differences in beta since differences in beta will lead to differences in alpha.<sup>15</sup> We can compute the portfolio betas at the beginning and at the end of the month. If all trades took place near the end of the month, the best estimate of the portfolio beta during the month would be that based on holdings at the beginning of the month. If all trades take place immediately after the beginning of the month, the best estimate would be based on holdings at the end of the month. Since we do not know the timing of trades during the month, we shall use an average of the beginning and end-of-month betas as our estimate of the beta on the portfolio over the month. Later, when we use the bottom-up approach to estimate betas using quarterly, semi-annual or annual periods, we will use as our estimate of beta for each month in the relevant period an average of betas measured at the beginning and end of the relevant period.

Table 1 provides information about the distribution of betas estimated from the bottom-up approach using monthly holding data across the funds in our first sample. The average beta of the funds in our sample is slightly below one with the market factor.

---

<sup>14</sup> If 36 months of data are not available, we estimate equation (1) using the longest time period available that is at least 12 months. We had less than 36 months in only 8% of the cases.

<sup>15</sup> There is a question about the estimation error in computing betas. This is examined in an appendix available from the authors..

However, our sample includes funds with a large spread in their sensitivity to the market factor. When we examine the small-minus-big factor, we see that the average beta is 0.1628, demonstrating a general tendency for funds to hold small stocks. However, over 25% of our funds have a negative beta with the size factor, which indicates they overweight large stocks. Examining the third factor, we see a slight tendency on average to hold value stocks, although again over 25% of the funds overweight growth stocks.<sup>16</sup>

The next question we examine is the stability of betas. If betas do not change, then having more frequent data is not important. In Table 1, row 4, we present the average absolute difference in betas from month to month for all funds in this sample.<sup>17</sup> The surprising result from this table is that the average absolute difference for each of the sensitivities is about the same size, a change from month to month of approximately 0.04. Not only is the average absolute difference in beta from month to month large, but the range of this statistic across funds is quite large (for example, an interquartile range of 0.273 for the market factor beta). Therefore, having frequent measures of the portfolio beta should be important.

---

<sup>16</sup> When we examine the four-factor model (adding a momentum factor), the average coefficients on the first three factors are almost identical, and the average coefficient on the fourth factor is very close to zero. The number of funds that trade on momentum appears similar to the number of funds that trade against it.

<sup>17</sup> We examined the causes of the change in the market factor beta in some detail. Approximately 50% of the beta change for mutual funds is due to weights changing due to differential returns on the various securities (passive weight change). More mutual funds change weights through security trading to exacerbate the effect of the passive change rather than trade securities to mitigate the effect (about 50% more). The percentage change in betas due to passive and active trades is 74% to 89% depending on the beta, which means 11% to 26% is due to changing estimates of security betas. It is possible that mutual funds react to passive weight changes with a lag. We repeated the analysis using 3-month intervals and got similar results.

We next examine whether betas estimated from different techniques are different from each other. It is simplest to compare the results of applying different techniques to a single technique. Since monthly estimates of bottom-up betas should produce better estimates than bottom-up betas measured at longer intervals, this is a natural metric for bottom-up betas and we select this as our standard for examining all techniques.

Denote the bottom-up monthly beta for fund  $i$  on factor  $k$  at time  $t$  as  $\beta_{i,k,t}$  and an estimate from any other method as  $b_{i,k,t}^m$  where the superscript  $m$  and lower case  $b$  signify that an alternative estimation method is used.

For all alternative methods we compute the average difference and the average absolute difference. The average difference for technique  $m$  is

$$(2) \quad \bar{D}_{i,k}^m = \frac{1}{T} \sum_{t=1}^T (b_{i,k,t}^m - \beta_{i,k,t})$$

The average absolute difference is

$$(3) \quad AAD_{i,k}^m = \frac{1}{T} \sum_{t=1}^T |b_{i,k,t}^m - \beta_{i,k,t}|$$

Table 2 shows the average difference between the beta for each forecasting technique and the bottom-up beta estimated using monthly holdings (equation (2)). The average difference between computing bottom-up betas using quarterly, semi-annual or annual holdings rather than monthly holdings is very small and insignificantly different from zero. The average differences for top-down betas are larger but only significantly so for the high-minus-low factor.

When we examine average absolute differences in Table 3, we see, as expected, larger differences from betas estimated monthly when betas are estimated at longer intervals. Moving from quarterly to semi-annual intervals increases the average absolute difference in bottom-up betas by more than 50% and moving to annual from semi-annual

results in another 50% increase in absolute difference for each of the betas in our model. All of the differences from adjacent intervals for bottom-up betas (e.g., quarterly versus semi-annual) are statistically significant at the 0.05 level. The differences from the time series regression are much larger, more than four times the difference from using quarterly holdings data, and are statistically different from the estimates using holdings data. Given these differences, the techniques will rank differently, with the biggest difference being between the top-down and bottom-up techniques. There is no difference in top-down absolute differences whether we update the regression monthly, quarterly, semi-annually or annually. Thus, in what follows we will follow tradition and measure the top-down betas revising the regression at annual intervals.

## **B. Differences in Measurement of Alpha**

In this section we will discuss the magnitude of the difference in estimates of portfolio performance (alpha) caused by the different estimates of sensitivity.

Assume we wish to calculate alpha over a year. With monthly holdings data we will calculate each month's alpha using an average of the sensitivities derived from holdings data at the beginning and end of each month. There will be a different set of betas each month. With quarterly weights we will calculate alphas each month over a quarter using sensitivities computed as the average of the beginning and end of quarter holdings betas. Sensitivities will be fixed over six months when semi-annual holdings data are used and 12 months when annual data are used. Once the sensitivities are computed, the monthly alpha will be calculated as the difference between the fund's return and the benchmark return. The benchmark return is the sensitivities times the realized returns on the factors plus the riskless rate. The average monthly alpha is simply the sum of the monthly alphas divided by twelve.

We will estimate top-down betas using the regression of fund returns on factors. To be consistent with normal practice, and given that in Table 2 there was little difference in errors over different intervals, we will estimate sensitivities each year by running a three-year regression of fund returns on the factors including data from the year over which we are computing alpha. These betas are then used to compute each month's alpha over the year.

Given that we found differences in sensitivities, we will find differences in alphas. However, the real issue for selecting mutual funds is whether the techniques rank funds in the same way. The fact that bottom-up techniques rank similarly is seen in Table 4, which shows the rank correlation across techniques. The rank correlation across the bottom-up techniques is very high, ranging between 0.994 between monthly and quarterly to 0.968 between monthly and annual. However, the correlation between alphas using bottom-up betas and alphas using top-down betas is not nearly as high. The rank correlation between bottom-up alphas using monthly intervals and top-down alphas is 0.762. Thus, using bottom-up or top-down alpha to select funds will result in different ranking.<sup>18</sup>

### **C. Forecasting**

While we have been examining differences in rankings of past performance, the principal purpose (and some would argue the only important use) of performance measurement is selecting funds that will do well in the future. In this section we will examine whether alternative measures of past performance predict future performance

---

<sup>18</sup> Examining the composition of deciles across techniques shows very different membership when bottom-up is compared to top-down ranking. This is especially prevalent in the top decile. This is further evidence that the different techniques are likely to exhibit very different performance as ranking devices.

and whether any technique can be used to select active funds that outperform passive funds.

