Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 363–386
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Benchmarking the performance of recommended allocations to equities, bonds, and cash by international investment houses ☆

Mary M. Bange [a], Kenneth Khang [b], Thomas W. Miller Jr. [c,*]

[a] Houston, TX 77007, United States
[b] College of Business Idaho State University Pocatello, ID 83209, United States
[c] John Cook School of Business, Saint Louis University, 3674 Lindell Boulevard, St. Louis, MO 63108, United States

Accepted 17 July 2007
Available online 17 October 2007

Abstract

We conduct performance tests of the recommended asset allocations made by a panel of international investment houses (the "Houses") from 1982 through 2005. We compare the returns and Sharpe Ratios from the recommended-weight portfolio against those of several benchmark portfolios and to a set of 10,000 returns and Sharpe Ratios from randomly shuffled-weight and shuffled-weight change portfolios. We find that the Houses generally fail to outperform the benchmarks. The shuffled-weight change benchmark exhibits a robust "style-preserving" property in that the average portfolio standard deviation is nearly equal to the portfolio standard deviation from the actual recommended weights.

JEL classification: G12; G14; G23
Keywords: Asset allocation; Portfolio performance measurement; Benchmark portfolio; Conditional performance; International investment houses

1. Introduction

Perhaps the most fundamental investment decision is how best to allocate portfolio dollars among equities, bonds, and cash. However, there is little published research concerning advice on this basic investment decision. By contrast, many studies examine the ability of investors to make superior decisions about market timing and security selection.

☆ We thank Mike Alderson, Chris Anderson, Andrew Ang, Brian Betker, Geert Bekaert, Matthew Billet, Werner De Bondt, Wayne Ferson, Shingo Goto, Jim Granato, Mike Hemler, Robert Hodrick, Alok Kumar, Bob Miller, Carolyn Moore, David Myers, Greg Niehaus, Ľuboš Pástor, Dave Rapach, Rudi Schadt, Sandra Sizer, Jack Strauss, Jonathan Taylor, Zhenyu Wang, Bonnie Wilson, and Guofu Zhou. We are especially grateful to Heber Farnsworth for his many insights and comments concerning this paper. We acknowledge the helpful suggestions from anonymous referees as well as seminar participants at Northern Illinois University, Saint Louis University, the University of Kansas, the University of South Carolina, and Washington State University. In addition, we received valuable comments from participants at the 2004 AFA Conference, the 2003 FMA Conference, and the 2003 Global Investment Conference in Banff, Alberta, Canada. We are grateful to Leigh Riddick, who provided valuable data. We also thank Chullawach Darongkamas and David Shrider for their high-quality research assistance. Of course, we are responsible for any remaining errors.
* Corresponding author. Tel.: +1 314 977 2476.
E-mail address: millertw@slu.edu (T.W. Miller).

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2007.07.005

Blake et al. (1999) underscore the importance of asset allocation when they show that potential gains to investors from asset allocation dwarf potential gains from market timing and security selection.[1]

A motivating question we address in this paper is how best to measure the performance of asset allocation recommendations in the face of differing investment "styles" and their associated ex-ante risks. We introduce two new performance measures that use randomization to generate a benchmark return distribution. In these randomization methods, the recommended portfolio weights, as well as the recommended portfolio weight changes, are repeatedly shuffled to generate two distributions of portfolio returns. We can then compare the actual performance to these distributions and calculate significance levels.

In our study, we examine the long-horizon performance of recommended asset allocation strategies made by a panel of international investment houses between April 1982 and July 2005. These investment houses (the "Houses") provide recommended asset weights to equities, bonds, and cash. The recommendations come from surveys published in the *Financial Report*, a confidential newsletter that was purchased by The *Economist Newspaper, Ltd.*, in 1989. In these surveys, money managers in different countries are asked to provide asset allocation recommendations for a hypothetical investor who has "no existing investments, no overriding currency considerations, and the investment objective of long-term capital growth." These criteria give us a sample that is well-suited to a detailed examination of the performance of asset allocation recommendations over a long horizon.

We look for performance attributable to style by separating the Houses into "equity champions" (Houses that recommend equity holdings that are greater than the average equity holding), "fixed-income champions" (Houses that recommend equity holdings that are less than the average equity holding), and by examining individual House performance. We find that the shuffled-weight change benchmark portfolio exhibits a robust "style-preserving" property. That is, the average portfolio return standard deviation is nearly equal to the portfolio return standard deviation from the actual recommended weights. This property is robust across time periods and return series. It is also robust as to whether the recommendations come from equity champions, fixed-income champions, individual Houses, or an overall consensus.

In our sample, as a group or individually, the Houses do not exhibit much, if any, skill in shifting asset allocations among equities, bonds, and cash. Independent and simultaneous research by Annaert et al. (2005) also concludes that recommendations from these Houses are unable to outperform passive portfolio benchmarks. However, our study differs in significant ways.

First, Annaert et al. (2005) do not incorporate transaction costs. In our study, we account for transaction costs using a method we develop to rebalance a portfolio to target weights when transaction costs differ among asset classes. Annaert et al. (2005) only use unconditional performance tests. In addition to the randomization methods, we measure portfolio performance using unconditional Sharpe Ratios and the Ferson and Khang (2002) conditional weight measure. We find that the Ferson and Khang (2002) conditional weight measure confirms the basic conclusion of this study. That is, the Houses exhibit little, if any, ability to make timely asset allocation shifts among equities, bonds, and cash.

Our findings have important implications for investors who rely on basic asset allocation advice. The scale of their collective decisions is colossal. Mutual fund investments in particular are vast. As of the second quarter of 2006, world-wide mutual fund assets totaled \$19.4 trillion, distributed among 59,385 funds across 41 countries (www.ici.org, November 6, 2006). It is possible that this enormous investment pool has benefited from knowledgeable asset allocation recommendations over time. However, asset allocation recommendations that are highly correlated with those in our sample would have had little value for investors.

The paper proceeds as follows. In Sections 2 and 3, we describe the return and weight data, respectively. In Section 4, we detail how we incorporate transaction costs into portfolio rebalancing. In Section 5, we summarize the unconditional portfolio performance results, and in Section 6, we present two robustness inquiries. Section 7 contains conditional portfolio performance and Section 8 concludes.

[1] For example, from April 1982 through October 2005, an investment of \$1 million would have grown to over \$2.1 billion if, at the beginning of each month, the investor placed their growing portfolio into the asset class (i.e., equities, bonds, or cash) that will perform best over the month (with transaction costs of 50 basis points per month, the portfolio still would have grown to over \$940 million). Three other studies on fundamental asset allocation are those by Brinson et al. (1986, 1991), and Annaert et al. (2005).

2. Data

The data comprise asset allocation recommendations from the Houses, as well as monthly returns for equities, bonds, and cash in various market markets. In the conditional performance tests, we use a traditional set of information variables, which are described in Section 7.2.

Our data have several important features to help test whether these recommendations provide returns in excess of benchmark returns. Because we have a set of portfolio recommendations made at specific dates, performance cannot be clouded by intra-period changes in asset allocations, as shown by Goetzmann et al. (2000). Given recommendation dates, we are assured that the only trading that occurs is to rebalance the portfolio to the newly recommended weights. We can also control for the common information set available at the time of the recommendations, which is necessary for conditional performance tests. Together, these features provide a clean setting to test the performance of recommended portfolio allocations among equities, bonds, and cash over a 23-year sample period.

2.1. Asset allocation recommendations

Twenty-six Houses voluntarily participate in the *Financial Report/Economist* surveys conducted between 1982 and 2005. We note that voluntary participation results in a non-random sample, because Houses that participate are potentially skilled and Houses that do not participate are potentially unskilled. However, if ability is to be found anywhere, it should be in a group of voluntary participants.

Seventeen Houses made 20 or more recommendations, and nine Houses offered ten or fewer recommendations. For all recommendation dates in the sample except the last one, there were at least six Houses polled, with a maximum of ten and an average of 8.3. Fig. 1 presents House participation by recommendation date.[2]

2.2. Returns

Consistent with other international financial market studies, we use the Morgan Stanley Capital International (MSCI) value-weighted index as the proxy for the world market portfolio. We use total returns (capital gains plus dividends) in U.S. dollars. In our robustness tests, we use the S&P 500 Total Return Index for U.S. equity returns, the All-share Index to proxy for U.K. equity returns, and the Nikko Securities total return index to proxy for Japanese equity returns. For bonds, we use the Total Return Bond Indices based on the yields on 10-year Government bonds for the world, the U.S., the U.K., and Japan. The return indexes for equities and bonds, as well as Treasury bill returns for the U.S., the U.K., and Japan, are from www.globalfinancialdata.com. For the world portfolio, we calculate the returns on cash using the one-month dollar LIBOR from *Datastream*.

3. Weight recommendations

3.1. Recommendation dates

For all recommendation dates in the sample, we do not know exactly when each House established its recommendations. Therefore, we must specify a "recommendation date." We specify the recommendation date in this study as the beginning of the month, if the actual publication date of each survey is before the fifteenth of the month. Otherwise, we assume that the recommendation date is the beginning of the subsequent month.[3]

[2] Throughout most of the sample, the Houses report their recommended portfolio allocations among only equities, bonds and cash, i.e., short-term government paper. However, in a few instances the surveys also provide recommended allocations to other assets such as real estate, gold, and fine art. The recommended percentage allocation to these other assets averages less than 1% throughout the sample. Therefore, we add this other allocation to the percentage allocated to cash.

