NBER WORKING PAPER SERIES

CONDITIONAL PERFORMANCE MEASUREMENT USING PORTFOLIO WEIGHTS:
EVIDENCE FOR PENSION FUNDS

Wayne Ferson
Kenneth Khang

Conditional Performance Measurement Using Portfolio Weights:
Evidence for Pension Funds
Wayne Ferson and Kenneth Khang
NBER Working Paper No. 8790
February 2002
JEL No. G12, G14, G23

ABSTRACT

This paper combines the use of portfolio holdings data and conditioning information to create a new performance measure. Our conditional weight-based measure has several advantages. Using conditioning information avoids biases in weight-based measures as discussed by Grinblatt and Titman (1993). When conditioning information is used, returns-based measures face a bias if managers can trade between observation dates. The new measures avoid this interim trading bias. We use the new measures to provide fresh insights about performance in a sample of U.S. equity pension fund managers.

Wayne Ferson
Carroll School of Management
Boston College
140 Commonwealth Avenue
Fulton Hall 330B
Chestnut Hill, MA 02467
and NBER
wayne.ferson@bc.edu

Kenneth Khang
School of Business Administration
University of Wisconsin
Milwaukee, WI 53201

1. Introduction

Measuring the performance of a managed portfolio has long been of interest to financial economists and practitioners alike. However, even after years of research, several issues remain unresolved. Among these is how to handle the dynamic behavior of a managed portfolio. Difficulties arise, not only because the required returns on the assets in a portfolio may be time-varying, but also because the portfolio structure may vary due to the manager's strategy or other influences, such as exogenous cash flows to the fund. These dynamics create problems in measuring performance.

Most current performance measurement techniques are returns-based and involve regressing the return of a portfolio on some benchmark return. The measure of performance, or alpha, is the intercept in the regression. A strength of returns-based methodologies is their minimal information requirements. One needs only returns on the managed portfolio and the benchmark. However, this ignores potentially useful information that is often available: the composition of the managed portfolio. Cornell (1979) was among the first to propose using portfolio weights to measure the performance of trading strategies. Copeland and Mayers (1982) modify Cornell's measure and use it to analyze Value Line rankings. Grinblatt and Titman (1993) propose a weight-based measure of mutual fund performance.[1]

Previous studies that use portfolio weights combine them with unconditional moments to measure performance. However, Ferson and Schadt (1996) put returns-based

---

[1] A number of studies have used the Grinblatt and Titman measure, including Grinblatt and Titman (1989a); Grinblatt, Titman, and Wermers (1995); Zheng (1999); and Wermers (1997).

measures into a conditional framework and find that it changes the results. This paper

considers conditioning information in weight-based measures of performance.[2]

The use of portfolio weights may be especially important in a conditional setting.

When expected returns are time-varying and managers trade between return observation

dates, returns-based approaches are likely to be biased. Even conditional returns-based

methods are affected. This bias, which we call the interim trading bias, can be avoided by

using portfolio weights in a conditional setting.

This paper develops a weight-based approach with conditioning information and

evaluates the relative advantages of this approach.[3] We illustrate the approach on a

sample of U.S. equity pension fund managers, 1985 to 1994. The growth-style pension

funds tend to trade more and follow a momentum strategy, compared with value-style

funds. These differences in fund strategy allow us to highlight several features of the

performance measures.

We find that a conditional weight-based measure has some important advantages.

It controls for interim trading bias in a setting where returns-based measures, including

conditional measures, are severely biased. The conditional weight measure controls for

trading on public information better than either returns-based measures or weight-based

---

[2] Daniel, Grinblatt, Titman, and Wermers (1997) create characteristic-based benchmarks to measure the performance of mutual funds. They condition the benchmark on the fund's characteristics.

[3] Recent studies have made limited use of portfolio weights in a conditional setting, by regressing returns on lagged weights and conditioning variables to see if the weights have marginal explanatory power e.g., Graham and Harvey, 1996; Becker, Ferson, Myers, and Schill, (1999). These studies are limited to market timing, and thus a single weight in stocks versus cash. We study the weights on the full vector of assets in a portfolio. Eckbo and Smith (1998) use a version of our measure, as developed in Khang (1997), in a study of insider trading.

measures that use unconditional means. Using both portfolio weights and conditioning variables, we also obtain more precision than with the unconditional weight measure.

In our sample of pension funds, returns-based measures suggest that the funds have positive abnormal returns, which is consistent with previous studies.[4] With conditioning information in a weight-based measure, the funds have neutral performance. Thus, previous estimates of abnormal pension fund performance may reflect biased measures.

The rest of the paper proceeds as follows: Section 2 motivates conditional weight-based measures of performance. Section 3 formulates the empirical measure and discusses estimation issues. The data are discussed in Section 4. Section 5 presents the empirical results, Section 6 explores their robustness, and Section 7 concludes the paper.

2. The measures

The intuition behind weight-based performance measures is simple. Suppose a manager has private information telling him when returns are likely to be higher or lower than expected by the market. Other things being equal, the manager can profit by shifting his portfolio weights toward those assets with returns that are likely to be higher than

---

[4] Lakonishok, Shleifer, and Vishny (1992) found inferior performance among pension fund managers using the Standard & Poor's (S&P) 500 as the performance benchmark. However, Christopherson, Ferson and Glassman (1998b) show this can be explained by the small stock exposure of the funds in their sample over a period in which small stocks perform poorly relative to the S&P 500. Coggins, Fabozzi, and Rahman (1993) find positive performance among 72 equity portfolios with an unconditional returns-based approach. Christopherson, Ferson, and Glassman (1998a) find evidence of persistent performance using a conditional returns-based approach.

expected and away from those assets with returns that are likely to be lower. This suggests that the covariance between the change in a portfolio's weights and subsequent abnormal security returns may be used to measure performance. Here, we define a security's abnormal return as the component of return not expected by the market.

Grinblatt and Titman (1993) point out that the covariance between the weights and subsequent abnormal returns need not be positive for every security. Consider two securities that are correlated with each other. This manager may choose to buy one and sell the other as a result of hedging considerations. Grinblatt and Titman (1989b) demonstrate, however, that the sum of the covariances across securities will be positive for an investor with private information, given nonincreasing absolute risk aversion as defined by Rubinstein (1973).

The conditional weight measure (CWM) is the conditional covariance between future returns and portfolio weight changes, summed across the asset holdings. The measure can be motivated, following Grinblatt and Titman (1993), with a single-period model in which an investor maximizes the expected utility of terminal wealth.

2.1. The conditional weight measure

Consider the following utility maximization problem,

$$\max_{\mathbf{w}} \quad E\left\{U(W_0(1+r_f)+W_0\mathbf{w}'\tilde{\mathbf{R}}) \mid Z, S\right\}, \tag{1}$$

where r_f is the riskfree rate; $\tilde{\mathbf{R}}$ is the vector of risky asset returns in excess of the riskfree rate; W_0 is the initial wealth; $\mathbf{w}$ is the vector of portfolio weights on the risky assets; Z is

public information available at time 0; and S is private information available at time 0.

Private information, by definition, is correlated with $\tilde{\mathbf{R}}$, conditional on Z. If we assume

that returns are conditionally normal and consider an investor with nonincreasing absolute

risk aversion, it follows from the optimization problem (see Khang, 1997) that

$$E\left\{\mathbf{w}(Z,S)'\left(\tilde{\mathbf{R}} - E\left(\tilde{\mathbf{R}} \mid Z\right)\right) \mid Z\right\} > 0, \tag{2}$$

where $\mathbf{w}(Z,S)$ is the optimal weight vector and $\tilde{\mathbf{R}} - E\left(\tilde{\mathbf{R}} \mid Z\right)$ are the unexpected, or

abnormal, returns from the perspective of an observer with the public information.

Because one of the variables has mean zero, Eq. (2) is also a conditional covariance. It

says that, conditional on the public information, the sum of the conditional covariances

between the weights of a manager with private information, S, and the abnormal returns

for the securities in a portfolio is positive.[5] If the manager has no private information, S,

then the covariance is zero.

To develop the empirical conditional weight measure, we introduce a benchmark

weight, $\mathbf{w}_b$, that is in the public information set Z, so Eq. (2) implies

$$E\left\{\left(\mathbf{w}(Z,S) - \mathbf{w}_b\right)'\left(\tilde{\mathbf{R}} - E\left(\mathbf{R} \mid Z\right)\right) \mid Z\right\} > 0, \tag{3}$$

if the manager has superior information, S. Because $\mathbf{w}_b$ is a constant given Z, it will not

affect the conditional covariance. By making $\mathbf{w}_b$ a function of the weights of the

portfolio in a past time period, we use weight changes in the empirical CWM. Weight

---

[5] Because this result uses conditional normality of the returns, given (Z,S), the risk of the asset returns depends on the precision of the signals but not on the realization of the signals. When there is private information, S, the product of the weight and the unexpected return will be nonnormal, conditional on Z, and the conditional covariance will depend on Z in general.

changes are advantageous on statistical grounds, as the levels of the weights may be

nonstationary.[6]

2.2. Interpreting the conditional measures

Conditional performance evaluation is usually motivated by the assumption of

market efficiency with respect to the public information Z (e.g., semi-strong form

efficiency as defined by Fama, 1970). However, in practice some investors may not

monitor all public information. We argue that even an investor who does not monitor the

public information should be interested in a conditional performance measure, because it

reveals the source of the manager's performance.

Consider the relation between unconditional and conditional weight-based

measures,

$$\sum_{j=1}^{N} Cov(\Delta w_j, r_j) = \sum_{j=1}^{N} E\{Cov(\Delta w_j, r_j \mid Z)\} + \sum_{j=1}^{N} Cov(E(\Delta w_j \mid Z), E(r_j \mid Z)), \quad (4)$$

where $\Delta w_j \equiv w_{jt} - w_{bjt}$. The left-hand side is the unconditional weight measure (UWM),

similar to Grinblatt and Titman (1993). The second term on the right-hand side of Eq. (4)

---

[6] Consider when a manager follows a buy-and-hold strategy starting with an initial set of weights w_{j0},
where j=1,...,N. At time t, the weight for security j would satisfy

$$\ln w_{jt} = \ln w_{j0} + \sum_{t=1}^{t} \ln\left(\frac{1 + r_{jt}}{1 + r_{pt}}\right),$$

where $r_{pt} \equiv \sum_j w_{jt-1} r_{jt}$. If the log excess returns follow a martingale difference sequence, the weights

w_{jt} follow a nonstationary, $I(1)$ process. Differenced weights, however, are stationary.

is the average conditional weight measure, equal to the unconditional mean of Eq. (3). The third term captures the effects of mechanical trading, based on the public information, Z. By comparing the conditional and unconditional measures, the third term may be calculated as a residual. This decomposition is useful for understanding the source of a portfolio's performance.