In this section we will evaluate subsequent performance using two metrics. First we will use alpha calculated using monthly bottom-up betas. If bottom-up betas produce the best estimate of alpha, more frequent calculation should be more accurate. Second, we will use alpha calculated from a time series regression on fund returns. This is the most common standard used in the literature.

As discussed earlier, there are conditions under which each of these techniques will produce more accurate determination of present and future alphas. If one ranking (forecasting) produces better future alphas regardless of whether those future alphas are computed on a bottom-up or a top-down approach, then that technique is indeed the better technique.

Each year where we have at least 40 funds, each of the techniques discussed in the previous section is used to rank funds. For each technique we will rank funds using the average alpha over the year. For each ranking criterion, funds will be divided into quintiles based on those average alphas. The ranking techniques include alphas computed using monthly, quarterly, semi-annual and annual bottom-up measures and the top-down annual measure.<sup>19</sup> Because of the wide use of the four-factor model, we will not only rank using the three-factor model but also rank using the four-factor model. This yields a total of 20 ranks, 10 for each of the two models. Then the actual subsequent performance (in the evaluation period) for the funds in each decile will be computed where actual performance is defined in two different ways:

1. Alpha from the monthly holdings data (bottom up)
2. Alpha from a time-series regression of the fund return (top down)

---

<sup>19</sup> Later in this paper we will also look at ranking and evaluation using betas conditional on a set of macroeconomic variables.

We compute alpha in the evaluation period in a different manner than was done in prior sections so that there is no overlap between data used in the evaluation and ranking period. We do not have three years of data after the ranking is completed. Because of this, we estimate betas using one year of weekly data in the evaluation period, which is the year following the ranking period. For example, if we rank using data from January 1998 through December 2000, we evaluate using data from January 2001 through December 2001. This gives a reasonable amount of data and no overlap between the ranking and evaluation periods. Weekly mutual fund returns were used to compute top-down betas. For the bottom-up method, we estimated betas over the evaluation year for individual securities from weekly data and computed portfolio betas in any month using the average of beginning and ending weights.<sup>20</sup> The monthly alpha was the fund's excess return (fund return minus 30-day T-bill) less the return on the benchmark portfolio computed using these betas and the factor returns.

For any ranking technique, we examine the probability that the realized alpha on the top quintile is greater than zero and the bottom quintile is below zero, each at a statistically significant level. An alpha greater than zero is a clear indication that active mutual funds that outperform index funds with the same risk can be selected.<sup>21</sup>

Tables 5 and 6 show the average alpha of funds that ranked in the top 20% or the bottom 20% of funds in the prior year by each of our ranking techniques.<sup>22</sup> Table 5 shows

---

<sup>20</sup> In any year, if a security had missing data but at least 26 weeks of complete return data, we estimated beta using the available data. If securities had less than 26 weeks of complete returns, we used the average beta on the portfolio. For securities other than common equity we used the same calculations as described in earlier sections of this paper, but applied to weekly data during the evaluation year. We also examined Performance using evaluation alphas computed using beginning-of-year weights. While individual fund

Alphas varied considerably, the average alpha in the top quintile was virtually unchanged.

<sup>21</sup> Since index funds have expense ratios, finding alphas greater than zero is even stronger evidence.

the results when ranking is done using the three-factor model. Table 6 shows the subsequent alphas when ranking is done using the four-factor model. Examining Tables 5 and 6 together shows that for every ranking technique, for both the three-factor model and the four-factor model and *for every method of measuring alpha in the subsequent (evaluation) period*, funds in the top quintile outperform funds in the bottom quintile. All of these differences are statistically significant at the 1% level.

When we examine different ranking techniques, we see some significant differences in predicting performance. In what follows, we concentrate on the results for the top quintile, since this is the group investors would want to hold.<sup>23</sup> First note that, for bottom-up ranking based on the three-factor model (Table 5), the subsequent bottom-up alphas are positive. The differences from zero are economically significant, ranging from 34 to 180 basis points annually depending on which evaluation model is used and the frequency of the data used. Although the top quintile is positive for all three-factor ranking techniques using holdings data, there are differences that depend on the frequency of the holdings data used to compute rankings. The general tendency is for the average alpha of the top quintile to be higher and have greater statistical significance when more frequent holdings data are used. A big reduction in alpha occurs when we move from monthly to quarterly intervals and from quarterly to semi-annual intervals. Semi-annual and annual alphas are about the same size. This is important, given the controversy over how often mutual funds should be required to report holdings. Recently

---

<sup>22</sup> When funds are divided into deciles, the magnitudes of the top and bottom deciles are very similar to those of the top and bottom quintiles, but due to the smaller sample size the significance is reduced.

<sup>23</sup> In addition, as the table shows, any technique of ranking funds identifies funds that will perform poorly and underperform both the average fund and index funds in subsequent periods. This raises the question of testing a portfolio that holds the top quintile and short sells the bottom quintile. We do not examine this strategy, because it is an infeasible strategy for investors since mutual funds can not be sold short.

the SEC shifted to requiring funds to report holdings quarterly rather than semi-annually. The higher alpha in the top quintile using quarterly data shows that investors do gain from this requirement. The difference in alpha between using monthly ranking and quarterly ranking is about the same as that between quarterly and semi-annual ranking. Thus investors would gain about the same amount by moving to monthly data as they would by moving to quarterly.

For the top quintile, all bottom-up techniques have higher ex-post alphas than top-down ranking techniques, the traditional method of estimating betas and alphas for mutual funds. It is important to note that this is true even when the evaluation is done using the top-down method of evaluating performance. The fact that higher ex-post performance is achieved when ranking is done using the bottom-up technique (compared to the top-down technique) and that this result holds whether the ex-post performance is evaluated using either the top-down or bottom-up technique is powerful evidence that the bottom-up technique is the superior method of selecting mutual funds.