[3] In the early days of the sample, data-gathering technology and lack of e-mail might not have allowed for a quick survey turn-around. Our assumption concerning the specification of the "recommendation date" is not critical. Test results assuming the "recommendation date" is always at the beginning of the subsequent month are similar to the reported test results.

3.2. Subperiods

Fig. 1 presents the survey participation dates for the Houses from April 1982 through October 2005. As Fig. 1 shows, the set of participating Houses generates four "natural" subperiods.

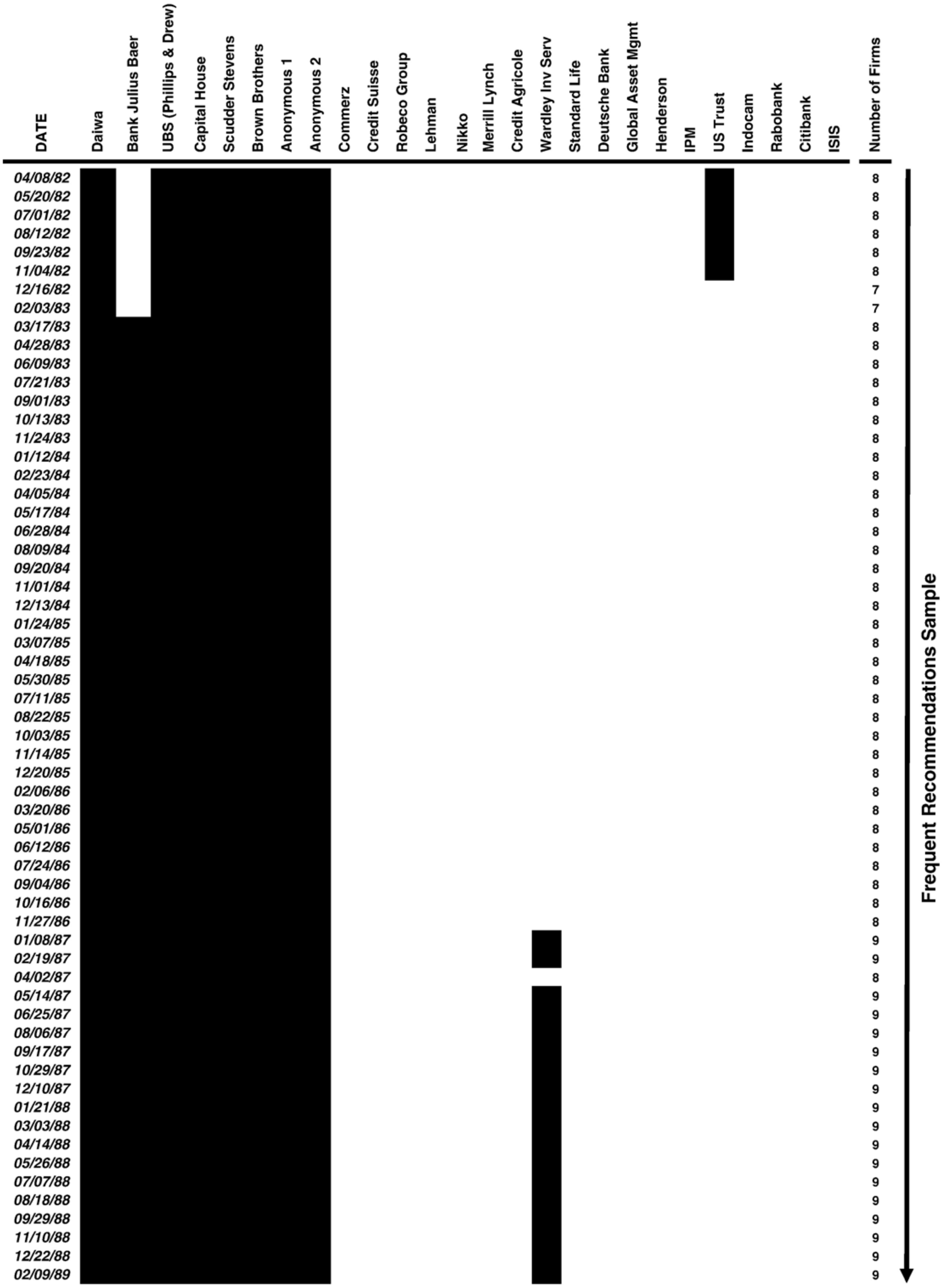

Fig. 1. Investment house participation, by number of recommendations (dark cells indicate "in-samples").

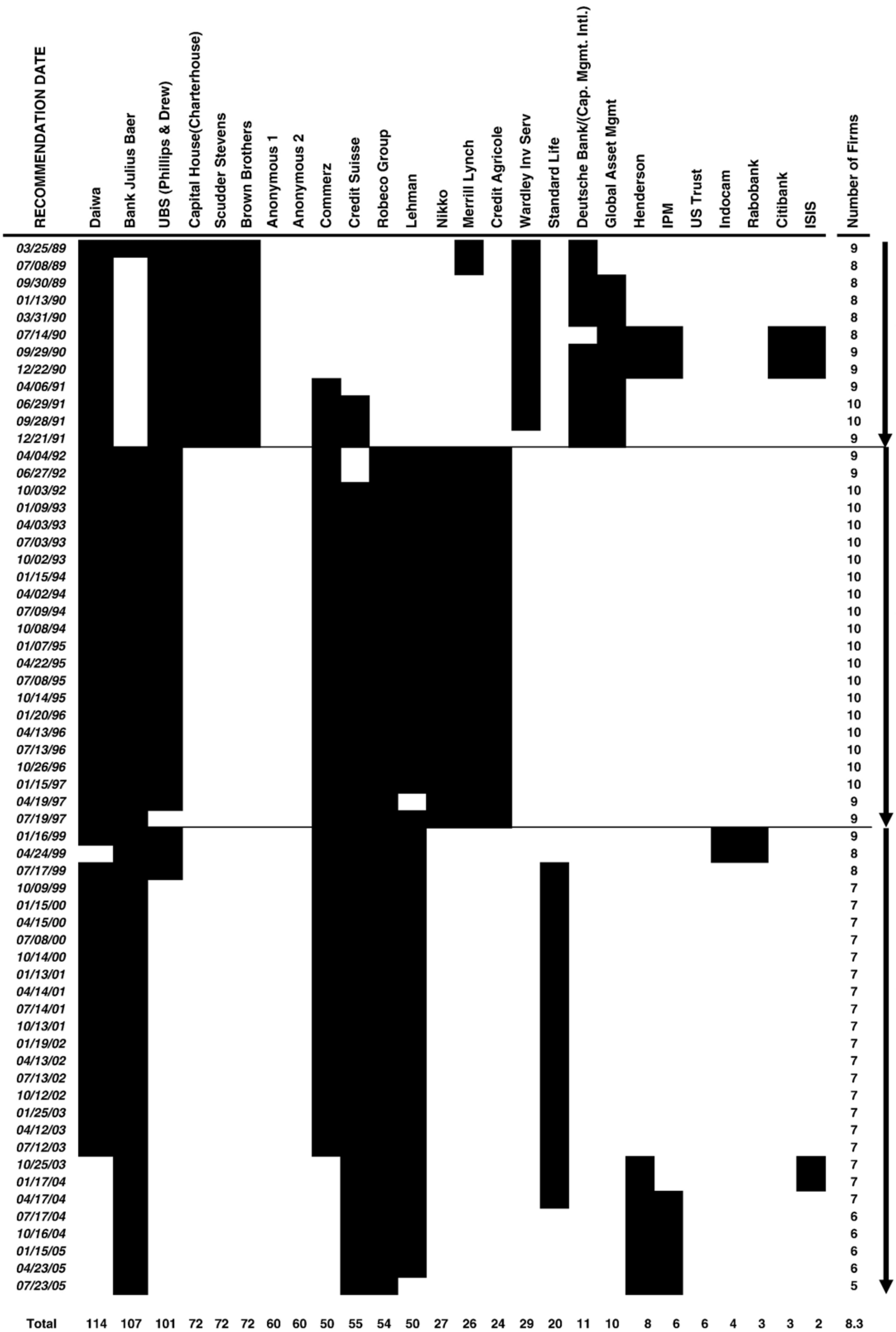

Fig. 1 (*continued*).

Beginning on April 8, 1982 the *Financial Report* published its survey every six weeks. The first subperiod runs from the April 8, 1982 survey through the February 9, 1989 survey. As shown in Fig. 1, this subperiod is characterized by frequent recommendations from essentially the same group of Houses.

Beginning on March 25, 1989, *The Economist* published the survey approximately every three months. In addition, the group of Houses that were surveyed changed significantly. In early 1992, *The Economist* made some additional substantial changes concerning which Houses comprised the survey. Thus, our second subperiod begins with the March 25, 1989 survey and ends with the December 21, 1991 survey, whose weights run through March 1992. In this subperiod, the group of Houses comprising the survey was largely unstable. The third subperiod begins with the April 4, 1992 survey and ends with the October 18, 1997 survey. In this subperiod, the group of Houses comprising the survey was largely stable.