Consider an example of a manager with a two asset portfolio, one risky and one riskless, and a single-period investment horizon. Suppose the manager gets a signal that the return on the risky asset for next period is expected to be higher than average (the asset's unconditional mean). In response, the manager raises the weight on the risky asset in his portfolio, resulting in a positive unconditional weight measure in Eq. (4). The interpretation of this measured performance depends on whether or not we maintain the assumption of semi-strong form market efficiency.

If markets are efficient, an investor cannot use publicly available information to generate abnormal returns. If the manager has only public information, the CWM is zero and a positive UWM comes from the right-hand term in Eq. (4). If the manager has private information, a positive UWM may indicate the use of either public or private information. A nonzero CWM indicates the manager is using more than the public information.

Comparing the conditional and unconditional measures, an investor can decompose the manager's return from active trading into a component attributable to private information and a component attributable to public information. These are the first and second terms on the right-hand side of Eq. (4). If the investor does not monitor the public information, the public information component can help to evaluate the result of the

decision to delegate the monitoring of public information to the manager. Its magnitude may be compared with the investor's cost of monitoring public information. The private information component is the performance the investor could not obtain without the manager, even if he chose to monitor the public information. Isolating this component enables an investor to compensate a manager for his use of private information.

If markets are not semi-strong efficient, then a public information signal indicating a higher expected return to the risky asset may not be associated with a rise in the required return. The higher expected return may be due to some market inefficiency. Thus, raising the weight can increase the total return on the portfolio without changing the return the market requires to invest in the portfolio of assets. A conditional performance measure would still be of interest to investors concerned with the source of the abnormal returns. By comparing the conditional and the unconditional measures, an investor could infer whether the manager's performance came from taking advantage of market inefficiencies discernible from the public information, or from employing private information.

2.3. Advantages of the conditional weight measure

The CWM has a number of advantages. Grinblatt and Titman (1993) note four cases in which an unconditional weight-based measure may indicate superior performance when none exists: (1) the investor targets stocks whose expected return and risk have risen temporarily (e.g., stocks subject to takeover or bankruptcy); (2) the investor exploits serial correlation in stock returns; (3) the investor exploits the January effect; and (4) a manager gradually changes the risk of his portfolio over time.

These problems may be addressed using a conditional approach because they all involve strategies based on publicly available information. In this paper, we include variables to adjust for serial correlation and the January effect. Other conditioning variables may be important to include in a conditional performance measure. Thus, we use other publicly available predictor variables as well. The sample used here does not include any funds that specialize in takeover or bankruptcy stocks.

The conditional weight-based approach can control a potentially important bias inherent in returns-based measures. This bias, the interim trading bias, arises when we depart from the assumption that returns are independently and identically distributed over time (iid) and is therefore especially relevant to a conditional setting. The problem arises when managers may trade between the dates over which returns are measured.

Consider when returns are measured over two periods, but a manager trades each period. The manager has neutral performance, but the portfolio weights for the second period can be a function of public information at the intervening date. If returns are iid, this creates no bias, as there is no information at the intervening date that is correlated with the second period return. However, if expected returns vary with public information, then a manager who observes and trades on public information at the intervening date generates a return for the second period from the conditional distribution. His two-period portfolio strategy will contain more than the public information at the beginning of the first period, and a returns-based measure over the two periods will detect this as superior information. Even a conditional returns-based measure will suffer from this interim trading bias, if it can only condition on information at the beginning of the first period. Goetzmann, Ingersoll, and Ivkovic (2000) devise a partial solution for interim trading bias

in returns-based measures, in which they simulate the multiperiod return generated by interim trading options.

A conditional weight-based measure examines the conditional covariance between the manager's weights at the beginning of the first period and the subsequent two-period returns. If the manager has no information beyond the public information at the beginning of the first period, this conditional covariance would be zero. The ability of the manager to trade at the intervening period thus creates no interim trading bias in a conditional weight-based measure.

Managers may engage in interim trading based on superior information to enhance performance. A weight-based measure will not record these interim trading effects. Interim trading thus presents a bias under the null hypothesis that managers possess only public information. Under the alternative hypothesis of superior ability, a weight-based measure should sacrifice some power to detect the ability. Thus, the cost of using a weight-based measure to avoid bias is a potential loss of power. However, compared with returns-based measures, the use of the portfolio weight data may result in improved power. The tradeoff becomes an empirical question. We find that the CWM is no less precise than returns-based measures, even those with conditioning information. At the same time, the CWM is less susceptible to biases from interim trading on public information. Thus, the empirical tradeoff favors the CWM over the other measures.

3. Implementing the CWM

In this section, we describe how we make the measures operational. First we explain our assumptions and choice of the benchmark weights. Then, we describe our econometric methods.

3.1. An empirical conditional weight measure

The conditional weight measure at time t is

$$CWM_t = E\left\{ \sum_{j=1}^{N} \left(\widetilde{w}_{jt} - w_{bjtk}\right)\left(\widetilde{r}_{jt+1} - E\left(\widetilde{r}_{jt+1} \mid Z_t\right)\right) \mid Z_t \right\}, \tag{5}$$

where $w_{bjtk} = w_{jt-k} \prod_{t=t-k+1}^{t}\left(\dfrac{1+\widetilde{r}_{jt}}{1+\widetilde{r}_{pt}}\right)$ is the benchmark weight of asset j at time t, based on a

lag of k periods, k being the number of periods until the weights become public information. The benchmark return for the manager is formed from the actual lagged weights of the fund at t - k, updated using a buy-and-hold strategy, where r_{pt} is a buy-and-hold portfolio return, starting with the weights $\left\{w_{jt-k}\right\}$, j=1…N, at time t - k. The performance benchmark is internalized, based on the manager's previous holdings. If an external benchmark is desired, one need only substitute those benchmark weights for w_{bjtk}.

With the buy-and-hold benchmark, we are measuring the deviations between a manager's weights and a strategy of no trading during the previous k periods. A manager with no information might follow a buy-and-hold strategy. The benchmark weight w_{bjtk} can also be motivated as a proxy for $E\left(w_{jt} \mid Z_t\right)$, assuming the manager with only public

information follows a buy-and-hold strategy. Because Eq. (5) is a conditional covariance, this is a natural interpretation. However, w_{bjtk} does to have to be a good proxy for $E(w_{jt} | Z_t)$. We use different values of k because of the uncertainty as to when we may consider the lagged weights to be public information.

3.2. An unconditional weight measure

For comparison purposes, we compute an unconditional weight-based measure,

$$UWM = E\left\{ \sum_{j=1}^{N} (\tilde{w}_{jt} - w_{bjtk})(\tilde{r}_{jt+1} - E(r_j)) \right\}. \tag{6}$$

This differs from Grinblatt and Titman's (1993) version of the measure in two ways. First, we use a buy-and-hold strategy to measure the weight change instead of the lagged weight itself. Grinblatt and Titman (1993) use the lagged weight, which implicitly assumes a rebalancing strategy. They indicate they also investigated the effects of using a buy-and-hold strategy during the quarter. Second, we use the demeaned returns, where $E(r_j)$ is the mean. If $E(\tilde{w}_{jt} - w_{bjtk}) \neq 0$, then $E((\tilde{w}_{jt} - w_{bjtk})\tilde{r}_{jt+1}) \neq Cov((\tilde{w}_{jt} - w_{bjtk}), \tilde{r}_{jt+1})$. Demeaning the returns allows for nonzero mean changes in the weights, as $E((\tilde{w}_{jt} - w_{bjtk})(\tilde{r}_{jt+1} - E(r_j))) = Cov((\tilde{w}_{jt} - w_{bjtk}), \tilde{r}_{jt+1})$ even when $E(\tilde{w}_{jt} - w_{bjtk}) \neq 0$. An analogous situation arises with the CWM. In that case, we demean the returns using the conditional mean.

3.3. Estimating the measures

For the CWM, we estimate the conditional expected returns of the individual

assets as

$$E\left(\widetilde{r}_{jt+1} \mid Z_t\right) = \mathbf{b}_j ' \mathbf{Z}_t , \qquad (7)$$

where $\mathbf{Z}_t$ is an L vector of information variables, including a constant. We assume that a

linear regression delivers the conditional expectation, following most of the literature on

conditional performance evaluation.

Next, we assume that the conditional covariance between the weight changes and

abnormal returns is approximated by a linear function of the conditioning information.

This assumption follows Christopherson, Ferson, and Glassman (1998a) and can be

motivated by a Taylor series approximation. This results in

$$CWM_t = CWM + \mathbf{g}' \mathbf{z}_t , \qquad (8)$$

where $\mathbf{z}_t = \mathbf{Z}_t - E(\mathbf{Z})$, excluding the constant. In Eq. (8), CWM is the average

conditional covariance, $E(CWM_t)$. This provides an estimate of the second term of the

decomposition of Eq. (4).

Our linearity assumptions follow much of the conditional asset pricing literature.

Like most of this literature, we ignore model error that may arise if linearity is a poor

approximation. See Eqs. (7) and (8). We also ignore the errors that arise from having an

incomplete set of information variables. If we could account for these additional errors in

the expectations of the model, the precision of our estimates would be reduced. However,

fully accounting for these forms of model error remains a challenge for future research. In

the empirical analysis, we consider the introduction of nonlinear terms. We include the

squares of the elements of Z_t in regressions such as Eq. (10) and test for their exclusion with F tests. We find no evidence to support a nonlinear specification.

We set up a generalized method of movements (GMM) system to carry out the estimation.

$$\tilde{e}_{jt+1} = \tilde{r}_{jt+1} - \mathbf{b}_j ' \mathbf{Z}_t , \tag{9}$$

$$\tilde{e}_{cwmt+1} = \sum_{j=1}^{N} \left(w_{jt} - w_{bjtk} \right) \left(\tilde{r}_{jt+1} - \mathbf{b}_j ' \mathbf{Z}_t \right) - \mathbf{g}' \mathbf{z}_t - CWM , \tag{10}$$

$$\tilde{e}_{jt+1}^{u} = \tilde{r}_{jt+1} - E\left(\tilde{r}_j \right), \tag{11}$$

$$\tilde{e}_{uwmt+1} = \sum_{j=1}^{N} \left(w_{jt} - w_{bjtk} \right) \left(\tilde{r}_{jt+1} - E\left(\tilde{r}_j \right) \right) - UWM . \tag{12}$$

The parameters of the model are $\mathbf{b}$, γ, CWM, UWM, and $E(\tilde{r})$, where $\mathbf{b}$ is the stacked $\mathbf{b}_j$s and $E(\tilde{r})$ is the stacked $E(\tilde{r}_j)$s. We form the set of moment conditions as

$$g_t = \begin{pmatrix} \mathbf{e}_{t+1}^{c} \otimes \mathbf{Z}_t \\ e_{cwmt+1} \otimes \mathbf{z}_{ct} \\ \mathbf{e}_{t+1}^{u} \end{pmatrix}, \tag{13}$$

where $\mathbf{e}_{t+1}^{c}$ is an N vector of error terms for the N asset equations in Eq. (9) and

$\mathbf{z}_{ct} = \begin{pmatrix} 1 \\ \mathbf{z}_t \end{pmatrix}$. $\mathbf{e}_{t+1}^{u}$ is an N + 1 vector of error terms from Eqs. (11) and (12). The vector of

sample moment conditions g is an (N + 1)(L + 1) x 1 vector, $g = 1/T \sum_{t=1}^{T} g_t$. The GMM

estimates are obtained by minimizing the quadratic form, $g'Wg$. Hansen (1982) shows

that the parameter estimates will be T-consistent and asymptotically normally distributed for any fixed positive-definite W.