The bottom-up and top-down techniques select a lot of funds in common. To further explore the difference between the bottom-up and top-down techniques we formed a sample of funds selected in the top quintile by one technique but not by the other. For this group of funds the difference in alpha between the unique funds selected by the bottom-up model and the unique funds selected by the top-down model varied from a low of 17.3 basis points per month to a high of 27 basis points per month. These differences are economically large. However, given the small sample the differences are not statistically different from zero at the 5% level except for three cases. These three cases are ranking by the four-factor model and evaluating by the four-factor model using bottom-up or top-down evaluation and ranking by the three-factor bottom-up model and evaluating by the four-factor top-down model. In later sections with a larger sample, we

test all the differences in alpha between the bottom-up and top-down technique and find them highly statistically significant.<sup>24</sup>

The third result to note is that, comparing Tables 5 and 6, all alphas are higher when we rank by the three-factor model rather than the four-factor model. This is true no matter whether we evaluate using top-down or bottom-up three- or four-factor models. To test the statistical difference, we once again examined the set of funds that are selected by one technique and not the other. For this group, all of the bottom-up rankings are statistically significantly better than the four-factor rankings, at least at the 5% level no matter how the evaluation is done. The top-down three-factor models are statistically better than the top-down four-factor models when we use top-down evaluation. Examining the loadings on the factors shows that the loading on the momentum factor is highly intertemporally unstable, which probably accounts for its poorer performance.

Tables 5 and 6 provide strong evidence that bottom-up alphas are useful in identifying funds with either positive or negative alphas in the subsequent period, and both tables give insight into the frequency of holdings data that is helpful. Furthermore, both tables produce strong evidence that ranking on the three-factor model produces superior alpha compared to other models. Thus, in future sections, we will emphasize the three-factor model.

In Table 7 we examine the ability of the technique that works best in obtaining positive alpha in the top decile to differentiate future alphas not just in the tails but across all deciles. Table 7 shows ex-post alphas when funds are ranked by the bottom-up

---

<sup>24</sup> Examining the two empirical distributions shows that the bottom-up distribution first order stochastically dominates the top-down distribution. This means that, drawing from the empirical distribution of funds not held in common, the odds are always higher of drawing a higher alpha fund from the bottom-up distribution. The same cases with significant  $t$  values are cases where we can say that the bottom-up distribution dominates at a statically significant level.

monthly three-factor model, where the future alphas are computed from the bottom-up and top-down alphas using the three- or four-factor model. The rank correlation coefficients are shown in each column, and all are statistically significant at the 1% level. Thus the bottom-up monthly ranking not only produces tails with positive alphas for the highest decile and negative for the bottom decile, but also produces ex-post alphas for the deciles that are highly correlated with the ex-ante ranking. Once again the results hold even when evaluation is done using top-down alphas.

Before ending this section, we examine in one additional way the ability of bottom-up monthly ranking versus top-down ranking to produce better results. While performance in the evaluation period has, to this point, been examined using measures advocated in the literature of financial economics, many individual investors and some institutional investors still judge performance by the return on a fund versus return on its benchmark.<sup>25</sup> With this evaluation criterion, both bottom-up and top-down rankings for the top quintile of funds show positive returns above the benchmark. The average return over the benchmark for funds not selected by both techniques was 57 basis points higher per year when ranking is done by bottom-up betas rather than by top-down betas. Once again, bottom-up betas produce better ex-post performance no matter how that performance is judged.

#### **IV. Other Ranking Techniques**

In the prior section we showed that estimating bottom-up betas led to selection of superior-performing mutual funds. One possible reason is that bottom-up betas capture changes in beta over time. There is an alternative way of capturing changes in beta: conditioning beta on macro variables. In this section we will explore two alternative

---

<sup>25</sup> The benchmark index for each fund was selected as that public benchmark index selected by Morningstar.

methodologies for mutual fund selection: first, the best-known method of capturing changes in beta based on conditioning on macro variables as proposed by Ferson and Schadt (1996); second, the best-known method of using holdings data directly to evaluate mutual funds, the Grinblatt and Titman (1993) measure.

#### **A. Conditional Betas and Mutual Fund Selection**

Ferson and Schadt (1996) (hereafter F&S) have explored the impact of measuring mutual fund performance when conditioning betas on a set of predetermined time-varying variables representing public information. F&S find that conditioning beta on a small set of variables changes many of the conclusions about the selection ability of mutual fund managers.

The justification for using conditional betas in our top-down estimation is that conditional betas may be better estimates of beta at each point in time. When we estimate bottom-up betas using conditional betas, these betas are a smoothed estimate of beta. Insofar as there may be substantial error in bottom-up betas, and insofar as management uses publicly available data in determining their exposure, conditional betas may be better estimates of actual beta than unconditional betas.

#### **B. The Model of Conditioning Betas**

We follow F&S in defining four variables to capture public information that might affect management's choice of beta.<sup>26</sup> The variables are: the one-month Treasury bill yield, the dividend yield of the CRSP value-weighted index of NYSE/AMEX, the term spread, and the quality (credit) spread in the corporate bond market

We follow F&S in defining each of the variables, lagging them one month, and in assuming that time-varying betas in the four-factor model are a linear function of the four conditioning variables discussed above. Designating these conditioning variables as  $Z_1$

---

<sup>26</sup> F&S also use a January dummy but find that it has virtually no effect, so we don't include it here.

to  $Z_4$  the conditional beta with respect to any beta for fund  $P$  is found from the following time series regression:

$$(4) \quad \beta_{Pjt} = C_{P0j} + \sum_{k=1}^4 C_{Pkj} Z_{kt} + \varepsilon_{Pjt}$$

Where

$\beta_{Pjt}$  is the unconditional bottom-up beta for portfolio  $P$  with respect to factor  $j$  at time  $t$ .

$C_{Pkj}$  is the coefficient of the  $j$ th factor conditioned on variable  $k$  for portfolio  $P$ .

$Z_{kt}$  is the value of conditioning variable  $Z_k$  at time  $t$ .

$\varepsilon_{Pjt}$  is the random error.

When we estimate top-down alphas using conditional betas, we follow F&S in directly substituting equation (4) in the return-generating process (equation (1)) and estimating alpha using the time series of returns. When we estimate bottom-up beta using conditional betas, we regress over time each beta for each fund against the conditioning variables (conditioning regression). We then use the coefficients estimated from this equation to compute a conditional beta in each month (the intercept plus the estimated coefficients of the conditioning regression times the macro variables in month  $t$ ). The conditional beta is then used to compute alpha.<sup>27</sup>

In examining the ability of the F&S procedure to select superior funds, we focus on the three-factor model because this model, as shown above, produced the highest future performance of all ranking techniques. The results are shown in Table 8.

There are certain similarities between results from conditional and unconditional rankings. First, note that bottom-rankings outperform top-down rankings with and without conditioning betas. This is true whether the evaluation of subsequent

---

<sup>27</sup> Note that in using bottom-up conditional betas only monthly conditional betas are computed since monthly betas work best when unconditional betas are used.

performance is done with the top-down or bottom-up estimation and whether we use conditional or unconditional betas in the subsequent period. Second, note that bottom-up betas with and without conditioning variables produce alphas that are significantly different from zero (and from the lower decile) under all methods of evaluating subsequent performance. Finally, basing ranking on conditional variables tends to cause a deterioration in the ability of ranked funds to produce future alphas under all methods of measuring future performance.