The Economist did not request asset allocation recommendations for equities, bonds, and cash from the December 1997 survey through the October 1998 survey. Therefore, the fourth subperiod begins with the January 16, 1999 survey and ends with the July 23, 2005 survey, whose weight recommendations are used through October 2005.

Because of the interruption in weight recommendations, the longest subperiod we can analyze begins in April 1982 and ends in October 1997. This subperiod is our fifth sample.

3.3. Average House weight recommendations

Except in our robustness inquiries, we investigate the performance of recommended portfolio weights by using the average weight recommendations of the participating Houses at each survey date. Given the fluid nature of the set of firms that comprise the survey, forming an average allows us to construct and study a monthly sample that spans about 23 years.

There are several other reasons for using the average, or consensus, recommendation. Long-time readers of the survey might construct such a weight because many Houses were surveyed throughout the sample period. In addition, any reader of the survey, at any time, could think that an average of the House recommendations is a “reasonable” weighting strategy. Finally, forming a consensus allows us to investigate whether the varied investment styles and strategies buttress or offset one another.

3.4. High-Equity and Low-Equity Houses

It is quite likely that investment strategies varied across Houses in our sample. Using a subset of these Houses, Bange and Miller (2004) find evidence of momentum trading in equities for Houses that often recommend a high percentage allocation to equities. Consequently, we examine the potential for an underlying “equity champion” or “fixed-income champion” recommendation bias.

At each date throughout the sample, we compute the average recommended equity allocation. We classify Houses that recommend equity holdings that are higher than the average equity holding as “equity champions.” We average the asset allocations of these Houses and name the group “High-Equity Houses.” We classify Houses that recommend equity holdings that are less than the average equity holding on that date as “fixed-income” champions. We average the asset allocations of these Houses and name the group “Low-Equity Houses.”

To examine the consistency of the classifications, we note that 17 of the 26 Houses made more than ten recommendations. Throughout the sample, 13 of the 17 Houses that made more than ten recommendations are consistently classified. That is, throughout the sample, only four Houses are ever classified as both a High-Equity House and a Low-Equity House.[4]

3.5. Recommended asset allocations among equities, bonds, and cash

Panel A of Table 1 contains the annualized returns for equities, bonds, and cash for various subperiods. For the April 1982 through October 1997 subperiod, the annualized returns for equities, bonds, and cash are 12.08%, 13.28%, and

[4] These Houses are Capital House, Commerz International, Daiwa Europe, and Merrill Lynch. Nine Houses are consistently classified as a High-Equity House: Anonymous One, Anonymous Two, Brown Brothers Harriman, Credit Agricole, Lehman Brothers, Nikko Securities, Scudder Stevens Clark, Standard Life, and Wardly Investor Services. Four Houses that are generally classified as a Low-Equity House are Bank Julius Baer, Credit Suisse, Robeco Group, and UBS Phillips and Drew.

Table 1
Recommended percentage allocations to equities, bonds, and cash, by sample period

	Sample period:				
	Apr-1982 Oct-1997	Apr-1982 Mar-1989	Apr-1989 Mar-1992	Apr-1992 Oct-1997	Jan-1999 Oct-2005
Panel A. Annualized benchmark returns and number of recommendations					
Annualized Average Returns (in %) for:					
World equity	12.08	19.80	−2.50	11.12	1.43
World gov't bonds	13.28	17.34	9.77	10.22	3.06
1-mo LIBOR	6.97	8.71	7.35	4.62	3.28
Monthly returns	187	84	36	67	82
Recommendations	94	60	12	22	27
Panel B. Recommended portfolio weights, all houses (in %)					
Equity:					
Average	67.2	73.5	63.9	51.8	52.7
Median	72.0	74.5	61.5	52.2	52.9
S.D.	10.6	5.6	7.8	2.5	1.5
Highest	82.7	82.7	75.6	58.9	54.7
Lowest	45.0	56.1	54.7	45.0	49.4
Bonds:					
Average	24.0	18.1	22.7	40.7	40.3
Median	19.4	16.5	19.6	40.6	39.7
S.D.	10.4	4.1	8.1	1.8	2.7
Highest	45.4	29.0	35.3	45.4	45.4
Lowest	12.8	13.3	12.8	36.3	35.1
Cash:					
Average	8.9	8.5	13.4	7.5	7.0
Median	7.8	7.0	12.0	7.3	6.4
S.D.	4.0	3.9	5.1	1.5	2.6
Highest	26.4	17.7	26.4	9.9	13.3
Lowest	2.6	2.6	8.2	4.5	3.3
Panel C. Recommended portfolio weights, high-equity houses (in %)					
Equity:					
Average	79.0	85.3	78.2	61.5	59.5
Median	85.0	86.7	78.6	60.6	59.6
S.D.	11.8	6.3	8.0	5.2	2.4
Highest	96.3	96.3	88.4	72.8	64.3
Lowest	51.3	60.0	67.2	51.3	55.6
Bonds:					
Average	13.6	7.5	9.9	32.5	37.6
Median	7.6	5.7	7.2	33.4	38.2
S.D.	11.9	5.6	7.4	4.0	2.4
Highest	37.0	28.3	23.5	37.0	42.0
Lowest	1.3	1.3	1.7	22.8	32.2
Cash:					
Average	7.4	7.0	12.0	6.1	2.7
Median	6.2	6.0	11.5	5.4	1.8
S.D.	3.7	3.1	3.7	3.6	2.2
Highest	17.8	15.3	17.8	15.0	9.2
Lowest	0.0	1.5	7.6	0.0	0.0
Panel D. Recommended portfolio weights, low-equity houses (in %)					
Equity:					
Average	50.9	54.2	51.1	41.7	43.6
Median	52.5	55.0	47.2	41.4	44.0
S.D.	8.0	6.4	8.8	3.0	3.8
Highest	67.5	67.5	64.6	47.8	50.0
Lowest	32.0	32.0	39.6	36.0	34.0

(continued on next page)

Table 1 (*continued*)

	Sample period:				
	Apr-1982	Apr-1982	Apr-1989	Apr-1992	Jan-1999
	Oct-1997	Mar-1989	Mar-1992	Oct-1997	Oct-2005
Panel D. Recommended portfolio weights, low-equity houses (in %)					
Bonds:					
Average	38.2	34.9	34.2	49.2	43.6
Median	38.3	34.7	32.4	48.7	42.5
S.D.	8.5	5.7	9.8	3.3	4.8
Highest	55.8	45.0	50.4	55.8	54.3
Lowest	19.0	24.5	19.0	45.0	34.0
Cash:					
Average	10.9	10.9	14.8	9.1	12.8
Median	9.9	9.3	13.0	9.3	13.3
S.D.	5.5	5.7	7.2	2.3	7.0
Highest	33.4	26.7	33.4	13.8	27.0
Lowest	1.7	1.7	6.2	4.0	2.0

Panel A contains the annualized average monthly return by asset category, the number of monthly returns in the sample period, and the number of asset allocation recommendations. Panels B, C, and D contain the grand average, highest average, and lowest average percentage asset allocation using data from all Houses, “High-Equity” Houses, and “Low-Equity” Houses, respectively. For each recommendation date, we compute the average equity holding recommendation across all Houses. We classify those Houses that recommend an equity holding greater than (less than) the average as “High-Equity” (“Low-Equity”) Houses.

6.97%, respectively. All three asset classes have their highest returns in the April 1982 through March 1989 sample period.

There is considerable variation among the periods with respect to which asset class has the highest return.In the April 1989 through March 1992 period, the returns on bonds (9.77%) and the returns on cash (7.35%) exceed the returns on equities (−2.50%). In the April 1992 to October 1997 period, the returns on equities (11.12%) and bonds (10.22%) exceed the returns on cash (4.62%). In the period January 1999 through October 2005, the returns on cash (3.28%) exceed bond (3.06%) and equities (1.43%) returns.

Panel B of Table 1 contains the average, median, and extreme recommended allocations to equities, bonds, and cash. On average across the April 1982 through October 1997 sample period, the Houses recommend that investors allocate 67.2% of their portfolios to equities; 24.0% to bonds; and 8.9% to cash.[5]

There is also considerable variation in the average recommended weights among periods. Panel B of Table 1 shows that the average equities allocation falls from 73.5% in the April 1982 through March 1989 period to 51.8% for the April 1992 to October 1997 period. There is a corresponding increase in bond allocation from 18.1% to 40.7%.

Fig. 2 illustrates the time variation in average asset allocations. At the beginning of the sample, the average percent equities, bonds, and cash allocations were 59.6%, 28.5%, and 11.9%, respectively. The allocations to equities (bonds) quickly rose (fell) and remained above 70% (below 20%) for much of the 1980s. In 1989, the Houses began to shift out of equities and into bonds. The allocation to equities decreases to the 50% range, and the allocation to bonds increases to the 40% range.[6]

[5] The lowest average equities recommendation is 45% in October 1992. By contrast, the highest average equities allocation, 82.7%, is in January 1984. In this survey, there are two Houses that recommend a 100% equities weight, and three other Houses that recommend a weight of at least 90%. At every survey date, each House recommends a weight of more than 20% in equities. Average recommended bond weights range from 12.8% to 45.4%. At no time does any House recommend more than a 70% bond weight. At least once during the sample period, six Houses recommend that bonds should receive zero weight.