The GMM system is exactly identified with $(N+1)(L+1)$ parameters and $(N+1)(L+1)$ orthogonality conditions. The GMM parameter estimates for the system may be obtained as follows. First, consider Eqs. (9) and (11). As shown in Appendix A, the GMM parameter estimates are equivalent to those obtained using ordinary least squares (OLS) on Eq. (9) and using the sample mean for $E(\tilde{r}_j)$. The remaining parameter estimates may be obtained by running an OLS regression of $\sum_{j=1}^{N}(w_{jt} - w_{bjtk})(\tilde{r}_{jt+1} - \mathbf{b}_j'\mathbf{Z}_t)$

on $\mathbf{z}_{ct}$, and setting UWM equal to the sample mean of $\sum_{j=1}^{N}(w_{jt} - w_{bjtk})(\tilde{r}_{jt+1} - E(\tilde{r}_j))$.

Appendix A shows that these point estimates are identical to the GMM estimates based on the full system.

Setting this problem in a GMM context has several advantages. The multiplicative terms in Eqs. (10) and (12) suggest that the error terms will be non-normal, and the GMM justifies the estimators without normality, allowing general forms of conditional heteroskedasticity. Another important advantage is the standard errors. The standard errors account for conditional heteroskedasticity that may be an arbitrary function of $\mathbf{z}_{ct}$. The system also allows accounting for the covariance between the estimates of CWM and UWM. This is important in computing a standard error for the difference between the two measures for the decomposition of performance in Eq. (4).

While estimating the full GMM system, as $T > (N + 1)(L + 1)$, is not feasible, the linear and block-diagonal structure of the system allows us to derive asymptotic standard

errors for the UWM and CWM that account for the estimation error in **b** and $E(\tilde{r})$. Let

$$
\mathbf{d} = \left(\begin{array}{cc} \left[\displaystyle\sum_{t=1}^{T} \mathbf{z}_{ct} \mathbf{z}_{ct}^{'} \right]^{-1} \left[\displaystyle\sum_{t=1}^{T} (\mathbf{w}_t - \mathbf{w}_{bt})^{'} \otimes \mathbf{z}_{ct} \mathbf{z}_{ct}^{'} \right] & \mathbf{0} \\ \mathbf{0} & \dfrac{1}{T} \displaystyle\sum_{t=1}^{T} (\mathbf{w}_t - \mathbf{w}_{bt})^{'} \end{array} \right), \text{ where } \mathbf{w}_t \text{ is the vector}
$$

of portfolio weights at t, $\mathbf{w}_{bt}$ is the vector of benchmark weights at t, and

$$
Cov(\mathbf{B}) = E\left[\begin{pmatrix} \mathbf{b} \\ E(\tilde{r}) \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ E(\tilde{r}) \end{pmatrix}^{'} \right] \text{ is a consistent estimator of the covariance matrix of the first}
$$

stage parameters. The standard errors for $\hat{\mathbf{a}} \equiv \begin{pmatrix} g \\ CWM \\ UWM \end{pmatrix}$ are derived in Appendix A as

$$
Cov(\hat{\mathbf{a}}) = \left[T \left(\frac{\partial g_2}{\partial \mathbf{a}} \right)^{'} Cov(g_2)^{-1} \left(\frac{\partial g_2}{\partial \mathbf{a}} \right) \right]^{-1} + \frac{1}{T} \mathbf{d} Cov(\mathbf{B}) \mathbf{d}^{'}. \tag{14}
$$

Examination of our conditional measures using artificial portfolios of only five assets and one conditioning variable, which allows us to compare the full system GMM standard errors to the solutions given by Eq. (14) is similar.

4. Data description

Three data sets are used for this study. The first consists of the quarterly portfolio holdings and reported quarterly returns for a sample of U.S. equity pension fund managers; the second, individual security returns; and the third the conditioning variables.

4.1. Pension manager data

We use the quarterly holdings and returns of 60 domestic equity portfolios during the time period from December 31, 1984 through December 31, 1994. Each of these portfolios was actively managed by a professional money manager for one of the Investment Measurement Services (IMS) clients of Callan Associates Inc. Quarterly holdings for some pension fund managers are available through Securities and Exchange Commission filings, though much of that is composite data; that is, the results of multiple portfolios are lumped together. The data here provide an opportunity to examine the performance of individual U.S. equity pension fund portfolios. The portfolios have between 20 and 40 quarters of data. The starting date for each portfolio corresponds to either when the plan sponsor became an IMS client, when the manager was newly hired by the plan sponsor, or the data were first gathered. No backfilling of data took place when managers are hired, which mitigates some selection biases. The ending date corresponds to either when the plan sponsor ceased being an IMS client, when the manager was terminated by the plan sponsor, when that the manager ceased operations, or the last quarter of 1994. Unlike most pension fund databases, ours includes managers that were discontinued prior to the end of the sample period. Thus, our sample provides some control of survivorship bias.

The capitalizations of the portfolios as of the end of either the last quarter available or September 30, 1994 ranged from $3.2 million to $1.5 billion with a median of $78.9 million. The number of positions, or individual stocks held by a fund, as of the end of the sample ranged from 26 to 165 with a median of 62. There are 22 growth managers and 29 value managers. One of the portfolios is a balanced fund. These designations are

provided to Callan Associates by the client and normally coincide with the manager's self-declared style.

4.2. Representativeness

Given that our sample consists of portfolios being managed for clients of a single consulting firm, the question arises as to whether the sample is unusual relative to the larger universe of pension portfolios. Data on pension fund returns have been used in several previous studies. In all cases, the data are self-reported by the fund manager to a pension consulting firm. These firms typically do not make the data available to the general public. Lakonishok, Shleifer, and Vishny (1992) use data from SEI. Coggin, Fabozzi, and Rahman (1993) and Christopherson, Ferson, and Glassman (1998a) use Frank Russell Co. (FRC) data. Delguercio and Tkac (1999) and Coggin and Trzcinka (1999) use data from Mobius. Ikenberry, Shockley, and Womack (1998) use data from DeMarche Associates. No published study has compared the characteristics of the data from these various sources in any detail.

We compare the quarterly reported return characteristics of our sample with data from the Frank Russell Co. prouded in Myers (1999). Panel A of Table 1 contains summary statistics from the Callan Associates data; panel B summarizes the FRC data. The reported returns are measured net of trading costs, but management fees have not been subtracted. The statistics include means, standard deviations, minimum, median, and maximum values for the quarterly returns of equally weighted portfolios. The equally weighted portfolios comprise all funds, surviving funds, and discontinued funds.

Discontinued funds are defined as those that do not exist in the fourth quarter of 1994.

We also include equally weighted portfolios for the categories, grouping the funds by style

(i.e. value and growth).

Insert Table 1

The data shown in Table 1 show that the funds in the FRC sample have slightly

higher average returns and standard errors than those in the Callan Associates sample.

However, the ratios of the returns to the standard errors are similar. Using all funds we

have $\bar{r}/s = 3.92/8.02 = 0.49$ in the FRC data; $\bar{r}/s = 3.69/7.40 = 0.50$ in the Callan

Associates data. The lower returns and standard deviations may reflect higher cash

holdings by the Callan Associates funds. We are told by Frank Russell representatives that

the firm encourages its funds to employ custodial, cash-sweeping services and that cash

holdings are typically small. There are larger cash holdings in some of the Callan

Associates funds.

The average reported returns of growth funds exceeds value funds over this period

(by 0.19% in Callan Associate, 0.47% in FRC). The average returns of discontinued

funds exceed surviving funds in both data sets: by 0.31% in Callan Associates and 0.15%

in FRC. This is not an artifact of our particular sample period, as Myers (1999) finds a

similar result for a sample of FRC funds with returns data from 1979. This contradicts the

usual assumption about mutual fund survival; that is, that the low-return funds are less

likely to survive. In pension fund databases, high-return funds may be discontinued when

a star manager leaves to form a new firm [see Christopherson, Ferson, and Glassman,

(1998a) for a discussion of these differences and Myers (1999) for an empirical analysis of

survival and persistence in pension fund performance].

With 40 quarterly returns, the standard error of an average return in Table 1 is about $s / \sqrt{T} = 1.25\%$. Thus, none of the differences between these two data sets is statistically significant. The similarities between the two samples indicate that the Callan Associates sample is probably not atypical relative to the larger universe of pension portfolios.

Our sample period, from 1985 to 1994, may not fairly represent a longer historical record. However, unlike returns-based measures, our weight-based measures should not be subject to biases as a result of long swings in the relative performance of different investment styles. This is because the benchmark returns are not constructed from style indexes, but from the managers' lagged weights.

4.3. The timing of benchmark weights

We assume that the benchmark weight w_{bjtk} in Eqs. (5) and (6) is in the public information set at time t. It is constructed using the actual portfolio weight at t - k. Thus, the date when past weights become public information is an issue, and this may differ depending on the circumstances.

Mutual funds' portfolio weights become publicly available each quarter, although with a reporting lag in some cases. In our application to pension funds, one may view the public information as that available to pension plan sponsors, if performance is being measured from their viewpoint. If a plan sponsor wishes to know the current holdings of a portfolio at any given time, a manager is likely to respond within days, if not hours. However, plan sponsors systematically examine holdings data on a less frequent basis. A

lag of one quarter may be a reasonable assumption; however, one could argue for a longer period. For example, a careful review of the holdings could take place on an annual basis with a more cursory review at quarterly reporting periods. The time when the weights become effectively contained in the relevant public information set is therefore not clear. We try various lags k to evaluate the sensitivity of the measures to this assumption.