It is worth noting that while bottom-up betas produce the highest ex-post alphas when unconditional estimation is used to rank funds, the alphas produced by this ranking are higher when subsequent performance is judged by a conditional model rather than the unconditional model.

In the next section we will examine whether our results hold up with a large sample, but first we will examine performance using an alternative way of using holdings data.

### C. Selecting Funds Using the Grinblatt and Titman Measure

An alternative way to rank funds is to employ a metric that uses changes in a portfolio's holdings to measure performance. The best known of these measures is the one developed by Grinblatt and Titman (1989). Grinblatt and Titman recommended ranking funds using the following measure for each fund:

$$(5) \quad m = \frac{\sum_{t=1}^T \sum_{j=1}^N (x_{jt} - x_{jt-k}) R_{jt+1}}{T}$$

$m$  is the performance measure

$x_{jt}$  is the weight of stock  $j$  in period  $t$

$k$  is the number of periods over which changes are calculated

$R_{jt}$  is the return on security  $j$  in month  $t$

$T$  is the number of time periods used to calculate  $m$

$N$  is the number of securities held

The weights add up to one in every period. Thus the sum of the positive changes in weights is the same as the sum of the negative changes, and the subsequent return can be viewed as the return on an arbitrage portfolio. The  $k$  controls how many periods are used to define the change in weights. The decision involves a choice of how many periods in the future the information that the trade was based on takes to be reflected in the markets. Grinblatt and Titman recommend three months and one year. Because of data limitations, we will only use three months. In addition, since we have monthly holdings data as well as quarterly holdings data, we will re-estimate the measure monthly as well as quarterly.

The Grinblatt and Titman measure has a number of potential problems when it is used as a ranking device. First, note that the metric measures performance gross of transaction costs and expenses, and thus does not capture the impact of expenses on performance. The second potential problem stems from the returns on securities that are held and neither bought nor sold. Note that the Grinblatt and Titman measure is concerned with the return on the securities bought less the securities sold. The return on the assets which are not traded during a period can have a real effect on the overall portfolio return. Both of these problems should introduce a lot of random error in the ranking.

The question remains whether, given these potential problems, the Grinblatt and Titman (G&T) measure is a useful ranking device for identifying mutual funds that will perform well in the future. We examine the ability of ranking based on the G&T measure to predict future values of the G&T measure and to predict the alpha from a multi-factor

model. To accomplish this latter step, we will use the same ex-post measures as used earlier in this paper, three-factor and four-factor alphas from both monthly holdings and from a time series of fund return. Since our data set, unlike Grinblatt and Titman's, reports holdings beyond traded equity, we need to discuss how return is computed for these assets. For cash we used the return on the one-month T-bill. For bonds we used the return on the relevant bond index. For options and futures we used the return on the underlying index or security where the weights are adjusted for the leverage inherent in the terms of the contract, and for non-traded assets we used the average return on the traded assets in the portfolio.

We computed the G&T measure for the same sample we used in the prior sections assuming that holdings were available every month and then assuming availability only over the quarter. The first change in holdings we can compute when we assume quarterly holdings is the change from December to March. In order to compute the G&T measure for 12 months we need return data for 12 months beginning in the first month after we examine the first change in holdings. This means that we compute the G&T measure through March of the second year. We then rank funds. To evaluate their performance, we use data from January through December of the second year. We compute the standard measures used in the prior section, bottom-up alpha and top-down alpha. In addition, since G&T showed that their measure predicted future values of their measure, we also compute their measure in the evaluation year. There is a three-month overlap in the ranking period and the evaluation period. Thus, some of the same mutual fund returns will be used in the ranking and evaluation period. This increases the probability that we will find that ranking on the G&T measure is related to future performance even if it is not.

In Table 9 we present the results of employing rankings based on the G&T measure to select funds that will perform well over a subsequent year. The strongest result arises from the use of the G&T measure computed monthly to select funds with positive G&T measure in the next period. The rank correlation across the 10 deciles using this evaluation criterion is 0.891, which is statistically significant at the 1% level. Computing the measure quarterly, the correlation is only 0.43, which is not statistically significant at any meaningful level. Note that whether computed monthly or quarterly, the G&T measure selects a top quintile of funds that has a positive average G&T measure in the evaluation period, while the bottom quintile has a negative value. If we judge the ability of the G&T measure to select funds that have positive alphas computed from either the top-down or the bottom-up three-factor model, the results are disappointing.<sup>28</sup> There is only a very small amount of correlation between the G&T measure and the standard measures of performance. Furthermore, the top 20% of funds selected by the G&T measure generally have large negative alphas in the year after they were selected.

In summary, an investor can use the G&T measure to help select funds that will do well on this measure in the future, but the investor will find very little information in this measure for selecting funds that will outperform index funds in the future.

## **V. Measuring Performance (Sample 2)**

The results to this point have been based on a sample of mutual funds where monthly holdings data were available. While the results are quite strong, they are based on a sample of 215 funds with restrictions on the amount of options and futures in each fund and restrictions on any investments not in the form of stock or cash. Two of the findings in the prior section are that, while forecasts of future performance based on betas

---

<sup>28</sup> The four-factor model gave similar results.

computed from monthly holdings data provide the best forecasts of future performance, using holdings data to compute betas on a quarterly basis 1) provides a large part of the benefit of using monthly data and 2) outperforms top-down measures of performance.

By restricting our analysis to funds which report holdings data at least quarterly, we can repeat the analysis in earlier sections on a much larger sample of funds over a longer period of time. This new sample consists of 1,255 funds with at least two years of quarterly data over the years 1995 to 2005.<sup>29</sup>

We employ the same methodology we employed in the previous section with an exception. Because our second sample of funds is over five times as large as our first sample, we can examine performance by deciles rather than quintiles. This allows us to concentrate more on the tails of the predicted distribution and in particular on the upper 10%. Those are the funds that the investor should potentially want to hold.

Tables 10, 11, 12 and 13 present our results for the larger sample. First note from Table 10 that for all ranking techniques and evaluation methods the average alpha is positive for the top decile and negative for the bottom decile, and that the difference from zero is significant at the 1% level in all cases but one, where it is significant at the 5% level.