[6] Given the changes in the set of Houses in the April 1989 to March 1992 sample period, it is unclear whether the increase (decrease) in equities (bonds) after 1989 stems from changes in the allocations or from changes in the set of Houses.

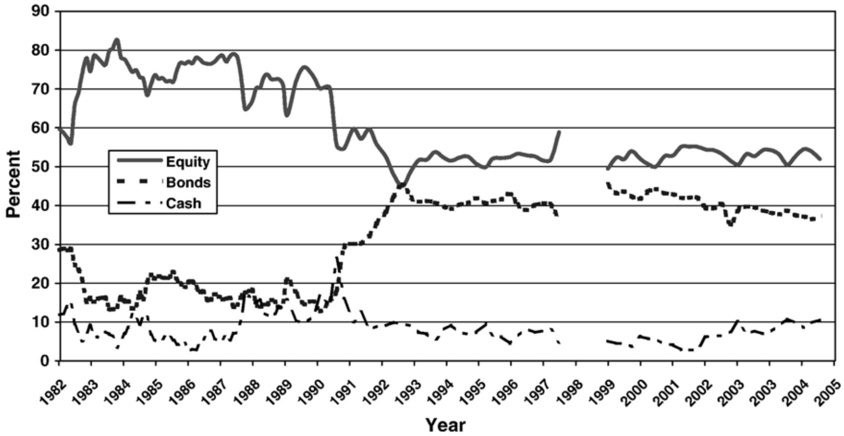

Fig. 2. Mean allocation to equity, bonds, and cash for the 24 investment houses in the sample, 4/82 through 10/97 and 1/99 through 10/05.

In July 1997, *The Economist* suspended their request for recommended allocations to equities, bonds, and cash. At that time, the average percent equities, bonds, and cash recommendations were 58.9%, 36.3%, and 4.8%, respectively. When *The Economist* resumed the request in January 1999, the average percent equities, bonds, and cash recommendations were 49.5%, 45.4%, and 5.1%. In the last recommendation in the sample (July 2005), the average equities, bonds, and cash allocations were 52.0%, 37.4%, and 10.6%.

Panels C and D of Table 1 contain the respective average portfolio weights for the High-Equity and Low-Equity Houses. High-Equity Houses and Low-Equity Houses recommended the highest allocation to equity in the April 1982 through March 1989 sample period, 85.3% and 54.2%, respectively. Major declines in their recommended portfolio weights to equity occur in subsequent sample periods.

The average equities allocation for High-Equity Houses falls to 59.5% by the January 1999 to October 2005 sample period. In this period, the average portfolio weight for bonds increases to 37.6%. Recommendations for the Low-Equity Houses show a similar pattern through October 1997. Portfolio weights for equities decrease to 41.7%, and weights to bonds increase to 49.2%. However, the shift from equities to bonds reverses slightly in the January 1999 to October 2005 sample period.

4. Transaction costs and portfolio rebalancing

We assume that investors rebalance their portfolios at each recommendation date. To rebalance, investors must make trades in the recommended portfolios (and some benchmark portfolios). Thus, we include transaction costs so that we can accurately compare recommended portfolio performance to the benchmark portfolio performance.

Consistent with Pesaran and Timmermann (1995) and Liu et al. (2003), we assume transaction costs are proportionate to the value of the trades necessary to rebalance the portfolio and symmetric for purchases and sales. However, unlike these studies, we do not assume that transaction costs are constant throughout the sample.

For equity trades, we assume one-way transaction costs in 1982 are 75 basis points, based on Jones (2002). These transaction costs reflect brokerage commissions and half the bid–ask spread. Thereafter, we assume that transaction costs throughout the sample follow the path documented by Jones (2002). With the exception of a five-basis-point increase in 1990, transaction costs fall from two to five basis points per year. By 1999, one-way transaction costs are 20 basis points. We assume this level holds to the end our of sample period.

Schultz (2001) finds that between 1995 and 1997, the round-trip cost for trading corporate bonds is about 27 basis points. Chakravarty and Sarkar (1999) report that during this same period, trading costs decline and that the costs of trading government bonds are lower than the trading costs for corporate bonds. Liu et al. (2003) assume a one-way trading cost for bonds of ten basis points from 1991 to 2002.

To capture the declining nature of transaction costs throughout our sample period, and the relatively lower trading costs of government bonds, we assume one-way transaction costs for bonds of 20 basis points for 1982 to 1991, 15 basis points from 1992 to 1995, and ten basis points from 1996 to 2005. We assume there are no transaction costs for changing cash balances. We also assume that the investors who follow the asset allocation recommendations in the surveys do not use futures contracts to make asset allocation shifts.

Consider a beginning of period arbitrary investment of \$1 in a portfolio comprised of j assets. At the end of the period, the investor will rebalance the portfolio, and after transaction costs are paid, we assume the proportionate holdings in the portfolio reflect the new weight recommendations. To accomplish this, first define

ω_j^I	start-of-period, or initial, weight in asset j
ω_j^T	end-of-period, or target, weight in asset j, after transaction costs
r_j	periodic return on asset j
r_p	periodic portfolio return before transaction costs
τ_j	percent transaction cost applied to a \$1 dollar transaction in asset j
Δ_j	dollar transaction in asset j to achieve rebalancing.

To account for the buying and selling of assets when an investor rebalances a portfolio, we define a "buy–sell" indicator for asset j, ι_j. This indicator has a value of either +1 or −1. The indicator is positive when the investor must buy more of asset j for the portfolio. That is, when

$$\omega_j^T > \omega_j^I \left(\frac{1+r_j}{1+r_p} \right). \tag{1}$$

Consequently, the following condition must hold for each asset j:

$$\omega_j^T \left[1 + r_p - \sum_{j=1}^{N} \tau_j \Delta_j \right] = \omega_j^I (1 + r_j) + \iota_j \Delta_j - \tau_j \Delta_j. \tag{2}$$

The bracketed term on the left-hand side in Eq. (2) represents the total dollar value of the portfolio after transaction costs are paid. Thus, the left-hand side of Eq. (2) is the dollar value in asset j after rebalancing (net of transaction costs). The right-hand side of Eq. (2) is the sum of three terms. These terms are: 1) the end-of-month buy-and-hold value of the dollar investment in asset j, plus; 2) the dollar adjustment in asset j needed to achieve the target weight in asset j, minus; 3) the dollar transaction costs of making this adjustment. Therefore, the right-hand side of Eq. (2) also equals the dollar value in asset j after rebalancing (net of transaction costs).

The rebalancing solution comes from simultaneously solving for all Δ_j. Rearranging Eq. (2), the following equation holds for each asset j:

$$\omega_j^T (1 - r_p) - \omega_j^I (1 + r_j) = \omega_j^T \left[\sum_{j=1}^{N} \tau_j \Delta_j \right] + \iota_j \Delta_j - \tau_j \Delta_j. \tag{3}$$

If the portfolio consists of three assets, solving the following system of equations results in the dollar transaction, Δ_j, in each asset needed to achieve target weights.

$$\begin{bmatrix} \omega_1^T (1 + r_p) - \omega_1^I (1 + r_1) \\ \omega_2^T (1 + r_p) - \omega_2^I (1 + r_2) \\ \omega_3^T (1 + r_p) - \omega_3^I (1 + r_3) \end{bmatrix} = \begin{bmatrix} (\omega_1^T \tau_1 + \iota_1 - \tau_1) & \omega_1^T \tau_2 & \omega_1^T \tau_3 \\ \omega_2^T \tau_1 & (\omega_2^T \tau_2 + \iota_2 - \tau_2) & \omega_2^T \tau_3 \\ \omega_3^T \tau_1 & \omega_3^T \tau_2 & (\omega_3^T \tau_3 + \iota_3 - \tau_3) \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix} \tag{4}$$

In matrix notation, this system of equations can be written as

$$D = M\Delta. \tag{5}$$

The solution for the dollar transaction in each asset required to achieve its target weight is

$$M^{-1}D = \Delta. \tag{6}$$

Given a time series of initial weights, target weights, asset returns, and transaction cost levels, we can solve for Δ at the end of each period. With this method, we can calculate the end-of-period portfolio values before and after transaction costs, and also the gross and net portfolio returns. This method is useful for accounting for transaction costs when a portfolio is rebalanced to reflect target portfolio weights.

5. Portfolio performance and unconditional measures

In this study, we can easily impose a no-trading constraint between recommendation dates. Given the time series of recommended weights and asset returns, we can measure portfolio returns generated by actual weight recommendations against portfolio returns generated by benchmark weights without the complications introduced by trades between recommendation dates.