4.4. Preliminary analysis of the portfolio holdings

To characterize differences in managers' trading styles, we compute measures of momentum investing and turnover. We refine a measure from Grinblatt, Titman, and Wermers (1995) to measure momentum

$$LM = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{N_t} \left(w_{jt} - w_{jt-1} \left(\frac{1 + r_{jt}}{1 + r_{pt}} \right) \right) r_{jt-1} \,, \tag{15}$$

where $r_{pt} = \sum_{j=1}^{N_t} w_{jt-1} r_{jt}$, and N_t is the number of positions held in the portfolio at time *t*.

Unlike the weight-based performance measures, the LM measure in Eq. (15) captures the covariance between a lagged return and a current weight change measured relative to a buy-and-hold strategy. If the manager, on average, raises the weights on those securities that had high returns last quarter, the LM measure will be positive. Thus, the measure captures whether a manager employs a momentum strategy in managing the portfolio.

Our implementation of the momentum measure differs from Grinblatt, Titman, and Wermers in three ways. First, they pair quarterly weight changes with each of the three monthly returns in the previous quarter. We use the lagged quarterly return. We also measure the weight changes relative to a buy-and-hold strategy as opposed to the simple changes. Thus, our measure avoids showing momentum for a buy-and-hold strategy. Finally, Grinblatt, Titman, and Wermers calculate their weights using the average of the beginning- and end-of-quarter share prices. We use the actual weight data, based on share prices for the date on which the holdings are measured.

Insert Table 2

Table 2 presents the momentum measures for the sample of managed portfolios. Panel A summarizes the LMs and their t-statistics by taking the means across the individual funds.[7] The data in Panel A give the impression that growth managers tend to be momentum investors while value managers tend to be contrarians. The results in Panel B show that an equally weighted portfolio of the growth managers also has a significantly positive LM, while the equally weighted portfolio of value managers has a significantly negative LM.

Panel C of Table 2 presents the medians and the proportions of managers with positive or negative momentum measures. Of 22 growth managers, 19, or 86%, here positive LMs, and 15 are significant at the 5% level. Of 29 value managers, 22, or 76%,

---

[7] The LM and standard error for any given portfolio are computed using the GMM, where the orthogonality condition is

$$e_{lmt} = \sum_{j=1}^{N}\left(w_{jt} - w_{jt-1}\left(\frac{1+r_{jt}}{1+r_{pt}}\right)\right)r_{jt-1} - LM \; ; \; E(e_{lmt}) = 0.$$

have negative LMs, and 15 are significant at the 5% level. We also compute a binomial

test on the proportions of managers with positive LMs.[8] The proportion of growth

managers with positive LMs is significantly greater than one-half, and the proportion of

value managers with positive LMs is less than one-half.

Overall, the data in Table 2 show that growth managers tend to follow momentum

strategies, while value managers tend to be contrarians. Grinblatt, Titman, and Wermers

(1995) find a similar result for mutual funds.[9] Thus, the relation between the value and

growth investment styles and momentum is not limited to the particular context of mutual

funds. The differences between growth and value funds will prove useful in evaluating the

different performance measures.

Turnover is another area of difference between growth and value managers. Our

measure of turnover is

$$Turnover = \frac{1}{2}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{Nt}\left|w_{it} - w_{it-1}\left(\frac{1+r_{it}}{1+r_{pt}}\right)\right|\right].$$ (16)

This measure examines the departure of a portfolio's actual weights at time t, w_{it}, from

the weights of a buy-and-hold strategy, starting from the actual weights at time t - 1.

---

[8] The binomial test is formed as follows:

$$t - stat = \frac{(P - 0.5)}{\sqrt{\frac{(0.5)(0.5)}{N}}}, \text{ where P equals the proportion} > 0.$$

If $P > 0.5$, then it tests whether the proportion positive is significantly above 0.5. If $P < 0.5$, then it tests
whether the proportion positive is significantly less than 0.5. We use a one-sided test.

Buying and selling within a quarter such that the end-of-quarter weights are unchanged will not be picked up by this measure.

Table 3 provides a comparison of the average quarterly turnover measures for growth and value managers. It shows that turnover is higher, on average, for growth managers than for value managers. Using a *t*-test for the difference in the means and a Wilcoxon-Mann-Whitney rank test, we can reject the null hypothesis that turnover is the same for both groups of managers at the 5% level. Thus, the growth managers tend to depart more from a buy-and-hold strategy than the value managers. On this basis, we may expect the interim trading bias of returns-based performance measures to differ for the two styles.

<div align="center">**Insert Table 3**</div>

4.5. Asset returns

The second data set contains returns for all the equity securities in the quarterly portfolio holdings data that could be associated with a security in the Center for Research in Security Prices (CRSP), NYSE, AMEX, and NASDAQ daily file. In addition, all identifiable cash holdings were summed and paired with the three-month Treasury bill return from the CRSP Fama riskfree file, and all of the identifiable bond holdings (which consist of mostly Treasury notes and bonds) were summed and paired with the 60- to 120-month bond portfolio returns from the CRSP Fama files. The fixed income holdings for

---

[9] A referee suggests a simple explanation. Because growth (value) funds invest in stocks with prices that tend to be high (low) relative to earnings, cash flow, etc., they are likely to be found moving into stocks with prices that have recently risen (fallen).

the majority of these portfolios are usually small, although the fixed income component is substantial for some portfolios and isolated periods for other portfolios.

Using the quarterly holdings data and the CRSP returns, we construct a hypothetical quarterly return series for each portfolio. The hypothetical quarterly returns assume a buy-and-hold strategy during the quarter. Appendix B contains a comparison between the reported quarterly returns in the Callan Associates data and the hypothetical quarterly returns we constructed using the portfolio weights and the CRSP data.

4.6. Instrumental variables

The publicly available information Z_t are predetermined variables that appear to predict stock returns over time according to previous research. These include a lagged dividend yield, default spread, term spread, T-bill yield, and a January indicator variable. We also include the lagged market return and the lagged return of individual stocks to capture the effects of serial correlation in returns.

The lagged dividend yield, DY(t), is the sum of the total dividends paid on the Standard & Poor's (S&P) 500 as reported in Citibase for the year preceding time t, divided by the value of that portfolio at time t. The default spread, DS(t), is defined as the difference between Moody's BAA-rated corporate bond yield and the AAA-rated corporate bond yield as reported in Citibase. The term spread, TS(t), is the difference between the constant maturity ten-year Treasury bond yield and the three-month Treasury bill yield as reported in Citibase. The three-month Treasury bill yield, TB(t), is also from Citibase.

These standard instruments, from the asset-pricing literature, are mostly marketwide variables. Thus, it naturally seems to follow that they would be more relevant to market timing or asset allocation than to active equity management. However, Ferson and Harvey (1999) find that asset-specific loadings on a similar list of variables have strong explanatory power for the cross-section of portfolio returns. Thus, by allowing stock-specific coefficients in the expected return models, we capture information about the cross-section of returns. The literature on momentum, following Jegadeesh and Titman (1993), suggests that own lagged returns also have explanatory power for the cross-section of returns. We conduct further experiments using firm-specific dividends and stock prices from CRSP to form additional instruments and find little difference in the results.

5. Empirical results

We first document the interim trading bias in returns-based measures and the performance using these measures for our sample of managers. This provides a comparison for the weight-based measures. The units of an alpha in a returns-based measure and the units of a weight-based measure are both excess returns per quarter.

5.1. Returns-based measures

The returns-based measures include the average returns in excess of a benchmark, unconditional alphas, and conditional alphas. We use the CRSP value-weighted index as the benchmark in Panel A of Table 4. Panel B, we use style indexes for value or growth portfolios. The value (growth) index is the average of a small-cap and a large-cap, high (low) book-to-market portfolio constructed as in Fama and French (1993, 1996).[10] For the conditional alphas, we use the methodology of Ferson and Schadt (1996) to incorporate the conditioning information. This involves estimating

$$\tilde{R}_{pt+1} = a_p + b_{0p}\tilde{R}_{mt+1} + \mathbf{b}_p'\left(\mathbf{z}_t \tilde{R}_{mt+1}\right) + \tilde{e}_{pt+1}, \tag{17}$$

where $\tilde{R}_{pt+1}$ is the excess return on portfolio p; a_p is the intercept term, the measure of alpha; b_{0p} is the average conditional beta; $\tilde{R}_{mt+1}$ is the excess return on the benchmark portfolio; $\mathbf{b}_p$ is the vector of conditional beta response coefficients with respect to the public information set; $\mathbf{z}_t$ is the vector of demeaned public information variables; and $\tilde{e}_{pt+1}$ is the regression error.

Insert Table 4

5.2. Interim trading bias

The first line of Table 4 presents the results of applying the returns-based measures to the quarterly returns of hypothetical, interim trading portfolios. The interim trading portfolios use only public information to trade each month within a quarter. The portfolios start with an initial set of 500 stocks randomly selected from the CRSP, NYSE,

---

[10] Our data for the size and book-to-market portfolios are the same as used by Ferson and Harvey (1999).

AMEX, and NASDAQ tape, with prices initially greater than or equal to five dollars per share. The initial weights are in proportion to market capitalization. The weights then evolve each month according to

$$w_{jt} = w_{jt-1}\left(\frac{1 + E\left(r_{jt+1} \mid Z_t\right)}{1 + E\left(r_{pt+1} \mid Z_t\right)}\right),$$ (18)

where $E\left(r_{pt+1} \mid Z_t\right) = \sum_{j=1}^{N} w_{jt} E\left(r_{jt+1} \mid Z_t\right)$. The $E\left(r_{jt+1} \mid Z_t\right)$ is modeled by a regression on the lagged instruments. Because the hypothetical portfolios do not consider the return volatility directly, they would not be optimal for risk averse investors. However, in many models, investors' optimal weights are linear in the conditional mean return (e.g., exponential utility with conditional normality), so our examples should provide a reasonable illustration of the potential biases. The data in Table 4 show the average results for 100 interim trading portfolios. Each portfolio starts with a different random selection of the initial stock holdings.

Interim trading bias has a significant impact on the returns-based performance measures. The average unconditional alpha of the strategies, relative to the value-weighted CRSP index, is 1.03% per quarter with an average t-ratio of 3.64. A conditional model seems to have little effect on the bias. The average conditional alpha is 1.00% with an average t-ratio of 3.27. Avoiding this severe interim trading bias provides a strong motivation for the conditional weight-based measures.

5.3. Returns-based fund performance

We summarize the returns-based measures for the funds in three ways. First, we look at cross-sectional means of the average performance measures. These weight each manager's results equally and, therefore, put greater weight on dates at the end of the sample period, which has more managers. Second, we examine equally weighted portfolios of managers. Because the managers have variable data histories, the weight a manager's performance receives is directly related to the length of the data series, and each date receives equal weight. Finally, we examine the median performance. If the cross-section of returns is skewed, means and medians can imply different inferences about fund performance as emphasized by Ikenberry, Shockley, and Womack (1998). Here we get similar inferences from the different statistics.