Second, as shown in Table 10, our analysis presents clear evidence that under each and every model used to evaluate performance, ranking by the bottom-up method produces higher ex-post alphas than ranking by the top-down method. To test this we took the 10% of funds ranked highest by the three- and four-factor bottom-up ranking models and the three- and four-factor top-down models and compared the ex-post

---

<sup>29</sup> While all of these funds are categorized as common stock funds, we no longer place restrictions on the percentage of the portfolio invested in stock and cash or the percentage in options and futures. As explained earlier, we do eliminate specialty, international, real estate, and index funds from our sample.

performance of these two groups under different evaluation schemes. To judge statistical significance, we had to recognize the fact that the top decile from the bottom-up ranking technique and the top decile from the top-down ranking technique have alphas which are correlated because they contain many of the same funds. To correct for this we discarded all funds that were in common and tested the significance of the difference on the remaining funds. Statistical significance was judged by the  $t$  test for the difference in means with unequal variances. As can be seen from Table 11 in all eight cases, no matter which method is used to evaluate funds, higher alphas are achieved when ranking is performed on the bottom-up method. Results are statistically significant at the 5% level or better in all but one case.<sup>30</sup> In that one case, ranking by the four-factor model and evaluation by the three-factor model, the direction is consistent with the other models but the difference is not significantly different from zero.<sup>31</sup>

---

<sup>30</sup> The numbers reported assume that the holdings data are available at the time they occur. This type of data should be available to an institution. Purchasers of the Morningstar database would get this data with a delay. The average delay was two months. We took the actual delay from the Morningstar files in getting the data (one to three months) for each fund and assumed that (1) funds were invested in a zero-alpha investment (T-bills or an index fund) for the appropriate lag, and (2) membership in the top decile was determined using only data available at the time the decile was formed. The alpha for the quarterly data using the three-factor model was 0.197% per month, while for the four-factor model it was 0.209 per month when alphas are calculated using bottom-up betas. This is still substantially and statistically higher than the alpha selected using the top-down betas.

<sup>31</sup> This is the least interesting of the cases, for if one were to believe that the four-factor model is more appropriate than the three-factor model, then we should evaluate the results using the four-factor model. Another way to examine the difference between bottom-up and top-down ranking is first order stochastic dominance. Under all ranking and evaluations, the bottom-up technique first order stochastically dominates the top-down technique. Using the Wilcoxon (Mann-Whitney) test to test if they are statistically different produces results with significance levels identical to Table 11. Thus at a statistically significant level an

The numbers in Table 10 are average numbers. But an investor is likely to select just a few funds. The odds of selecting a single fund with a positive alpha in the next year when ranking is done by the three-factor bottom-up alpha and a fund is selected from the top decile is 69.6% when evaluation is done using the bottom-up three-factor alpha and 66.3% when evaluation is done using the top-down three-factor model.<sup>32</sup>

Are the results likely to arise because one of the models is selecting a particular type of fund rather than managers who do well? While it is not possible to examine this for all possible partitions of the mutual funds selected, we did look at the major classifications of the set of funds picked by the bottom-up or top-down model for both the three-factor and four-factor model. As shown in Table 13 the composition of the funds selected by the top-down and bottom-up methods are very close. There is no evidence that the bottom-up or top-down method of estimating beta selects very different types of funds.

We next examine the future alpha when ranking is done by the three-factor model rather than the four-factor model. As shown in Table 10 in all cases, the three-factor model produces a higher ex-post alpha. To judge the statistical significance of the difference we use the methodology described above. Funds not in common in the top 10% of ranked alpha were identified and a *t* test for difference in means with unequal variances was employed.

When we look at funds not in common, the alphas produced by selection (ranking) on the three-factor model are always higher than those produced by the four-factor model no matter which technique is used to evaluate the alphas. However, the results are only statistically significant at the 5% level in four out of the eight cases

---

investor has higher odds of getting a high alpha fund using bottom-up ranking rather than top-down ranking.

<sup>32</sup> If evaluation is done by the four-factor model, the number decreases slightly to 63.1% for both cases.

examined, although two more are significant at the 10% level and almost significant at the 5% level.

We have seen that the quarterly bottom-up ranking leads to identifying funds that outperform index funds in the future. While it successfully identifies the tails, it is interesting to see if it leads to a correct ordering of funds across the deciles. Table 14 presents the ex-post alpha associated with each decile.

Note that when quarterly bottom-up rankings are used, the order of the deciles by ex-post performance is very consistent. The Spearman rank correlation coefficient ranges from 0.988 evaluated by the top-down three-factor model to 0.873 when evaluated by the four-factor model. The results from examining the larger less restrictive sample reinforce the results found with the smaller sample.

## **VI. Possible Biases**

The final question we address is whether either requiring monthly or quarterly holdings data, or requiring two years of data, introduces a bias.

There are two possible sources of bias. First, funds that voluntarily provide monthly or quarterly holdings data may be different from those that do not. Second, even if funds that provide monthly or quarterly holdings are no different from those that do not, requiring two years of monthly or quarterly holdings may bias the results in two ways. When we require two years of monthly or quarterly holdings data we are excluding funds that merged or liquidated. Also, when we require two years of monthly or quarterly holdings data we are excluding funds that reported monthly (or quarterly) holdings data in the ranking period but did not report every month (or quarter) in the subsequent year. Each of these potential sources of bias will now be examined.

The first question is whether the characteristics of funds that voluntarily report holdings monthly are different from the general population. Ge and Zheng (2006) examine whether funds that report voluntarily on a quarterly basis are different from those that report semi-annually. They found that those that reported voluntarily had 0.04% lower expenses, 10% less turnover, were less likely to commit fraud, and differed somewhat in performance. For our study, it is the possibility of difference in performance that needs to be examined. To examine this, we performed the following two analyses. For each fund in Sample 1 we randomly selected 10 funds with the same investment objective that did not report monthly holdings data. We then computed alpha for each fund in the random sample in a manner identical to the method we used to compute the top-down alphas for our sample. Since Sample 2 is so much larger, we matched each fund in our sample with a fund of the same objective. The difference in average alpha between our first sample and the matching sample was three basis points, while for the second sample it was four basis points, neither of which is statistically significant at any meaningful level. Results are not due to an upward bias in mean alpha for our sample relative to the population.

The second possible source of bias is survivorship bias. To analyze this, we examined all funds that met our criteria for inclusion in the ranking period but merged or liquidated in the evaluation period. There were 24 such funds out of 128 for Sample 1 and 8 out of 63 in Sample 2. Typical of funds that merge, the performance in the period before the merger was poor. Thus almost all of these funds would have ranked in the bottom group. Of the 24 funds in Sample 1 that had monthly data in the first year and merged in the second, only one ranked in the top quintile in the first year (it was in the second decile). Since the top quintile is the quintile of interest to investors, we need to examine the effect of inclusion of this fund on the alpha of this quintile. We cannot

compute bottom-up alpha for this fund in the evaluation period. However, we did compute its alpha in the evaluation period using the betas from the three-year time series regression computed through the last year of the ranking period. Its alpha in the evaluation period was positive and slightly above the alphas for the rest of the funds in the first quintile using the same calculation technique.<sup>33</sup> Including this fund in the first quintile would increase slightly the alpha earned by an investor who selected this quintile.<sup>34</sup> In the second sample none of the funds that merged in the second period would have been in the first decile and thus their exclusion does not affect our results.