5.1. The benchmark portfolios

5.1.1. Passive benchmarks

We specify two passive benchmark portfolios. The first is a common "Robot Blend" which distributes assets 60% to equities, 30% to bonds, and 10% to cash. Arshanapalli et al. (2001) found that their sample of broker recommendations from June 1989 to March 1999 had a average recommendation of 57.57% to equities, 32.12% to bonds, and 10.21% to cash. Brinson et al. (1991) found that their sample of pension fund allocations for the period December 1977 to December 1987 averaged 59.6% to equities, 26.9% to bonds, and 13.6% to cash. Our robot blend represents a rounded version of these averages. We assume that this benchmark portfolio is rebalanced monthly to the robot weights.

The weights for the second passive benchmark portfolio begin with 60% to equity, 30% to bonds, and 10% to cash. However, in this benchmark portfolio, we assume that there is no monthly rebalancing to the initial weight levels. Therefore, these resulting portfolio returns represent a simple "buy-and-hold" strategy. We note that these weights are dynamic because relative weights will fluctuate, or drift, over time as asset returns vary.

5.1.2. Randomized benchmarks

To calculate the portfolio return from the recommended weights over a subperiod of length N months, the first step is to multiply the matrix of recommended weights and the return matrix. For example, if the allocations are to equities, bonds, and cash, the first step is

$$\mathrm{A} = \begin{bmatrix} \omega_{1,e}R_{1,e} & \omega_{1,b}R_{1,b} & \omega_{1,c}R_{1,c} \\ \omega_{2,e}R_{2,e} & \omega_{2,b}R_{2,b} & \omega_{2,c}R_{2,c} \\ \omega_{3,e}R_{3,e} & \omega_{3,b}R_{3,b} & \omega_{3,c}R_{3,c} \\ \vdots & \vdots & \vdots \\ \omega_{N,e}R_{n,e} & \omega_{N,b}R_{n,b} & \omega_{N,c}R_{n,c} \end{bmatrix} = \begin{bmatrix} \omega_{1,e} & \omega_{1,b} & \omega_{1,c} \\ \omega_{2,e} & \omega_{2,b} & \omega_{2,c} \\ \omega_{3,e} & \omega_{3,b} & \omega_{3,c} \\ \vdots & \vdots & \vdots \\ \omega_{N,e} & \omega_{N,b} & \omega_{N,c} \end{bmatrix} \# \begin{bmatrix} R_{1,e} & R_{1,b} & R_{1,c} \\ R_{2,e} & R_{2,b} & R_{2,c} \\ R_{3,e} & R_{3,b} & R_{3,c} \\ \vdots & \vdots & \vdots \\ R_{N,e} & R_{N,b} & R_{N,c} \end{bmatrix} \tag{7}$$

where # represents elementwise multiplication. Then, we calculate monthly portfolio returns, R_i, by summing across the rows in Matrix A above. For all portfolios, we calculate the annualized geometric average monthly return over the subperiod as

$$R_{\text{portfolio}} = \left[\prod_{i=1}^{n} (1 + R_i)\right]^{12/n} - 1 \tag{8}$$

We call the first randomized benchmark portfolio the "Shuffled-Weight" portfolio. To create one shuffled-weight recommendation matrix, we randomly reorder the matrix containing the ordered time series of recommended weights to form a new matrix that contains a new time series of "shuffled" weights. We only shuffle the weights across time, not against time and asset class. That is, we maintain the recommendations to equities, bonds, and cash in each row. Of note is that we only shuffle, or re-assign, weights to those months where there was a new recommended weight. In the months in-between recommendation months, the weights fluctuate, or drift, with relative asset returns.[7]

As an example, suppose $N=12$ and there are quarterly weight updates that occur in months 1, 4, 7, and 10. We begin to calculate the shuffled-weight portfolio return by multiplying the matrix of shuffled weights and the return matrix. That is,

$$B=\begin{bmatrix}\omega_{4,e}R_{1,e} & \omega_{4,b}R_{1,b} & \omega_{4,c}R_{1,c}\\ \omega^*_{4,e}R_{2,e} & \omega^*_{4,b}R_{2,b} & \omega^*_{4,c}R_{2,c}\\ \omega^*_{4,e}R_{3,e} & \omega^*_{4,b}R_{3,b} & \omega^*_{4,c}R_{3,c}\\ \omega_{7,e}R_{4,e} & \omega_{7,b}R_{4,b} & \omega_{7,c}R_{4,c}\\ \omega^*_{7,e}R_{5,e} & \omega^*_{7,b}R_{5,b} & \omega^*_{7,c}R_{5,c}\\ \omega^*_{7,e}R_{6,e} & \omega^*_{7,b}R_{6,b} & \omega^*_{7,c}R_{6,c}\\ \omega_{10,e}R_{7,e} & \omega_{10,b}R_{7,b} & \omega_{10,c}R_{7,c}\\ \omega^*_{10,e}R_{8,e} & \omega^*_{10,b}R_{8,b} & \omega^*_{10,c}R_{8,c}\\ \omega^*_{10,e}R_{9,e} & \omega^*_{10,b}R_{9,b} & \omega^*_{10,c}R_{9,c}\\ \omega_{1,e}R_{10,e} & \omega_{1,b}R_{10,b} & \omega_{1,c}R_{10,c}\\ \omega^*_{1,e}R_{11,e} & \omega^*_{1,b}R_{11,b} & \omega^*_{1,c}R_{11,c}\\ \omega^*_{1,e}R_{12,e} & \omega^*_{1,b}R_{12,b} & \omega^*_{1,c}R_{12,c}\end{bmatrix}=\begin{bmatrix}\omega_{4,e} & \omega_{4,b} & \omega_{4,c}\\ \omega^*_{4,e} & \omega^*_{4,b} & \omega^*_{4,c}\\ \omega^*_{4,e} & \omega^*_{4,b} & \omega^*_{4,c}\\ \omega_{7,e} & \omega_{7,b} & \omega_{7,c}\\ \omega^*_{7,e} & \omega^*_{7,b} & \omega^*_{7,c}\\ \omega^*_{7,e} & \omega^*_{7,b} & \omega^*_{7,c}\\ \omega_{10,e} & \omega_{10,b} & \omega_{10,c}\\ \omega^*_{10,e} & \omega^*_{10,b} & \omega^*_{10,c}\\ \omega^*_{10,e} & \omega^*_{10,b} & \omega^*_{10,c}\\ \omega_{1,e} & \omega_{1,b} & \omega_{1,c}\\ \omega^*_{1,e} & \omega^*_{1,b} & \omega^*_{1,c}\\ \omega^*_{1,e} & \omega^*_{1,b} & \omega^*_{1,c}\end{bmatrix}\#\begin{bmatrix}R_{1,e} & R_{1,b} & R_{1,c}\\ R_{2,e} & R_{2,b} & R_{2,c}\\ R_{3,e} & R_{3,b} & R_{3,c}\\ R_{4,e} & R_{4,b} & R_{4,c}\\ R_{5,e} & R_{5,b} & R_{5,c}\\ R_{6,e} & R_{6,b} & R_{6,c}\\ R_{7,e} & R_{7,b} & R_{7,c}\\ R_{8,e} & R_{8,b} & R_{8,c}\\ R_{9,e} & R_{9,b} & R_{9,c}\\ R_{10,e} & R_{10,b} & R_{10,c}\\ R_{11,e} & R_{11,b} & R_{11,c}\\ R_{12,e} & R_{12,b} & R_{12,c}\end{bmatrix} \quad (9)$$

where # represents elementwise multiplication, and ω^* represents a weight that drifts with relative monthly asset returns. Then, we calculate the shuffled monthly portfolio returns, R_i, by summing across the rows in Matrix B above and calculate an annualized geometric average return. We repeat the process of shuffling the weights and calculating a shuffled-weight annualized geometric average portfolio return, portfolio standard deviation, and Sharpe Ratio 10,000 times.[8]

We call the second randomized benchmark portfolio the "Shuffled-Weight Changes" portfolio. The process of creating one shuffled-weight change recommendation matrix is exactly like the Shuffled-Weight method with one difference. The essential difference is that we randomly reorder a matrix containing the ordered time series of recommended weight changes.

As an example, suppose $N=12$ and there are quarterly weight updates that occur in months 1, 4, 7, and 10. We create a vector containing the changes in weight recommendations. In months where there is a recommended weight change, we calculate the weight recommendation change as $\delta_{t,s}=\omega_{t,s}-\omega_{t-i,s}$, where t represents the month of the new weight recommendation, $t-i$ represents the month of the previous weight recommendation (made i months ago), and s represents the security. We assume that $\delta_{1,s}=0$ for all securities. In months without a weight change recommendation, we allow the weights to drift with the relative asset returns during that month. We shuffle the rows of the vector containing the changes in weight recommendations and, in

[7] For clarification, each subperiod starts at the recommended allocation at the beginning of that subperiod. Also, the shuffled portfolios only shuffle the weights in the subperiod considered, and not of the entire sample.