The averages of the returns-based performance measures, taken across groups of individual funds, are percent per quarter. All the measures are positive, with the exception of the excess return of value funds relative to the style index. For equally weighted portfolios the point estimates of performance are similar and are statistically significant in seven of the 15 cases. Finding significant positive returns-based alphas for pension funds is consistent with Coggin, Fabozzi, and Rahman (1993) and Christopherson, Ferson, and Glassman (1998b). With the exception of value portfolios using a style index, no significant difference exists between the conditional and unconditional returns-based alphas at the aggregate levels, which is consistent with Christopherson, Ferson, and Glassman (1998b).

The performance of the median manager provides a qualitatively similar picture. The median performance is positive for each group with one exception and statistically significant in all but two cases. Finally, Table 4 reports binomial tests, which find that the

proportion of positive alphas is significantly greater than 50% in 13 of the 15 groups of portfolios.

Given that the returns-based measures attribute abnormal performance to the interim trading portfolios, the evidence for the funds is difficult to interpret. The alphas of the interim trading portfolios are larger than the alphas that we measure for the actual funds. However, the actual funds pay transaction costs while the interim trading portfolios do not. (Pension plan sponsors pay additional management fees that are not reflected in these returns.) The returns-based measures are also sensitive to the benchmark. For example, using the S&P 500 and the equally weighted fund portfolios, value funds outperform growth funds and produce a significant unconditional alpha (t-ratio, 2.21). Using the style benchmarks, growth funds outperform value funds by 0.5% per quarter. Internalizing the benchmark, based on a fund's lagged weights, is thus an attractive feature of the weight-based measures.

5.4. Weight-based measures meet artificial strategies

We next present performance measured with the CWM and UWM. The difference, UWM – CWM is our estimate of $\sum_{j=1}^{N} Cov\big(E(\Delta w_j \mid Z), E(r_j \mid Z)\big)$, the public information term in the decomposition of Eq. (4). Table 5 uses the weights lagged one quarter to construct the benchmark (k = 1). The expected return equation for each security uses the lagged security return as an element of Z_t in Eq. (9), and the expected return is estimated using available data for 1984-94. We also measure performance with

the weights lagged two quarters (k = 2), along with the two lagged quarterly returns in the information set in Eq. (9). We do not report these results in a table, as they are similar to the k = 1 case. Table 6 uses the weights lagged four quarters (k = 4) with the four lagged quarterly returns. The CWM in Eq. (10) includes a lagged market return, but no lagged individual security returns.

Insert Tables 5& 6 near here

The data in Panel A show the results for the interim-trading portfolios. The measured performance is much lower than with the returns-based measures. However, the performance measures are statistically significant for the UWM at 0.25% per quarter. Using the conditioning information in the CWM lowers the point estimate to –0.03% and eliminates any significant performance. This shows that the CWM can control the interim trading bias. The differences UWM - CWM are positive and strongly significant, which confirms that $Cov(E(\Delta w \mid Z), E(r \mid Z)) > 0$ for these strategies.

The insignificant performance of the interim trading strategy under the CWM is not due to a noisier measure. The conditional weight-based measures appear more precise than the returns-based measures. For example, a conditional alpha of about 0.5% per quarter would get a t-ratio of two in Table 4, with returns-based measures. While none of the CWMs is significant, it would take an abnormal return of less than 0.2% per quarter to achieve a t-ratio of two in Table 5. The conditional weight-based measures appear precise and well specified in the face of interim trading, which increases our confidence in using them to measure actual fund performance.

5.5. Weight-based performance of pension funds

Panel B of Tables 5 and 6 displays averages for the UWM and CWM in the manager sample. The growth managers show small positive performance under all versions of the UWM. Their performance under the CWM is close to zero. The alphas are typically much closer to zero than with the returns-based measures. The weight-based measures thus produce smaller t-ratios, in spite of their smaller standard errors. Given their greater precision and the ability of the conditional weight-based measures to control the interim trading bias, the inference of no abnormal performance using the CWM is more compelling than the returns-based measures of abnormal performance. Thus, our evidence supports the hypothesis that the funds have no abnormal performance, once we control for public information and interim trading bias.

Considering equally weighted portfolios of the funds (all, growth, and value), the point estimates of performance are similar to the previous averages. The growth managers appear to fair slightly better under the UWM than the CWM. The value managers, meanwhile, appear to fair better under the conditional measures.[11]

The differences between the two measures, UWM - CWM, are the covariances between the portfolio weight changes and the conditional expected returns given the public information, Z. These differences are positive for growth funds and negative for value funds, but not significant. This makes sense in view of the tendency of growth funds to follow momentum strategies, which imply a positive covariance with public

---

[11] We check the sensitivity of these results to the use of marketwide instruments by adding the individual stock dividend yield and the inverse of the price per share. Such cross-sectional instruments can be motivated by the model in Bossaerts and Green (1989). We find similar results with the addition of these instruments.

information, while value funds tend to be contrarian. These differences explain why the

growth funds look worse under the CWM, where they get no credit for $Cov(\Delta w, E(r\,|\,Z))$.

Similarly, value funds are not penalized for a negative covariance under the CWM, so their

performance typically looks better.

Panel D of Tables 5 and 6 displays the performance of the median managers for the

entire sample and the subsamples of growth and value funds. The median performance

numbers mirror those of Panels B and C. Panel D also reports a binomial test, which finds

the proportion of growth managers with positive results to be significantly greater than

50% when measuring performance with the UWM, but not significantly different from

50% when measuring performance with the CWM. For value managers, the proportion

with positive abnormal returns is not significantly different from 50% in either model.

5.6. The question of power

The CWM is less susceptible to biases from interim trading on public information

than the UWM or the returns-based measures. However, by avoiding interim trading,

weight-based measures may sacrifice power to detect superior ability. A comparison of

the standard errors indicates that the CWM has good precision, suggesting that the use of

the additional portfolio weight information may mitigate the loss of power. However, low

power to detect ability may be reflected in small point estimates of the performance, so the

standard errors are not sufficient to evaluate power.

To examine the relative power, we create monthly interim trading portfolios with

superior information. These portfolios combine public information with information about

the actual future returns. We start with randomly selected stocks, such as the interim

trading portfolios. The weights then evolve each month according to

$$w_{jt} = w_{jt-1}\left(\frac{1+(r)r_{jt+1}+(1-r)E(r_{jt+1} \mid Z_t)}{1+(r)r_{pt+1}+(1-r)E(r_{pt+1} \mid Z_t)}\right).$$ (19)

We compound the monthly returns to quarterly returns for the returns-based measures and

sample the weights quarterly for the weight-based measures.

The parameter r indicates the level of embedded superior information and

controls the correlation between the log weight changes and the actual future return.

When $r = 1$, the weight changes fully anticipate future returns and the correlation is

perfect, while $r = 0$ is the previously analyzed interim trading portfolio, which has only

public information. In this case the correlation is the square root of the regression R-

squared for the return on the lagged instruments.

Table 7 lists the average measured performance taken across 100 interim trading

portfolios that use only public information ($r = 0$) and differ by the initial, random

selection of 500 stocks. Below this is the relative performance for three different levels of

r. The relative performance is the difference between the measured performance for the

superior information portfolios and the corresponding portfolios with $r = 0$.

Concentrating on the relative performance, we adjust for the biases in the various

measures under the null hypothesis of no performance ($r = 0$) and interim trading. Thus,

the relative performance provides information about the size-adjusted power of the

different measures.

The data in Table 7 show that, for every experiment, the CWM with k = 1 has the highest power, among all the measures, to detect superior ability. The t-ratio for the CWM exceeds that of the returns-based measures and that of the UWM for every nonzero level of r. Thus, any concerns about loss of power in the face of informed interim trading, relative to returns-based measures, do not appear to represent a liability for the CWM with k = 1. For larger values of k, the power of the CWM does appear to suffer relative to the returns-based measures. However, the average t-statistic remains larger than the UWM in all cases. This suggests that future applications of the CWM should involve care in choosing the benchmark weights.

5.7. Survivorship issues

Table 1 compares discontinued and surviving funds. Unlike the evidence for mutual funds, discontinued pension funds have higher average returns. Discontinued pension funds also have larger standard deviations in our sample.

Panel A of Table 8 provides the comparison between the risk-adjusted performance of discontinued and surviving funds. For all measures and styles (with one exception), the discontinued funds have higher risk-adjusted performance, consistent with the tendency of star performers to leave the sample. However, the differences are seldom statistically significant and the economic magnitudes of the differences are small. These

differences are unlikely to affect our main conclusions about the relative reliability and precision of the various performance measures. We still find that the typical performance is neutral under the CWM.

While our sample includes discontinued funds, the potential for a selection bias still exists because we use only managers with at least 20 quarters of holdings history. Managers may have fewer than 20 quarters of holdings history for several reasons. First, they may have been hired and then terminated within 20 quarters by a client because of poor performance; a change in the strategy, personnel, or requirements of the plan sponsor; or a change in the personnel of the money manager. Second, the pension client may have remained an IMS client but decided to receive performance data on a returns-only basis, which is a less costly service. Third, the pension client may have discontinued being an IMS client. This may have occurred for a variety of reasons, including the desire to hire a different consultant. And finally, a pension fund may have become an IMS client or a manager may have been hired by a client recently enough so that 20 quarters of data had not been accumulated by the final quarter of 1994.

While we have no data on the funds that do not meet the 20-quarter selection criterion, Panel B of Table 8 provides some information on the effects of sample selection. We split the sample into those funds with more than 30 quarters of data and those with 30 or fewer quarters. We compare the means and standard deviations of return over a common time period, 1990-94. The funds with the shorter data histories have somewhat higher returns (by 0.2% per quarter for all funds) and standard deviations (by 0.62% per quarter), indicating a mild selection bias. However, the selection bias is unlikely to change our inferences about the relative performance of the measures.

6. Robustness

One of the main differences between the UWM and the CWM is the use of unconditional versus conditional expected returns. This section explores the robustness of the results to the way the expected returns are measured. We report the results of two experiments. In the first, we replace the regression-based expected returns with expected returns from an asset-pricing model. In the second, we vary the instruments used in the regression-based expected returns.

Modeling expected returns by regression on the lagged instruments is likely to involve a substantial amount of noise. Similarly, using a security's past average return for the unconditional mean produces a noisy estimate. While the standard errors account for noise, an alternative approach is considered. Here, we use a simple capital asset pricing model (CAPM). The expected returns thus are $b_j E(R_{mt+1} | Z_t)$, where b_j is the simple regression beta for asset j and $E(R_{mt+1} | Z_t) = \mathbf{d}_m' \mathbf{Z}_t$ is the regression-based estimate for the market index.[12] In the UWM, we use $b_j E(R_m)$ in place of $E(r_j)$.[13] The results are in Panel A of Table 9. Overall, the results are similar to those obtained in Table 5, with growth manager performance under the CWM being closer to zero than with the UWM.