Another bias could arise in Sample 1 if funds that had one complete year of monthly holdings data but not a second year stopped reporting monthly data because either their performance changed or they realized that they were performing not as well as the funds that continued reporting monthly data. The funds that continued to exist but did not have 12 months of reported holdings fell into two classes: those that switched to quarterly reporting and those that were missing a few months of data. Of the 104 funds that had 12 months of data in one year but less than 12 months of data in the subsequent year, only four switched to quarterly reporting.<sup>35</sup> Turning first to the 100 funds that did not switch to quarterly reporting, we find that about 75% of these funds have holdings reports for 10 or 11 months out of the possible 12. For the remaining funds, data appear to be missing at erratic intervals with no discernible pattern with respect to calendar

---

<sup>33</sup> With the inclusion of more funds, the definition of quintiles changes slightly. Examining funds that would be added to the top quintile shows this has little or no effect on our results.

<sup>34</sup> In calculating alphas for merged funds, we assumed that an investor liquidated the position when the fund merged and invested equally in all funds in the first quintile after the merger; the alpha is a combination of the alpha on the fund plus the after-merger alpha.

<sup>35</sup> In the case where observations were missing in one or more months toward the end of a year, data were examined in the following year to ascertain whether the funds were switching to quarterly reporting.

months. The random nature of reporting months suggests that missing months is a problem of data collection rather than a strategic decision by funds. As a final check, we examined directly whether funds that did not report every month in the second year have different alphas than those that did. We computed top-down betas for the 104 funds in the ranking and evaluation years. There was a general, but not significant, tendency for these funds to have slightly higher alphas in the evaluation period than the funds that reported monthly holdings data in that period (an average of one basis point).

When we looked at the four funds which switched to quarterly reporting, we again found that they performed no worse than the funds that continued to report holdings on a monthly basis. In Sample 2 there were no instances of funds that went to semi-annual reporting over the period in which only semi-annual reporting was required. Based on the data, there is no reason to believe that omitting funds that were missing some months of holdings data biased the results reported in this article.

## **VII. Conclusions**

In this paper we have explored the use of several alternative techniques for measuring performance to determine whether they lead to the identification of mutual funds that will outperform the average actively managed mutual fund and passive index fund in subsequent periods.

The ranking measures we investigate are based on the standard Fama and French three-factor model and a four-factor version including a momentum measure. A unique part of this study was the use of data on the monthly holdings of securities in each fund in our sample to estimate betas and alphas at a moment in time.

We find that the use of holdings data to compute betas and alphas leads to superior selection of mutual funds compared to selecting on the basis of the alphas from a time-series regression on fund returns. This result held whether we evaluated subsequent

performance using alphas computed from holdings data on a monthly basis or alphas computed using a time-series regression on mutual fund returns. Interestingly, ranking on the three-factor model led to better results than ranking on the four-factor model whether performance in the subsequent period was measured using either the three- or four-factor model.

We examined two other ranking devices, one using conditional betas and another using the Grinblatt and Titman measure. Neither was as successful as bottom-up alpha in selecting funds with high alpha in a subsequent period. However, the Grinblatt and Titman measure was able to select funds that ranked high on the Grinblatt and Titman measure in subsequent periods.

When we compared results assuming holdings data were available at different intervals, we found that, in general, the less frequently they were available, the poorer the predictive power. Ranking using the 3-factor model rather than the four-factor model produced higher alpha in almost all cases, whether realized alphas were measured by the three- or four-factor model. Thus, in examining the effect of frequency on realized alpha, more emphasis should be placed on the results from ranking using the three-factor model. Here the results are consistent across all methods and intervals. The shorter the interval between holding reports, the more accurately superior performance can be identified. These results provide strong evidence that the SEC improved the ability of investors to evaluate funds by its recent decision to require quarterly holdings data. Our analysis suggests that a further gain of the same size could be achieved by requiring that monthly holdings be reported.

## Bibliography

Baker, M.; Wurgler, J.; Litov, L.; and Wachter, J.. “Can Mutual Fund Manager Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements.” *Journal of Financial and Quantitative Analysis*, forthcoming.

Blake, C. R.; Elton, E. J.; and Gruber, M. J. “The Performance of Bond Mutual Funds.” *Journal of Business*; 66 (1993), 371-403.

Brown, S.; Gallagher, D.; Steenbeek, O.; and Swan, P. “Double or Nothing: Patterns in Equity Fund Holdings and Transactions.” Working Paper, NYU Stern School of Business (2005).

Busse, J. and Irvine, P. Bayesian alphas and mutual fund persistence. *Journal of Finance*, LXI(6) (2006), 2251-2288.

Carhart, M.. Persistence in mutual fund performance. *Journal of Finance*, 52 (1997), 57-28.

Dybvig, P.; Ross, S. Differential information and performance measurement using the security market line. *Journal of Finance* 40 (1985), 383-399.

Elton, E. J.; Gruber, M. J.; and Blake, C. R. The persistence of risk-adjusted mutual fund performance. *Journal of Business* 69 (1996), 133-157.

Elton, E. J.; Gruber, M. J.; Brown, S.; and Goetzman, W. Modern portfolio theory and investment analysis. Seventh edition, John Wiley & Sons (2007).

Elton, E. J.; Gruber, M. J.; Krasny, Y.; and Ozelge, S. The effect of the frequency of holding data on conclusions about mutual fund management behavior. Unpublished manuscript, New York University (2006).

Ferson, W. E.; Schadt, R.W. “Measuring Fund Strategy and Performance in Changing Economic Conditions.” *The Journal of Finance* 51 (1996), 425-461.

Ge, W.; Zheng, L. “The Frequency of Mutual Fund Portfolio Disclosure.” Working Paper, Ross School of Business, the University of Michigan (2005).

Grinblatt, M.; Titman, S. “Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns.” *Journal of Business* 66 (1993), 47-68.

Gruber, M. J. Another Puzzle: The Growth in Actively Managed Mutual Funds. *Journal of Finance* 51 (1996), 783-810.

Investment Company Institute. *Survey of Fund Groups' Portfolio Disclosure Policies: Summary of Results*. Washington, D.C.: Investment Company Institute (2001).

Mamaysky, H.; Spiegel, M.; and Zhang, H. “The Dynamic Behavior of Mutual Fund Styles.” *Review of Financial Studies*, 21(2008), 233-264.

Wermers, R. “The Potential Effects of More Frequent Portfolio Disclosure on Mutual Fund Performance.” *Investment Company Institute Perspective* (2001), pp. 1-12.

Wermers, R. “Performance Evaluation with Portfolio Holdings Information.” *North American Journal of Economics and Finance*, 17 (2006), 207-230.

Zheng, L. “Is money smart? – a study of mutual fund investors’ fund selection ability.” *Journal of Finance* 54 (1999), 901-933.