[8] Instead of shuffling the weights across rows, we can generate rows of random positive weights that sum to one for all rows in the weight matrix. Our preliminary investigations into the performance of these "monkey weights" suggest that they suffer greatly from paying high transaction costs. The high transaction costs are a result of the high variation in weights that occurs when weights are chosen randomly from month to month. We note that there is considerable inertia in the actual weights recommended by the Houses. By contrast, the monkey weights do not necessarily exhibit inertia. For example, equity allocations with random weights could be 10% in month one, 90% in month two, and 30% in month three. These large reallocations generate considerable transaction costs.

this example, add these four shuffled changes to the original recommended weights. The resulting shuffled weight change matrix is,

$$C = \begin{bmatrix} \omega_{1,e}+\delta_{7,e} & \omega_{1,b}+\delta_{7,b} & \omega_{1,c}+\delta_{7,c} \\ (\omega_{1,e}+\delta_{7,e})^{*} & (\omega_{1,b}+\delta_{7,b})^{*} & (\omega_{1,c}+\delta_{7,c})^{*} \\ (\omega_{1,e}+\delta_{7,e})^{*} & (\omega_{1,b}+\delta_{7,b})^{*} & (\omega_{1,c}+\delta_{7,c})^{*} \\ \omega_{4,e}+\delta_{10,e} & \omega_{4,b}+\delta_{10,b} & \omega_{4,c}+\delta_{10,c} \\ (\omega_{4,e}+\delta_{10,e})^{*} & (\omega_{4,b}+\delta_{10,b})^{*} & (\omega_{4,c}+\delta_{10,c})^{*} \\ (\omega_{4,e}+\delta_{10,e})^{*} & (\omega_{4,b}+\delta_{10,b})^{*} & (\omega_{4,c}+\delta_{10,c})^{*} \\ \omega_{7,e}+\delta_{4,e} & \omega_{7,b}+\delta_{4,b} & \omega_{7,c}+\delta_{4,c} \\ (\omega_{7,e}+\delta_{4,e})^{*} & (\omega_{7,b}+\delta_{4,b})^{*} & (\omega_{7,c}+\delta_{4,c})^{*} \\ (\omega_{7,e}+\delta_{4,e})^{*} & (\omega_{7,b}+\delta_{4,b})^{*} & (\omega_{7,c}+\delta_{4,c})^{*} \\ \omega_{10,e}+\delta_{1,e} & \omega_{10,b}+\delta_{1,e} & \omega_{10,c}+\delta_{1,c} \\ (\omega_{10,e}+\delta_{1,e})^{*} & (\omega_{10,b}+\delta_{1,e})^{*} & (\omega_{10,c}+\delta_{1,c}) \\ (\omega_{10,e}+\delta_{1,e})^{*} & (\omega_{10,b}+\delta_{1,e})^{*} & (\omega_{10,c}+\delta_{1,c}) \end{bmatrix} \tag{10}$$

where $(\omega+\delta)^{*}$ represents a weight that drifts with relative monthly asset returns.[9]

We repeat the process of shuffling the weights and weight changes and calculating a shuffled-weight and shuffled-weight change geometric average portfolio return, portfolio standard deviation, and Sharpe Ratio 10,000 times. We can then compare the portfolio return using the actual recommended weights to the distribution of the returns generated by the shuffled-weight portfolios and the shuffled-weight change portfolios. Using these benchmark portfolios allows a direct test of whether there is information in the order, or sequence, of the recommended weight or in the sequence of recommended weight changes. From Noreen (1989), the significance level, α, of such a test is

$$\alpha = \frac{\mathrm{NGE}+1}{\mathrm{NS}+1} \tag{11}$$

where NGE is the number of times the return from the shuffled portfolio equals or exceeds the return from the recommended weight portfolio, and NS is the number of shuffles. We report the number of times the return from the recommended-weight portfolio exceeds the return from the shuffled portfolio, or NLT. To calculate the significance level, we note that $\mathrm{NGE}=\mathrm{NS}-\mathrm{NLT}$.

5.1.3. Performance measures

Many portfolio managers often refer to their management style. Style could refer to restricting investment into subsets of securities or to the zeal exhibited by the manager in shifting asset weights. Style also could refer to how the manager sets average weights (i.e., high versus low equity) or to how the manager allows trends to appear in the weights over time (i.e., gradually shifting from high equity to low equity). Management style could make it difficult to choose an appropriate benchmark portfolio because a pre-specified benchmark might not share the same management style as the portfolio manager.

Consider the Buy-and-Hold portfolio style or the Robot weight style. Both of these styles are styles that an investor could have chosen. But, suppose investors chose to follow the advice of Houses that clearly have a different investment style from the style of these two benchmarks. In this case, the Buy-and-Hold portfolio and the Robot weight portfolio are probably best viewed as foregone opportunity benchmarks.

A strong candidate for a benchmark portfolio should match the style of the portfolio manager. That is, it should match 1) the types of assets or securities held; 2) the magnitude of asset allocation shifts; 3) trends in asset allocations over time, and; 4) "normal," or average, allocations over the measurement period. At the least, a benchmark portfolio should share the same risk profile as the managed portfolio. That way, if the managed portfolio return exceeds the benchmark portfolio return, the managers performance could be judged superior because the two portfolios have the same risk.

[9] We do not constrain the weights to be positive. However, in our sample, $(\omega+\delta)$ is always positive.

Table 2
Annualized return performance of the mean recommended weights for equities, bonds, and cash, by time period

Portfolio consists of: world equity, world gov't bonds, and 1-mo. LIBOR	Sample period:				
	Apr-1982 – Oct-1997	Apr-1982 – Mar-1989	Apr-1989 – Mar-1992	Apr-1992 – Oct-1997	Jan-1999 – Oct-2005
Panel A. Shared benchmark returns:					
1-mo LIBOR	6.97%	8.71%	7.35%	4.62%	3.28%
Monthly returns	187	84	36	67	82
Robot weights (60%/30%/10%)	12.16%	18.20%	2.25%	10.31%	2.35%
S.D.	9.88%	10.13%	12.76%	7.26%	9.18%
Sharpe Ratio, SR_{rob}	0.53	0.94	−0.40	0.78	−0.10
Buy-and-hold (initial: 60%/30%/10%)	12.10%	18.20%	2.49%	10.27%	2.11%
S.D.	10.32%	10.70%	12.34%	7.43%	8.80%
Sharpe Ratio, SR_{bah}	0.50	0.89	−0.39	0.76	−0.13
Panel B. Returns for all houses using:					
Actual weights	12.32%	18.96%	1.48%	10.37%	2.83%
S.D.	10.65%	11.64%	13.14%	7.00%	8.38%
Sharpe Ratio, SR_{act}	0.50	0.88	−0.45	0.82	−0.05
Shuffled weights (10,000 trials)	11.76%	18.37%	2.09%	10.53%	2.98%
Avg. S.D.	10.55%	11.33%	12.85%	6.86%	8.34%
Avg. Sharpe Ratio, SR_{shu}	0.45	0.85	−0.41	0.86	−0.04
Times actual weights outperform:	8897	9324	2499	731	1240
Times SR_{act} exceeds SR_{shu}:	8333	7871	2793	180	1270
Shuffled weight changes (10,000 trials)	12.24%	18.84%	1.45%	10.37%	2.82%
Avg. S.D.	10.65%	11.66%	13.14%	7.02%	8.39%
Avg. Sharpe Ratio, SR_{chg}	0.49	0.87	−0.45	0.82	−0.06
Times actual weights outperform:	7430	7502	5485	5249	5477
Times SR_{act} Exceeds SR_{chg}:	7240	7311	5419	5874	5456
Panel C. Returns for "High-Equity" houses using:					
Actual weights	12.09%	19.13%	−0.06%	10.46%	2.98%
S.D.	11.81%	12.85%	14.75%	7.72%	9.21%
Sharpe Ratio, SR_{act}	0.43	0.81	−0.50	0.76	−0.03
Shuffled weights (10,000 trials)	11.53%	18.59%	0.24%	10.61%	2.83%
Avg. S.D.	11.72%	12.62%	14.59%	7.51%	9.23%
Avg. Sharpe Ratio, SR_{shu}	0.39	0.78	−0.49	0.80	−0.05
Times actual weights outperform:	8656	8914	3837	2532	8026
SR_{act} exceeds SR_{shu}:	8116	8013	3956	1259	8028
Shuffled weight changes (10,000 trials)	11.96%	18.89%	0.03%	10.43%	2.96%
Avg. S.D.	11.83%	12.90%	14.56%	7.76%	9.23%
Avg. Sharpe Ratio, SR_{chg}	0.42	0.79	−0.50	0.75	−0.03
Times actual weights outperform:	7428	7469	4367	5620	5321
Times SR_{act} exceeds SR_{chg}:	7374	7516	5094	6156	5314
Panel D. Returns for "Low-Equity" houses using:					
Actual weights	12.32%	18.40%	2.81%	10.26%	2.62%
S.D.	9.26%	9.82%	11.75%	6.37%	7.31%
Sharpe Ratio, SR_{act}	0.58	0.99	−0.39	0.88	−0.09
Shuffled weights (10,000 trials)	12.02%	17.70%	3.64%	10.33%	3.11%
Avg. S.D.	9.14%	9.67%	11.38%	6.28%	7.23%
Avg. Sharpe Ratio, SR_{shu}	0.55	0.93	−0.33	0.91	−0.02
Times actual weights outperform:	8051	9329	2093	3291	397
SR_{act} exceeds SR_{shu}:	7200	8533	2353	1910	422
Shuffled weight changes (10,000 trials)	12.19%	18.19%	2.67%	10.24%	2.58%
Avg. S.D.	9.30%	9.85%	11.92%	6.40%	7.34%
Avg. Sharpe Ratio, SR_{chg}	0.56	0.96	−0.39	0.88	−0.10
Times actual weights outperform:	7269	7157	5989	5647	5577
Times SR_{act} exceeds SR_{chg}:	7338	6983	5728	6153	5525

A potential weakness of a Sharpe ratio lies in an issue concerning ex-ante versus ex-post risk. Two investment strategies could have distinctively different ex-ante risks, but could be judged as having similar risk ex-post simply by chance. For example, a hedge fund that using a strategy that purports to eliminate risk could earn a return greater than a T-bill that has the same ex-post total risk. The Sharpe Ratio for the hedge fund would be higher, indicating that the hedge fund was a superior investment compared to the T-bill. However, it is possible that the hedge fund strategy might not work effectively-resulting in a much higher ex-post risk and a lower Sharpe Ratio than the T-bill. Therefore, comparing the Sharpe Ratios of the hedge fund and the T-bill might be invalid because ex-ante risks differ. However, because Sharpe Ratios are the generally accepted empirical performance measures for diversified portfolios, we calculate Sharpe Ratios for both the recommended portfolios and the benchmark portfolios.