---

[12] Because the expected return in this model, $b_j \mathbf{d}_m \mathbf{Z}_t$, is nonlinear in the parameters, our solution for the standard errors does not apply. While the point estimates remain consistent, the standard errors in Panel A of Table 9 do not account for estimation of $\mathbf{d}_m$ and probably are understated.

In our second experiment we further examine robustness to the choice of the lagged instruments. We examine the distributions of the coefficients across the individual securities and find that several of the instruments produce significant coefficients relatively infrequently. The two most important instruments appear to be the lagged Treasury bill and dividend yield. Ferson and Schadt (1996) also find these to be the most important instruments for mutual fund returns. We therefore run the models again using only these two instruments. The results are in Panel B of Table 9. Again, we find results similar to Table 5.

7. Conclusion

We combine a weight-based approach to measuring performance with conditioning information to create a conditional weight-based performance measure. We examine this methodology and compare it with previous approaches. In view of biases associated with dynamic trading and the use of public information, this comparison is important and we document that the new measures are superior to previous approaches. Finally, we illustrate the use of the new measures on a small sample of U.S. equity pension fund managers, during 1985-94.

---

[13] We assume constant betas, as in the classical latent variable model of Gibbons and Ferson (1985). Ferson and Harvey (1991) argue that most of the time variation in conditional expected returns should

We find that the conditional weight-based measures can control an interim trading bias, caused by the ability to use publicly available information to trade between return measurement periods. Using conditioning information in a returns-based measure does not control for this bias, which can be substantial. Unconditional weight-based measures, as described by Grinblatt and Titman (1993), are also affected. By using instruments for public information combined with portfolio weights, our new measures provide relatively precise and reliable estimates of performance.

In our sample of equity pension fund managers, traditional returns-based measures imply positive and significant abnormal performance. However, the performance is smaller than the potential effects of interim trading bias on the returns-based performance measures. Growth-style funds, relative to value funds, tend to adopt a momentum trading strategy, which shows up as positive covariances between their weight changes and expected returns formed from public information. Under unconditional weight-based measures, growth funds appear to outperform value funds. However, using a conditional weight-based measure, the performance difference is reversed, and abnormal performance is not significant.

come from the market premium, not from the betas.

Appendix A

Following the notation in the text, we write the GMM system Eqs. (9), (10), (11), and (12) to isolate the first and second stage parameters as

$$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{T} \sum_t g_{1t} \\ \frac{1}{T} \sum_t g_{2t} \end{pmatrix},$$

(A.1)

where $g_{1t} = \begin{pmatrix} \mathbf{e}_{t+1}^c \otimes \mathbf{Z}_t \\ \tilde{\mathbf{r}}_{t+1} - E(\tilde{\mathbf{r}}) \end{pmatrix}$ and $g_{2t} = \begin{pmatrix} e_{cwmt+1} \otimes \mathbf{z}_{ct} \\ e_{uwmt+1} \end{pmatrix}$. The system is block triangular in

the parameters, as $g_1 = g_1(\mathbf{B})$ and $g_2 = g_2(\mathbf{B}, \mathbf{a})$, where $\mathbf{a} = \begin{pmatrix} \left(g', CWM \right)' \\ UWM \end{pmatrix}$ and

$\mathbf{B} = (\mathbf{b}', E(\tilde{r})')$. The system is also linear in the parameters and may be written as

$$g_1 = h_1 + \left(\frac{\partial g_1}{\partial \mathbf{B}} \right) \mathbf{B} \qquad \text{and}$$

(A.2)

$$g_2 = h_2 + \left(\frac{\partial g_2}{\partial \mathbf{a}} \right) \mathbf{a},$$

(A.3)

where h_2 is linear in $\mathbf{B}$ but does not depend on $\mathbf{a}$. Because the system in Eq. (A.1) is

exactly identified, the GMM parameter estimates are obtained by setting $g_1 = 0$ in Eq.

(A.2) and $g_2 = 0$ in Eq. (A.3). The solution gives $\hat{\mathbf{B}}$ as the first stage OLS estimator and

$\hat{\mathbf{a}}$ as the second stage estimator described in the text. The estimate of $\mathbf{a}$ may be

expressed

$$\hat{\mathbf{a}} = \mathbf{d}_0 - \mathbf{d}\hat{\mathbf{B}}, \tag{A.4}$$

where $\mathbf{d} = \begin{pmatrix} \left[\sum_{t=1}^{T}\mathbf{z}_{ct}\mathbf{z}_{ct}^{'}\right]^{-1}\left[\sum_{t=1}^{T}(\mathbf{w}_t - \mathbf{w}_{bt})^{'} \otimes \mathbf{z}_{ct}\mathbf{z}_{ct}^{'}\right] & \mathbf{0} \\ \mathbf{0} & \dfrac{1}{T}\sum_{t=1}^{T}(\mathbf{w}_t - \mathbf{w}_{bt})^{'} \end{pmatrix}$ and

$\mathbf{d}_0 = \begin{pmatrix} \left[\sum_{t}\mathbf{z}_{ct}\mathbf{z}_{ct}^{'}\right]^{-1}\left[\sum_{t}\mathbf{z}_{ct}(\mathbf{w}_t - \mathbf{w}_{bt})^{'}\mathbf{r}_t\right] \\ \dfrac{1}{T}\sum_{t}(\mathbf{w}_t - \mathbf{w}_{bt})^{'}\mathbf{r}_t \end{pmatrix}$. We use the delta method to obtain the

asymptotic variance of $\hat{\mathbf{a}}$, where

$$0 = g_2(\hat{\mathbf{B}}, \hat{\mathbf{a}}) \approx g_2(\mathbf{B}, \mathbf{a}) + \frac{\partial g_2}{\partial \mathbf{B}}(\hat{\mathbf{B}} - \mathbf{B}) + \frac{\partial g_2}{\partial \mathbf{a}}(\hat{\mathbf{a}} - \mathbf{a}). \tag{A.5}$$

The asymptotic covariances satisfy

$$Cov(g_2) = \frac{\partial g_2}{\partial \mathbf{B}}Cov(\hat{\mathbf{B}})\left(\frac{\partial g_2}{\partial \mathbf{B}}\right)^{'} + \frac{\partial g_2}{\partial \mathbf{a}}Cov(\hat{\mathbf{a}})\left(\frac{\partial g_2}{\partial \mathbf{a}}\right)^{'} + 2\frac{\partial g_2}{\partial \mathbf{B}}Cov(\hat{\mathbf{B}}, \hat{\mathbf{a}})\left(\frac{\partial g_2}{\partial \mathbf{a}}\right)^{'}. \tag{A.6}$$

Eq. (A.4) implies $Cov(\hat{\mathbf{B}}, \hat{\mathbf{a}}) = -Cov(\hat{\mathbf{B}})\mathbf{d}^{'}$. Substituting this expression into Eq. (A.6) and

simplifying gives the asymptotic variance of $\hat{\mathbf{a}}$. Thus, $\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a})$ converges to a normal

with mean zero and variance

$$Cov(\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a})) = \left[\left(\frac{\partial g_2}{\partial \mathbf{a}}\right)^{'}Cov(g_2)^{-1}\left(\frac{\partial g_2}{\partial \mathbf{a}}\right)\right]^{-1} + \mathbf{d}Cov(\hat{\mathbf{B}})\mathbf{d}^{'}. \tag{A.7}$$

Appendix B

The reported quarterly returns for the funds, which we use in our analyses of the returns-based measures, are computed by Callan Associate's IMS group based on daily holdings and transactions data, net of estimated trading costs. The reported returns are gross of management fees, however, because different clients may be charged different fees see Halpern and Fowler, (1991). The hypothetical returns assume that stocks held at the beginning of a quarter stay in the portfolio for the entire quarter with dividends reinvested. Khang (1997) examines the correlation between the reported and hypothetical returns as well as other summary statistics comparing the two return series.

Differences between the reported and hypothetical returns reflect four potential factors. First, stock purchases and sales within quarters are not included in the hypothetical return calculation. Second, the reported returns are net of estimated transaction costs, while the hypothetical returns reflect no transaction costs. Third, securities not identified are not included in the hypothetical return calculation. Fourth, the reported returns reflect fixed income returns on an individual security basis, while the hypothetical returns assume fixed income returns are equal to either the three-month Treasury bill return or the return on the 60- to 120-month bond portfolio from the CRSP Fama files.

Khang (1997) finds that the reported and hypothetical returns are similar for most of the portfolios. However, we cannot infer from this that interim trading effects are small because of the other differences between the reported and hypothetical returns. The standard deviations of the difference between the reported and hypothetical returns range from 0.48% to 3.22% per quarter. For 42 of the 60 funds, the standard deviation is below

1%. This compares with a standard deviation of 0.57% for the difference between the S&P 500 and the CRSP Value Weighted Index.

The average difference between the reported and hypothetical returns is 0.07% per quarter, although it ranges between -0.73% and 0.68% across the 60 funds. The differences in alpha average 0.08% and have a smaller range. Thus the averages for the reported portfolio returns are slightly higher than for the hypothetical portfolios, even though the reported returns are net of transaction costs while the hypothetical returns pay no transaction costs.   This may indicate that there are total return gains from intra-quarter trading that exceed the transaction costs borne by institutional investors. Evaluating whether these differences reflect differences in risk due to intraquarter trading is not possible, without higher frequency data.

References

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

Bossaerts, P., Green, R., 1989. A general equilibrium model of changing risk premia: theory and tests. Review of Financial Studies 2, 467-493.

Christopherson, J., Ferson W., Glassman D., 1998a. Conditioning manager alphas on economic information: another look at the persistence of performance. Review of Financial Studies 11, 111-142.

Christopherson, J., Ferson W., Glassman D., 1998b Conditional measures of performance and persistence for pension funds. In: Research in Finance, vol. 16. JAI Press, Stamford, Connecticut., pp. 1-46.

Coggin, T., Fabozzi, F., Rahman S., 1993. The investment performance of U.S. equity pension fund managers. Journal of Finance 48, 1039-1056.

Coggin, T., Trczinka C., 1999. The performance of institutional U.S. equity portfolio managers. Unpublished working paper, State University of New York at Buffalo.

Copeland, T., Mayers D., 1982. The value line enigma (1965-1978): a case study of performance evaluation issues. Journal of Financial Economics 10, 289-321.

Cornell, B., 1979. Asymmetric information and portfolio performance measurement. Journal of Financial Economics 7, 381-390.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035-1058.

Delguercio, D., Tkac P., 1999. The determinants of the flow of funds of managed portfolios: mutual funds versus pension funds. Unpublished working paper. University of Oregon.