**Table 1**

**Monthly Betas Estimated from Fund Holdings**

	Market Minus Risk-Free Rate	Small Minus Big	High Minus Low
Average Beta	0.9843	0.1628	0.0853
Top 25%	1.1057	0.4115	0.2446
Bottom 25%	0.8329	-0.1388	-0.0295
Absolute Value of Monthly Change in Beta	0.0437	0.0397	0.0398
Standard Deviation of Monthly Beta	0.1064	0.1179	0.1076

For each fund we first compute an average monthly beta using the data over all periods where we have monthly data. Likewise, for each fund we compute the average absolute change in beta between months and the standard deviation of the monthly betas. The numbers in the table are the average of these calculations or refer to the distribution across funds.

**Table 2****Mean Differences in Beta**

	Market Minus Risk-Free Rate		Small Minus Big		High Minus Low	
	Bottom-Up	Top-Down	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Monthly		-0.0169		0.0133		-0.1395
Quarterly	-0.0045	-0.0167	0.0004	0.0147	-0.0024	-0.1413
Semi-Annual	0.0009	-0.0179	0.0037	0.0226	0.0005	-0.1480
Annual	0.0009	-0.0206	0.0084	0.0361	-0.0079	-0.1578

Difference in beta is defined as the estimate produced by the techniques noted in the table minus the bottom-up monthly beta. Differences between any two techniques can be determined by differencing the entries in the table.

**Table 3****Average Absolute Difference in Beta**

	Market Minus Risk-Free Rate		Small Minus Big		High Minus Low	
	Bottom-Up	Top-Down	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Monthly		0.1304		0.1266		0.1958
Quarterly	0.0318	0.1297	0.0287	0.1257	0.0298	0.1955
Semi-Annual	0.0482	0.1284	0.0475	0.1260	0.0498	0.2004
Annual	0.0748	0.1276	0.0734	0.1256	0.0750	0.2065

This table shows the average absolute value of the difference for all mutual funds between the beta using the technique indicated and the beta estimated each month from security holdings (bottom-up).

**Table 4****Rank Correlations Between Techniques**

	<b>Bottom-Up</b>				<b>Top-Down</b>
	Monthly	Quarterly	Semi-Annual	Annual	Annual
<b>Bottom-Up</b>					
Monthly	1	0.994	0.986	0.968	0.762
Quarterly		1	0.990	0.971	0.757
Semi-Annual			1	0.987	0.775
Annual				1	0.784
<b>Top-Down</b>					
Annual					1

This Table shows the Spearman rank correlation in alpha between each pair of techniques used to estimate alpha.

**Table 5****Ex-Post Alpha Using the Three-Factor Model as a Ranking Device**

<b>Panel A: Evaluation Using the Three-Factor Model</b>				
	Top Quintile		Bottom Quintile	
Ranking Technique	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Bottom-up monthly	0.158**	0.088	-0.413**	-0.538**
Bottom-up quarterly	0.125*	0.056	-0.435**	-0.567**
Bottom-up semi-annual	0.104*	0.031	-0.385**	-0.512**
Bottom-up annual	0.100*	0.028	-0.412**	-0.574**
Top-down	0.089	0.015	-0.411**	-0.542**
<b>Panel B: Evaluation Using the Four-Factor Model</b>				
	Top Quintile		Bottom Quintile	
Ranking Technique	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Bottom-up monthly	0.132*	0.075	-0.333**	-0.436**
Bottom-up quarterly	0.129*	0.050	-0.352**	-0.465**
Bottom-up semi-annual	0.122*	0.033	-0.300**	-0.405**
Bottom-up annual	0.124*	0.033	-0.322**	-0.469**
Top-down	0.063	0.012	-0.335**	-0.458**

This table shows the alpha earned in the period subsequent to the period used for ranking.

Alphas in the columns labeled “bottom-up” are computed using monthly holdings and security betas computed over the year using weekly data. Alphas in the columns labeled “top-down” are computed using a time series regression of weekly fund returns. For both types of alphas, it is assumed that an equal amount is invested in each fund in the quintile indicated. Double asterisk indicates statistical significance different from zero at 1% level and single asterisk at 5% level.

**Table 6**

**Ex-Post Alpha Using Four-Factor Model as a Ranking Device**

<b>Panel A: Evaluation Using the Three-Factor Model</b>				
	Top Quintile		Bottom Quintile	
Ranking Technique	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Bottom-up monthly	0.051	-0.060	-0.438**	-0.536**
Bottom-up quarterly	0.054	-0.063	-0.410**	-0.522**
Bottom-up semi-annual	0.034	-0.089	-0.444**	-0.565**
Bottom-up annual	0.027	-0.094	-0.429**	-0.557**
Top-down	-0.035	-0.174**	-0.327**	-0.448**
<b>Panel B: Evaluation Using the Four-Factor Model</b>				
	Top Quintile		Bottom Quintile	
Ranking Technique	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Bottom-up monthly	0.091	-0.041	-0.383**	-0.448**
Bottom-up quarterly	0.085	-0.049	-0.342**	-0.419**
Bottom-up semi-annual	0.072	-0.068	-0.385**	-0.469**
Bottom-up annual	0.054	-0.080	-0.354**	-0.461**
Top-down	-0.017	-0.142*	-0.274**	-0.373**

This table shows the alpha earned in the period subsequent to the period used for ranking. Alphas in the columns labeled “bottom-up” are computed using monthly holdings and security betas computed over the year using weekly data. Alphas in the columns labeled “top-down” are computed using a time series regression of weekly fund returns. For both types of alphas, it is assumed that an equal amount is invested in each fund in the quintile indicated. Double asterisk indicates statistical significance different from zero at 1% level, single asterisk at 5% level.

**Table 7****Decile Alphas from Monthly Bottom-Up Alphas: 3-Factor Model**

Decile	Evaluated by Three-Factor Monthly Estimate		Evaluated by Four-Factor Monthly Estimate	
	Bottom- Up	Top- Down	Bottom-Up	Top-Down
1	-0.472	-0.523	-0.332	-0.374
2	-0.355	-0.551	-0.334	-0.496
3	-0.214	-0.470	-0.103	-0.335
4	-0.120	-0.301	-0.082	-0.260
5	-0.216	-0.373	-0.260	-0.386
6	0.072	-0.071	0.033	-0.071
7	-0.020	-0.170	-0.034	-0.145
8	0.010	-0.083	0.045	-0.080
9	0.164	0.076	0.126	0.058
10	0.151	0.100	0.138	0.092
Spearman Rank Correlation	0.915**	0.939**	0.939**	0.867**

\*\* Significant at the 1% level.

This table shows the average ex-post alpha in each decile where the ex-ante ranking is done on the basis of the monthly bottom-up 3-factor alpha and evaluation is based on bottom-up monthly and top-down alphas using the three- and four-factor models.