For the shuffled-weight and shuffled-weight change benchmarks, we report the average Sharpe Ratio from 10,000 iterations. We also calculate the number of times returns and Sharpe Ratios from the recommended-weight portfolios outperform returns and Sharpe Ratios for the two types of shuffled-weight portfolios.

5.2. Performance of average recommendations for equities, bonds, and cash

Our sample covers April 1982 through October 1997 and January 1999 through October 2005. Table 2 contains the annualized returns, their respective standard deviations, and Sharpe Ratios, using recommended weights and the various benchmark weights for the previously described sub-periods.

Over the longest sample, the April 1982 through October 1997 sample, the performance of the Houses seems to compare favorably to the benchmarks. The actual return exceeds the benchmark returns by eight to 56 basis points, and the Sharpe Ratio of the recommended portfolio equals or exceeds the Sharpe Ratio of three of the four benchmark portfolios.

In the shortest sample, April 1989 through March 1992, the return from the recommended portfolio exceeds the shuffled weight changes portfolio by only three basis points. The Sharpe Ratio of the recommended portfolio equals or exceeds one benchmark Sharpe Ratio. In two of the other three sample periods, the returns and Sharpe Ratios of the average of the House recommendations equal or exceed at least three benchmarks (and equal or exceed two benchmarks in the third).

Across benchmarks, the returns from recommended portfolio exceed at least three benchmarks in four of the five sample periods. For Sharpe Ratios, the number ranges from two to five. In each sample period, the standard deviation of returns for the recommended portfolio is different by only one or two basis points from the average standard deviation for the shuffled weight-change benchmark. Returns and Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark in all five sample periods. However, based on the significance levels of Noreen (1989), this performance is not significant. The number of times that the recommended returns exceeds the shuffled weight-change return ranges from 5249 to 7502 times. For Sharpe Ratios, the range is 5419 to 7311.

5.3. Performance of the recommendations by High-Equity and Low-Equity Houses

Badrinath and Wahal (2002) and Lakonishok et al. (1992) find that varied investment styles and strategies tend to offset one another. If so, an average of House recommendations could blur differences in measuring the differences in strategies.

Panels C and D of Table 2 summarize the performance of High-Equity and Low-Equity Houses. There are some differences between the performance of the High-Equity Houses and Low-Equity Houses that could be a reflection of a difference in strategies. The main difference is that returns and Sharpe Ratios for recommendations from the Low-

Notes to Table 2

Portfolio returns are net of transaction costs. Panel A contains data on the risk-less interest rate, the number of observations in the holding period, return performance for a Robot-weight strategy, and return performance for a buy-and-hold strategy. Panels B, C, and D contain results for all Houses, “High-Equity” Houses, and “Low-Equity” Houses, respectively. For each recommendation date, we compute the average equity holding recommendation across all Houses. We classify those Houses that recommend an equity holding greater than (less than) the average as “High-Equity” (“Low-Equity”) Houses. “Actual Weights” is the annualized mean monthly return in the period using the mean weight allocation recommended by the Houses. Return performance is also reported from two benchmarks created by shuffling actual weights and changes to actual weights 10,000 times.

Equity Houses equal or exceed the Robot benchmark and the Buy and Hold benchmark in at least four sample periods. For the High-Equity Houses, returns exceed these two benchmarks three times each, but the Sharpe Ratio is higher in only one sample period (January 1999 through October 2005).

Except for the April 1989 to March 1992 sample period, for both the High-Equity and Low-Equity Houses, the standard deviation of returns for the recommended portfolio is only two to five basis points different from the average standard deviation from the shuffled weight-change benchmark.[10] For both sets of Houses, returns and Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark in all five sample periods. However, this performance is also insignificant. The number of times that the recommended returns exceed the shuffled weight-change returns ranges from 4367 to 7469 times. For Sharpe Ratios, the range is 5094 to 7516.

6. Robustness inquiries

In previous sections, we assume that investors following advice concerning allocations to equities, bond, and cash: 1) form a portfolio using a “world” portfolio, and; 2) create a “consensus” recommendation by averaging House recommendations at each recommendation date. In this section, we investigate the potential effects of each of these assumptions. We present the results for two sample periods: April 1982 through March 1989 and April 1992 through October 1997. As can be seen in Fig. 1, a stable set of individual Houses participated throughout most of these two sample periods. Therefore, variations in performance cannot be generated by changes in the group of Houses that comprise the consensus.[11]

6.1. Investigating the effects of different portfolios

In this section, we report the results under the assumption that the reader of the *Financial Times/Economist* poll exhibits a “home-bias” and forms their equity, bond, and cash portfolio in the United States, the United Kingdom, and Japan. In addition, forming these portfolios can help us examine benchmark performance as the set of underlying assets changes (which is an aspect of style).

The results in Table 3 support the results reported in Table 2. In the April 1982 through March 1989 holding period, returns and Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark for all four portfolios. However, this performance is insignificant. The number of times that the recommended returns exceeds the shuffled weight-change return ranges from 6666 to 7502 times. For Sharpe Ratios, the range is 6645 to 7311.

In the April 1992 through October 1997 holding period, Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark for all four portfolios. However, returns for the recommended portfolio exceed the shuffled weight-change benchmark for the Japan portfolio only. Again, this performance is insignificant. The number of times that the recommended returns exceed the shuffled weight change returns ranges from 4844 to 5500 times. For Sharpe Ratios, the range is 5702 to 5874.

Panels B and D of Table 3 also illustrate an advantage of the shuffled weight-change method over the shuffled weight method. The number of times the recommended return exceeds the shuffled weight-change return is more similar to the number of times the Sharpe Ratio from the recommended portfolio exceeds the shuffled weight-change portfolio. Most likely, this result reflects the fact that the risk of the shuffled weight-change benchmark is close to the risk of the recommended portfolio. Therefore, both the returns and the Sharpe Ratios yield similar inferences.

This result does not hold for the shuffled weight benchmark. For example, in Panel B using Japanese returns and a one-tailed 5% significance level test, one would reject the hypothesis of no outperformance because the actual return bests the shuffled weight benchmark return 9,508 times. A similar test using Sharpe Ratios does not result in a rejection because the actual Sharpe Ratio exceeds the Sharpe Ratio from the shuffled weight portfolio 8287 times. A similar result appears in Panel D using U.S. returns. In this case, there is a significant underperformance using Sharpe Ratios, because the Sharpe Ratio from the actual portfolio exceeds the Sharpe Ratio from the shuffled weight portfolio only 401 times. However, using returns, one would conclude that there is to significance to the underperformance because the actual return exceeds the return from the shuffled weight portfolio 1622 times.

[10] In the April 1989 to March 1992 sample period, the difference was 19 and 17 basis points, respectively.

[11] The results, however, for the other three holding periods are similar even though individual House participation was somewhat spotty.