Eckbo, B., Smith, D., 1998. The conditional performance of insider trades, Journal of Finance 53, 467-498.

Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25, 383-417.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55-87.

Ferson, W., Harvey C., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385-415.

Ferson, W., Harvey C., 1999. Conditioning variables and cross-section of stock returns. Journal of Finance 54, 1325-1360.

Ferson, W., Schadt, R., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51, 425-461.

Gibbons, M., Ferson, W., 1985. Testing asset pricing models with changing expectations and an unobservable market portfolio. Journal of Financial Economics 14, 216-236.

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

Graham, J., Harvey, C., 1996. Market timing ability and volatility implied in investment newsletters' asset allocation recommendations. Journal of Financial Economics 42, 397-422.

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

Grinblatt, M., Titman, S., 1989b. Portfolio performance evaluation: old issues and new insights. Review of Financial Studies 2, 393-421.

Grinblatt, M., Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. Journal of Business 60, 97-112.

Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. American Economic Review 85, 1088-1105.

Halpern, P., Fowler, I., 1991. Investment management fees and the determinants of pricing and structure in the industry. Journal of Portfolio Management, Winter, 74-79.

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

Ikenberry, D., Shockley, R., Womack K., 1998. Why active fund managers often underperform the S&P 500: the impact of size and skewness. Journal of Private Portfolio Management, vol.1 issue 1, 13-26.

Jegadeesh, N., Titman S., 1993. Returns to buying winners and selling losers. Journal of Finance 48, 65-91.

Jensen, M., 1968. The performance of mutual funds in the period 1945-1964. Journal of Finance 23, 389-416.

Khang, K., 1997. Performance measurement using portfolio weights and conditioning information: an examination of pension fund equity manager performance. Unpublished Ph.D. dissertation. University of Washington.

Lakonishok, J., Shleifer A., Vishny R., 1992. The structure and performance of the money management industry. Brookings Papers on Economic Activity: Microeconomics 1992, 339-391.

Myers, D., 1999. Survival and the persistence of pension fund performance. Unpublished working paper. University of Washington.

Rubinstein, M., 1973. A comparative statics analysis of risk premiums. Journal of Business 46, 605-615.

Wermers, R., 1997. Momentum investment strategies of mutual funds, performance persistence, and survivorship bias. Unpublished working paper. University of Colorado.

Zheng, L., 1999. Is money smart?: a study of mutual fund investors' fund selection ability. Journal of Finance 54, 901-933.

Table 1

Statistics for two samples of pension fund managers

The mean, standard deviation, minimum, median, and maximum are presented for the quarterly returns of equally weighted portfolios. The returns are for all funds existing each quarter and for each category of surviving and discontinued funds by style. Discontinued funds are defined as those that do not exist in the fourth quarter of 1994. Panel A shows data from Callan Associates, Inc. It consists of the 60 funds (45 survivors and 15 discontinued) in the Callan Associates data that existed between January 1985 and December 1994 and that had at least 20 quarters of returns. Styles in this case are declared by the managers. Panel B shows data from the Frank Russell, from Myers (1999). It consists of the 243 funds in the Frank Russell data that were classified in either growth or value and that existed between January 1985 and December 1994. Styles in this case are defined by the Frank Russell. All is Growth plus Value, and Total is Survivors plus Discontinued.

Fund Group	Mean	Standard Deviation	Minimum.	Median	Maximum	Number of Funds
A: Callan Associates data						
Total						
All	3.69	7.40	-19.66	4.57	19.17	60
Growth	3.79	8.27	-19.64	4.40	21.88	22
Value	3.60	6.92	-20.14	4.36	17.32	29
Survivors						
All	3.64	7.18	-18.77	4.49	18.34	45
Growth	3.67	7.87	-18.27	4.53	20.65	16
Value	3.57	6.85	-19.41	4.28	17.10	23
Discontinued						
All	3.95	8.20	-22.03	5.22	21.11	15
Growth	4.14	9.36	-22.14	4.59	23.94	6
Value	4.03	7.50	-23.06	4.61	18.12	6
B: Frank Russel data						
Total						
All	3.92	8.02	-21.23	4.58	20.80	243
Growth	4.17	9.21	-21.72	4.82	23.94	111
Value	3.70	7.20	-20.79	3.77	17.75	132
Survivors						
All	4.00	8.16	-21.38	4.71	21.04	195
Growth	4.20	9.22	-21.92	4.84	23.78	96
Value	3.78	7.31	-20.80	3.77	17.82	99
Discontinued						
All	4.15	7.79	-20.80	5.31	20.13	48
Growth	4.57	9.79	-20.88	5.49	24.63	15
Value	3.96	7.18	-20.75	4.74	17.60	33

Table 2

Momentum investing measures for the managed portfolios

The LM is the measure of momentum investing in Eq. (15) that pairs the portfolio weight deviations from a buy-and-hold strategy, with quarterly returns from the preceding quarter. The t-statistics in Panels A and B are White-Hansen. Panel A shows the cross-sectional means for LM and t-statistic for the entire sample, for those portfolios labeled as growth and value. Panel B shows the LM for equally weighted portfolios of all funds, growth funds, and value funds. Panel C shows the median LM and the proportion of funds with positive LMs along with a t-statistic and p-value from a binomial test of whether the median is positive (or negative), which is the same as testing whether the underlying proportion is higher (or lower) than 0.50. Panel D shows the Bonferroni p-values for the entire sample and the two categories of growth and value. Panel E shows a t-test for the difference in two means and a Wilcoxon-Mann-Whitney rank test. The null hypothesis is that no difference exists between the LMs for growth managers and those for value managers. EW denotes an equally-weighted portfolio.

Fund Group	LM	t-statistic	p-value
Mean			
All funds	0.12	0.25	
Growth funds	0.58	2.89	
Value funds	-0.21	-1.76	
Equally weighted			
All funds	0.09	1.90	0.06
Growth funds	0.57	6.09	0.00
Value funds	-0.24	-6.19	0.00
Median			
All funds	-0.03		
Growth funds	0.59		
Value funds	-0.31		
Proportion>0			
All funds	0.48	-0.26	0.80
Growth funds	0.86	3.41	0.00
Value funds	0.24	-2.79	0.01
Bonferroni p-value	**Maximum.**	**Minimum.**	
All funds	0.000	0.000	
Growth funds	0.000	0.000	
Value funds	0.000	0.000	
Difference in LM,	**t-statistic**		
EW growth - EW value	7.99		
Wilcoxon-Mann-Whitney,	**z-statistic**		
Value has less momentum	-4.66		

Table 3

Turnover measures for the managed portfolios

Quarterly turnover is derived from the portfolio weight changes, in comparison to a buy-and-hold strategy. Panel A shows the cross-sectional means for the quarterly turnover estimates for the entire sample, for those portfolios labeled as growth and as value. Panel B shows the quarterly turnover for equally weighted portfolios of growth funds and value funds. Panel C shows the median quarterly turnover. Panel D shows a t-test for the difference in two means and a Wilcoxon-Mann-Whitney rank test. The null hypothesis for each is that no difference exists between the quarterly turnover for growth managers and value managers.

Fund Group	Quarterly Turnover
Mean	
Growth funds	17%
Value funds	13%
Equally-weighted	
Growth funds	14%
Value funds	10%
Median	
Growth funds	16%
Value funds	12%
Difference in Mean,	t-statistic
Mean growth - Mean value	2.82
Wilcoxon-Mann-Whitney,	z-statistic
Value has lower turnover	-2.43

Table 4

Returns-based measures of pension fund performance

Excess return is the mean quarterly return in excess of the benchmark index (percent per quarter). In panel A the benchmark is the Center for Research in Security Prices (CRSP) value-weighted NYSE, AMEX, and NASDAQ index; in Panel B it is the style index (growth or value) specified for the fund. The unconditional alpha is Jensen's (1968) measure and the conditional alpha is Ferson and Schadt's (1996) conditional Jensen measure. The conditioning variables are the lagged dividend yield on the Standard and Poor's (S&P) 500, the lagged three-month T-bill yield, a January indicator variable, the lagged difference between the constant maturity ten-year Treasury bond yield and the three-month T-bill yield, the lagged difference between Moody's BAA-rated corporate bond yield and AAA-rated corporate bond yield, and the lagged quarterly and annual return on the benchmark index. All are quarterly returns in percent. The interim trading portfolios start with value weights on 500 randomly selected stocks from the CRSP, NYSE, AMEX, and NASDAQ universe with a price that is greater than or equal to five dollars prior to the beginning of the first quarter. The portfolios then evolve using monthly rebalancing based on conditional expected returns using public information. For the actual funds the table shows the cross-sectional means and t-statistics for the entire sample as well as for those portfolios labeled as growth and value. The table also shows the performance for the equally weighted portfolios of the funds. The median measures and the proportion of funds with positive performance measures include a t-statistic and p-value from a binomial test of whether the proportion positive is higher (or lower) than 0.50. The data cover the 1985-94 period (40 quarterly observations).

Fund Group	Excess return	t-statistic	Unconditional alpha	t-statistic	Conditional alpha	t-statistic
Mean						
Interim-trading	1.00	3.60	1.03	3.64	1.00	3.27
All funds	0.19	0.35	0.29	0.73	0.21	0.58
Growth funds	0.33	0.58	0.25	0.57	0.02	0.10
Value funds	0.08	0.12	0.32	0.81	0.40	0.93
Equally-weighted						
All funds	0.16	1.08	0.32	2.64	0.30	2.55
Growth funds	0.26	0.91	0.22	0.75	0.16	0.59
Value funds	0.06	0.27	0.37	2.21	0.43	2.33
Median						
All funds	0.17		0.32		0.34	
Growth funds	0.34		0.30		0.07	
Value funds	0.04		0.31		0.51	
Proportion of funds with positive performance						
All funds	0.68	2.84	0.78	4.39	0.72	3.36
Growth funds	0.82	2.98	0.73	2.13	0.55	0.43
Value funds	0.59	0.93	0.83	3.53	0.90	4.27
Mean						
Growth funds	0.52	0.84	0.94	1.88	1.00	1.95
Value funds	-0.21	-0.30	0.49	0.70	1.31	2.01
Equally-weighted						
Growth funds	0.52	1.06	0.95	2.78	1.11	3.29
Value funds	-0.21	-0.37	0.45	1.00	1.03	2.59
Median						
Growth funds	0.55		0.98		1.14	
Value funds	-0.23		0.48		1.19	
Proportion of funds with positive performance						
Growth funds	0.86	3.41	1.00	4.69	1.00	4.69
Value funds	0.17	-3.53	0.90	4.27	1.00	5.39

Table 5

Weight-based measures of pension portfolio performance using k = 1 lag

The UWM is the unconditional weight measure. The CWM is the conditional weight measure. Difference is the UWM - CWM. All are quarterly and in percent. The information lag used to create the benchmark weight is one quarter. The t-statistics in Panels A though C are White-Hansen. Panel A shows the cross-sectional means for the performance measures and t-statistics for 100 interim trading portfolios. Each of these portfolios starts with value weights on 500 randomly selected stocks from the Center for Research in Security Prices (CRSP), NYSE, AMEX, and NASDAQ universe with an initial price that is greater than or equal to five dollars per share. The portfolio then evolves using monthly rebalancing based on conditional expected returns using public information. Panel B shows the cross-sectional means for the performance measures and t-statistics for the entire sample of actual funds, as well as for growth and value funds. Panel C shows the performance measures for equally weighted portfolios of funds. Panel D shows the median measure and the proportion of funds with positive performance measures along with a t-statistic and p-value from a binomial test of whether the median is positive (or negative), which is the same as testing whether the underlying proportion is higher (or lower) than 0.50.