**Table 8****Ex-Post Alpha Using the Three-Factor Model as a Ranking Device**

<b>Evaluation Using the Three-Factor Model</b>				
	Top Quintile			
Ranking Technique	Bottom-Up Monthly	Conditional Bottom-Up Monthly	Top-Down	Conditional Top-Down
Bottom-up monthly	0.158**	0.277**	0.088	0.131**
Conditional bottom-up monthly	0.139**	0.249**	0.073	0.107*
Top-down	0.089	0.184**	0.015	0.055
Conditional top-down	-0.070	-0.038	-0.080	-0.008

This table shows the alpha earned in the period subsequent to the period used for ranking.

Conditional alphas in the columns labeled “bottom-up” are computed using 13-month

regressions of bottom-up betas on 4 conditional variables over the evaluation year plus the

prior December. Conditional alphas in the columns labeled “top-down” are computed using a

time series regression of 36 monthly excess fund returns on the factors plus the cross-products

of the factors and the 4 conditional variables over the 3 years starting in January of the

evaluation year. The alpha is the intercept plus the average of the 12 monthly residuals in the

evaluation year. For both types of alphas, it is assumed that an equal amount is invested in each

fund in the quintile indicated. Double asterisk indicates statistical significance different from

zero at 1% level and single asterisk at 5% level.

**Table 9****Ranking using Monthly and Quarterly G&T Measures**

<b>Panel A: Evaluation Using G&amp;T Monthly Measure and Three-Factor Alphas</b>						
	Top Quintile			Bottom Quintile		
Ranking Technique	G&T	Bottom-Up	Top-Down	G&T	Bottom-Up	Top-Down
G&T monthly	0.072*	-0.201**	-0.332**	-0.086**	-0.182*	-0.284**
G&T quarterly	0.027	-0.054	-0.236**	-0.007	-0.226**	-0.315**
<b>Panel B: Decile Rank Correlations</b>						
Ranking Technique	G&T		Bottom-Up		Top-Down	
G&T monthly	0.891**		0.055		0.152	
G&T quarterly	0.430		0.261		-0.200	

This table shows the performance in the next year when funds are ranked by the Grinblatt and Titman (G&T) measure. Performance is measured by the G&T measure and alpha computed using monthly holdings and betas computed using weekly data or alpha computed using weekly returns. An equal amount is assumed invested in each fund in the quintile. Double asterisk indicates significance at 1% level and single asterisk at 5% level.

**Table 10****Panel A: Ex-Post Alpha Using the Three-Factor Model as a Ranking Device**

	Evaluation Using the Three-Factor Model			
	Top Decile		Bottom Decile	
	Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
Bottom-up Ranking	0.246**	0.418**	-0.261**	-0.291**
Top-down Ranking	0.151**	0.301**	-0.261**	-0.291**
	Evaluation Using the Four-Factor Model			
	Top Decile		Bottom Decile	
	Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
Bottom-up Ranking	0.283**	0.350**	-0.167**	-0.276**
Top-Down Ranking	0.200**	0.252**	-0.177**	-0.282**

**Panel B: Ex-Post Alpha Using Four-Factor Model as a Ranking Device**

	Evaluation Using the Three-Factor Model			
	Top Decile		Bottom Decile	
	Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
Bottom-up Ranking	0.197**	0.382**	-0.332**	-0.324**
Top-down Ranking	0.111*	0.235**	-0.300**	-0.309**
	Evaluation Using the Four-Factor Model			
	Top Decile		Bottom Decile	
	Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
Bottom-up Ranking	0.259**	0.333**	-0.235**	-0.308**
Top-down Ranking	0.156**	0.195**	-0.218**	-0.317**

This table shows the alpha earned in the period subsequent to the period used for ranking.

Alphas in the columns labeled “bottom-up evaluation” are computed using monthly holdings and security betas computed over the year following the ranking using weekly data. Alphas in the columns labeled “top-down evaluation” are computed using a time series regression of weekly fund returns in the year following the ranking. For both types of alphas, it is assumed that an equal amount is invested in each fund in the decile indicated. Double asterisk indicates statistical significance different from zero at 1% level, and single asterisk at 5% level.

**Table 11**

**Ex-Post Difference in Performance for Funds Selected By Bottom-Up Ranking Minus Those Selected By Top-Down Ranking (positive numbers show bottom-up ranking produces higher alpha)**

**Panel A: Ranking by the Three-Factor Model**

Evaluation by the Three-Factor Model		Evaluation by the Four-Factor Model	
Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
0.264*	0.311**	0.223*	0.269**

**Panel B: A: Ranking by the Four-Factor Model**

Evaluation by the Three-Factor Model		Evaluation by the Four-Factor Model	
Bottom-Up Evaluation	Top-Down Evaluation	Bottom-Up Evaluation	Top-Down Evaluation
0.217	0.226**	0.186*	0.210**

\*\* Significant at the 1% level

\* Significant at the 5% level

This table shows the difference in alpha for the top decile (and its statistical significance) for the evaluation year (the year subsequent to the time of ranking). The difference is computed for the funds selected by the bottom-up method minus those selected by the top-down method for all funds which are not selected by both methods.

**Table 12**

**Types of Funds Selected by Each Ranking Device**

	Three-Factor		Four-Factor	
	Bottom-Up	Top-Down	Bottom-Up	Top-Down
Aggressive Growth	47.83%	45.83%	38.18%	35.78%
Long-Term Growth	40.22%	44.79%	46.36%	55.96%
Growth & Income	11.96%	9.38%	15.45%	8.26%

The percentages in the table represent the composition by objective of the funds in the top decile for the model indicated at the head of each column.

**Table 13****Decile Alphas from Quarterly Bottom-Up Alphas: 3-Factor Model**

Decile	Evaluated by Three-Factor Quarterly Estimate		Evaluated by Four-Factor Quarterly Estimate	
	Bottom- Up	Top- Down	Bottom-Up	Top-Down
1	-0.261	-0.291	-0.167	-0.276
2	-0.348	-0.256	-0.244	-0.234
3	-0.311	-0.186	-0.233	-0.175
4	-0.194	-0.183	-0.126	-0.157
5	-0.243	-0.163	-0.173	-0.157
6	-0.170	-0.041	-0.121	-0.048
7	-0.043	0.038	0.020	0.029
8	-0.124	0.007	-0.091	-0.020
9	-0.050	0.159	0.003	0.112
10	0.246	0.418	0.283	0.350
Spearman Rank Correlation	0.915**	0.988**	0.873**	0.985**

\*\* Significant at the 1% level.

This table shows the average ex-post alpha in each decile where the ex-ante ranking is done on the basis of the quarterly bottom-up 3-factor alpha and evaluation is based on bottom-up quarterly and top-down alphas using the three- and four-factor models. This table presents realized alpha for the years 1996 – 2005.