Fund Group	UWM	t-statistic	CWM	t-statistic	Difference	t-statistic
Interim-trading						
Mean	0.25	5.13	-0.03	-0.47	0.28	7.34
Mean						
All funds	0.08	0.38	0.04	0.30	0.04	0.14
Growth funds	0.09	0.45	0.01	-0.04	0.08	0.57
Value funds	0.05	0.26	0.05	0.46	0.00	-0.21
Equally-weighted						
All funds	0.06	1.14	0.04	0.78	0.02	0.54
Growth funds	0.05	0.45	-0.03	-0.18	0.07	0.68
Value funds	0.04	0.76	0.06	0.88	-0.03	-0.45
Median						
All funds	0.07		0.03		0.04	
Growth funds	0.06		0.01		0.07	
Value funds	0.06		0.04		-0.01	
Proportion of funds with positive performance						
All funds	0.72	3.36	0.60	1.55	0.67	2.58
Growth funds	0.82	2.98	0.55	0.43	0.91	3.84
Value funds	0.62	1.30	0.62	1.30	0.45	-0.56

Table 6

Weight-based measures of pension portfolio performance using k = 4 lags.

The UWM is the unconditional weight measure. The CWM is the conditional weight measure. Difference is the UWM - CWM. All are quarterly and in percent. The information lag used to create the benchmark weight is four quarters. The t-statistics in Panels A though C are White-Hansen. Panel A shows the cross-sectional means for the performance measures and t-statistics for 100 interim trading portfolios. Each of these portfolios starts with value weights on 500 randomly selected stocks from the Center for Research in Security Prices (CRSP), NYSE, AMEX, and NASDAQ universe with an initial price that is greater than or equal to five dollars per share. The portfolio then evolves using monthly rebalancing based on conditional expected returns using public information. Panel B shows the cross-sectional means for the performance measures and t-statistics for the entire sample of actual funds, as well as for growth and value funds. Panel C shows the performance measures for equally weighted portfolios of funds. Panel D shows the median measure and the proportion of funds with positive performance measures along with a t-statistic and p-value from a binomial test of whether the median is positive (or negative), which is the same as testing whether the underlying proportion is higher (or lower) than 0.50.

Fund Group	UWM	t-statistic	CWM	t-statistic	Difference	t-statistic
Interim-Trading						
Mean	0.44	2.94	0.04	0.36	0.40	3.59
Mean						
All funds	0.17	0.35	0.03	0.12	0.14	0.25
Growth funds	0.18	0.45	0.02	0.14	0.16	0.43
Value funds	0.09	0.15	0.01	0.09	0.08	0.02
Equally-weighted						
All funds	0.12	0.63	0.03	0.13	0.09	0.53
Growth funds	0.24	0.72	0.03	0.06	0.21	0.65
Value funds	0.01	-0.06	-0.02	-0.09	0.01	0.04
Median						
All funds	0.11		0.00		0.09	
Growth funds	0.13		-0.06		0.13	
Value funds	0.10		0.00		0.02	
Proportion of funds with positive performance						
All funds	0.72	3.36	0.48	-0.26	0.67	2.58
Growth funds	0.77	2.56	0.41	-0.85	0.77	2.56
Value funds	0.62	1.30	0.48	-0.19	0.55	0.56

Table 7

Performance measures using superior-information portfolios

Excess return is the mean quarterly return in excess of the Center for Research in Security Prices (CRSP) value-weighted index. The unconditional alpha is Jensen's (1968) measure with the CRSP value-weighted index as the benchmark portfolio. The conditional alpha is the Ferson and Schadt (1996) conditional Jensen measure. The UWM is the unconditional weight measure. The CWM is the conditional weight measure. ρ is the level of superior information. The table shows the cross-sectional means for the performance measures and t-statistics for 100 interim trading portfolios. Each portfolio starts with value weights on 500 randomly selected stocks from the CRSP, NYSE, AMEX, and NASDAQ universe with a price that is greater than or equal to five dollars prior to the beginning of the first quarter. The portfolio then evolves using monthly rebalancing based on the following equation:

$$w_{jt} = w_{jt-1}\left(\frac{1+(r)r_{jt+1}+(1-r)E(r_{jt+1}\mid Z_t)}{1+(r)r_{pt+1}+(1-r)E(r_{pt+1}\mid Z_t)}\right).$$

The data cover the 1985-94 period (40 observations). Relative performance is the difference between the performance of the portfolio in question and the corresponding $\rho = 0$ portfolio.

Value of ρ Performance Measure	Mean Excess Return	Mean Unconditional Alpha	Mean Conditional Alpha	K = 1 Mean UWM	CWM	K = 2 Mean UWM	CWM	K = 4 Mean UWM	CWM
$\rho = 0$	1.00	1.03	1.00	0.25	-0.03	0.40	0.02	0.44	0.04
$\rho = 0.10$									
Relative performance	0.06	0.07	0.06	0.03	0.04	0.02	0.04	0.01	0.03
(t-statistic)	(6.96)	(6.79)	(6.39)	(6.03)	(8.03)	(1.66)	(4.54)	(0.58)	(2.33)
$\rho = 0.20$									
Relative performance	0.12	0.12	0.12	0.07	0.08	0.04	0.07	0.01	0.06
(t-statistic)	(6.38)	(6.39)	(6.27)	(6.02)	(8.00)	(1.55)	(4.45)	(0.52)	(2.24)
$\rho = 0.50$									
Relative performance	0.25	0.27	0.29	0.15	0.19	0.07	0.15	0.01	0.14
(t-statistic)	(4.90)	(5.09)	(5.23)	(6.04)	(7.90)	(1.33)	(4.36)	(0.39)	(2.05)

Table 8

Effects of survival criteria

Panel A shows across fund means of various measures of performance. Discontinued funds are those that left the database before December 1994. The unconditional alpha is Jensen's (1968) measure with the Center for Research in Security Prices (CRSP) value-weighted index as the benchmark portfolio. The conditional alpha is the Ferson and Schadt (1996) conditional Jensen measure. The conditioning variables are the lagged dividend yield on the Standard & Poor's (S&P) 500, the lagged three-month Treasury bill yield, a January indicator variable, the lagged difference between the constant Maturity ten-year Treasury bond yield and the three-month Treasury bill yield, the lagged difference between Moody's BAA-rated corporate bond yield and the AAA-rated corporate bond yield, and the lagged quarterly return on the CRSP value-weighted index. The UWM is unconditional weight measure, and the CWM is the conditional weight measure. The measures are quarterly and in percent. Panel B shows the mean and standard deviation of the quarterly returns of equally weighted portfolios from 1990 to 1994. The returns are for all funds existing in each quarter for each category. Funds with 31 to 40 quarters are all funds with at least 31 quarters over the full sample. Funds with 20 to 30 quarters are all funds with between 20 and 30 quarters over the full sample.

A	Fund Group	Unconditional alpha	Conditional alpha	$K=1$ UWM	$K=1$ CWM	$K=2$ UWM	$K=2$ CWM	$K=4$ UWM	$K=4$ CWM
	Mean								
	All Funds	0.29	0.21	0.08	0.04	0.11	0.05	0.17	0.03
	15 Discontinued Funds	0.33	0.37	0.14	0.09	0.17	0.05	0.28	0.04
	All Growth	0.25	0.02	0.09	0.01	0.14	0.07	0.18	0.02
	6 Discontinued Growth	0.36	0.33	0.09	0.03	0.15	-0.01	0.28	0.04
	All Value	0.32	0.40	0.05	0.05	0.05	0.04	0.09	0.01
	6 Discontinued Value	0.45	0.57	0.16	0.13	0.20	0.09	0.37	0.08

B		All funds Mean	All funds Standard Deviation	Growth funds Mean	Growth funds Standard Deviation	Value funds Mean	Value funds Standard Deviation
Funds with 31 to 40 quarters of data		2.42	5.92	2.58	6.73	2.26	5.57
Funds with 20 to 30 quarters of data		2.62	6.54	2.64	7.61	2.45	5.95

Table 9

Weight-based measures using alternative measures of expected returns

The UWM is the unconditional weight measure. The CWM is the conditional weight measure. Difference is the UWM - CWM. All are quarterly and in percent. The information lag is k = 1. The t-statistics are White-Hansen. Panel A uses the Capital Asset Pricing Model to calculate expected returns. It shows the cross-sectional means for the performance measures and t-statistics for the entire sample of actual funds, as well as for growth and value funds, and the performance measures for equally weighted portfolios of funds. Panel B uses a subset of the conditioning variables (dividend yield and three-month T-bill yield to estimate expected returns. It shows the cross-sectional means for the performance measures and t-statistics for the entire sample of actual funds, as well as for growth and value funds, and the performance measures for equally weighted portfolios of funds.

Fund Group	UWM	t-statistic	CWM	t-statistic	Difference	t-statistic
Mean						
All funds	0.08	0.44	0.07	0.42	0.02	0.13
Growth funds	0.12	0.67	0.09	0.54	0.03	0.34
Value funds	0.04	0.20	0.04	0.24	0.00	-0.09
Equally-weighted						
All funds	0.05	1.46	0.02	0.90	0.03	1.39
Growth funds	0.06	0.73	0.01	0.19	0.04	1.10
Value funds	0.02	0.46	0.01	0.45	0.00	0.20
Mean						
All funds	0.08	0.33	0.03	0.13	0.05	0.23
Growth funds	0.09	0.40	-0.04	-0.13	0.12	0.90
Value funds	0.05	0.21	0.06	0.29	-0.02	-0.29
Equally-weighted						
All funds	0.06	1.14	0.04	0.78	0.02	0.54
Growth funds	0.05	0.45	-0.07	-0.50	0.12	1.41
Value funds	0.04	0.76	0.07	0.72	-0.03	-0.50