Table 3
Annualized return performance of the mean recommended weights for equities, bonds, and cash, by portfolio for two holding periods

Holding period: April 1982 through March 1989				
	Portfolio held by investor:			
	World	United States	United Kingdom	Japan
Panel A. Benchmark returns:				
1-mo interest rate	8.71%	7.20%	10.64%	4.75%
Monthly returns	84	84	84	84
Robot weights (60%/30%/10%)	18.20%	13.07%	20.00%	18.55%
S.D.	10.13%	11.66%	11.58%	10.71%
Sharpe Ratio, SR_{rob}	0.94	0.50	0.81	1.29
Buy-and-Hold (initial: 60%/30%/10%)	18.20%	12.78%	20.35%	20.24%
S.D.	10.70%	12.09%	13.69%	12.18%
Sharpe Ratio, SR_{bah}	0.89	0.46	0.71	1.27
Panel B. Returns for all houses using:				
Actual Weights	18.96%	12.84%	21.26%	21.13%
S.D.	11.64%	13.61%	14.22%	12.39%
Sharpe Ratio, SR_{act}	0.88	0.41	0.75	1.32
Shuffled weights (10,000 trials)	18.37%	12.79%	21.09%	20.51%
Avg. S.D.	11.33%	13.37%	13.90%	12.22%
Avg. Sharpe Ratio, SR_{shu}	0.85	0.42	0.75	1.29
Times actual weights outperform:	9324	5893	6691	9508
Times SR_{act} exceeds SR_{shu}:	7871	5077	5132	8287
Shuffled weight changes (10,000 trials)	18.84%	12.73%	21.15%	21.04%
Avg. S.D.	11.66%	13.64%	14.26%	12.41%
Avg. Sharpe Ratio, SR_{chg}	0.87	0.41	0.74	1.31
Times actual weights outperform:	7502	7245	6666	6949
Times SR_{act} exceeds SR_{chg}:	7311	7190	6645	7292
Panel C. Benchmark returns:				
1-mo interest rate	4.62%	4.20%	6.31%	1.76%
Monthly returns	67	67	67	67
Robot weights (60%/30%/10%)	10.31%	12.23%	14.38%	2.21%
S.D.	7.26%	7.34%	8.39%	11.41%
Sharpe Ratio, SR_{rob}	0.78	1.09	0.96	0.04
Buy-and-Hold (initial: 60%/30%/10%)	10.27%	12.48%	14.59%	2.31%
S.D.	7.43%	7.70%	8.82%	10.59%
Sharpe Ratio, SR_{bah}	0.76	1.08	0.94	0.05
Panel D. Returns for all houses using:				
Actual weights	10.37%	11.88%	14.05%	3.84%
S.D.	7.00%	7.12%	7.63%	9.74%
Sharpe Ratio, SR_{act}	0.82	1.08	1.01	0.21
Shuffled weights (10,000 trials)	10.53%	11.98%	14.12%	4.11%
Avg. S.D.	6.86%	7.04%	7.59%	9.69%
Avg. Sharpe Ratio, SR_{shu}	0.86	1.10	1.03	0.24
Times actual weights outperform:	731	1622	2127	1530
Times SR_{act} exceeds SR_{shu}:	180	401	1960	1467
Shuffled weight changes (10,000 trials)	10.37%	11.88%	14.05%	3.81%
Avg. S.D.	7.02%	7.15%	7.67%	9.80%
Avg. Sharpe Ratio, SR_{chg}	0.82	1.07	1.01	0.21
Times actual weights outperform:	5249	4887	4844	5500
Times SR_{act} exceeds SR_{chg}:	5874	5801	5825	5702

Portfolio returns are net of transaction costs. Panels A and C contains data on the risk-less interest rate, the number of observations in the holding period, return performance for a Robot-weight strategy, and return performance for a buy-and-hold strategy. Panels B and D contain results for "Actual Weights," which is the annualized mean monthly return in the period using the mean weight allocation recommended by the Houses. Return performance is also reported from two benchmarks created by shuffling actual weights and by shuffling actual weight changes 10,000 times. Portfolio returns are computed using total stock return indices, government bond return indices, and 1-month interest rates for the World, the U.S., the U.K., and Japan, respectively.

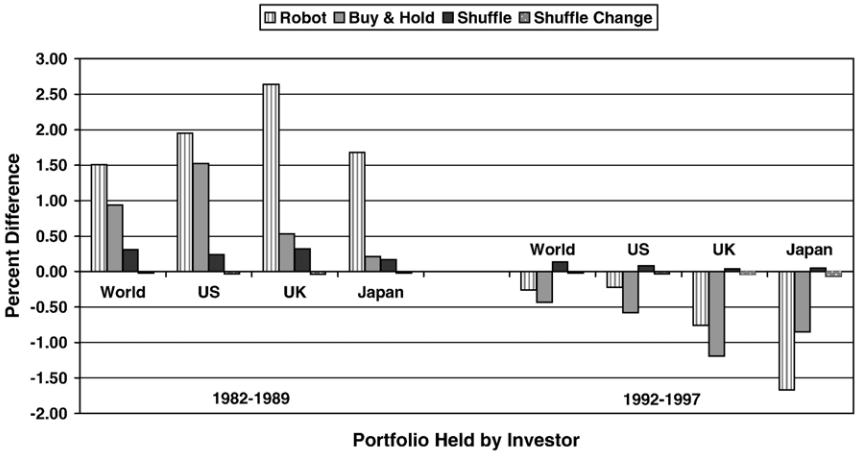

Fig. 3. Percent difference between recommended portfolio standard deviation and benchmark portfolio standard deviation, 1982–1989 and 1992–1997, by portfolio held by investor.

Using the data in Table 3, we conclude that the shuffled weight-change benchmark portfolio has a "style-preserving" and risk matching property. That is, the average portfolio return standard deviation is nearly equal to the portfolio return standard deviation from the actual recommended weights. As shown in Fig. 3, this property is robust to holding periods and return series. Given the results in Table 2, the property is also robust to whether the recommendations come from equity champions, fixed-income champions, or an overall consensus.

6.2. Investigating the effects of consensus weights

We make the common assumption that House forecasts about relative asset returns are independent. However, it is possible that the Houses set their weight recommendations in a similar manner. Similar forecasts, and weight recommendations, are possible if the Houses share common macroeconomic data to make forecasts by using widely shared techniques. As a result, the House forecasts could be highly correlated. However, Badrinath and Wahal (2002) and Lakonishok et al. (1992) find that varied investment styles and strategies tend to offset one another. Therefore, this is one reason to examine the performance of individual House recommendations. Another reason is that to the extent that individual Houses have varying styles, we can examine how the benchmarks perform given variations in style.

In our examination of individual House recommendations, we acknowledge that perhaps the analysts at various Houses could have more information and insights into their "home" markets. Therefore, we separate the Houses whose recommendations appear in a majority of the holding period into three geographic regions, the U.S., the U.K., and Japan. Of course, many of these Houses have international operations, and some Houses (i.e., Crédit Agricole) appear in the U.K. region even though they are not based in the U.K. We use the returns from the U.S., the U.K., and Japan, respectively, when we generate portfolio returns.

The results in Table 4 support the results reported in Tables 2 and 3. In the April 1982 through March 1989 holding period, returns and Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark in all cases for the individual Houses. Again, however, this performance is insignificant. The number of times that the recommended returns exceed the shuffled weight-change returns ranges from 5633 to 7935 times. For Sharpe Ratios, the range is 5411 to 8376.

In the April 1992 through October 1997 holding period, returns and Sharpe Ratios for the recommended portfolio exceed the shuffled weight-change benchmark for eight of the ten Houses. The number of times that the recommended returns exceed the shuffled weight-change returns ranges from 4484 to 7867 times. For Sharpe Ratios, the range is 5433 to 7077.

The performance of the Houses, when grouped as High-Equity or Low-Equity Houses, or when examined individually mirrors the consensus results. Therefore, we do not think that grouping the Houses into a consensus masks performance in this study. In fact, if anything, in this sample, the consensus House weight recommendations

Table 4
Annualized return performance of individual house recommended weights for equities, bonds, and cash, for two holding periods

Panel A. Holding period: April 1982 through March 1989

	United States		Japan	United Kingdom				
	Scudder Stevens	Brown Brothers	Daiwa	UBS Phillips	Capital House	Anon. One	Anon. Two	Bank Julius B.
1-mo interest rate	7.20%	7.20%	4.75%	10.64%	10.64%	10.64%	10.64%	10.54%
Robot weights (60%/30%/10%)	13.07%	13.07%	18.55%	20.00%	20.00%	20.00%	20.00%	18.61%
S.D.	11.66%	11.66%	10.71%	11.58%	11.58%	11.58%	11.58%	12.27%
Actual weights	10.78%	11.48%	15.42%	18.06%	17.32%	18.53%	17.15%	16.23%
S.D.	8.75%	8.75%	8.09%	9.16%	7.05%	7.36%	7.68%	8.11%
Shuffled weight changes (10,000 trials)	10.77%	11.39%	15.24%	17.82%	17.08%	18.32%	17.04%	16.04%
Avg. S.D.	8.76%	8.78%	8.09%	9.13%	7.11%	7.38%	7.66%	8.04%
Times actual weights outperform:	5633	6289	7935	7097	6554	7300	7369	6297
Times SR_{act} exceeds SR_{chg}:	5799	6423	8376	6513	6115	7193	7241	5411
Equity corr. with consensus weights	0.57	0.76	0.82	0.15	0.71	0.81	0.45	0.43
Bond corr. with consensus weights	0.81	0.63	0.75	0.17	0.76	0.80	−0.13	0.27

Panel B. Holding period: April 1992 through October 1997

	United States		Japan		United Kingdom					
	Lehman	Merrill	Daiwa	Nikko	UBS Phillips	Robeco	Commerz	Credit Agricole	Bank Julius B.	Credit Suisse
1-mo interest rate	4.12%	4.12%	1.91%	1.91%	6.26%	6.26%	6.26%	6.26%	6.26%	5.86%
Robot weights (60%/30%/10%)	10.50%	10.50%	3.05%	3.05%	14.12%	14.12%	14.12%	14.12%	14.12%	12.45%
S.D.	6.56%	6.56%	11.56%	11.56%	8.26%	8.26%	8.26%	8.26%	8.26%	7.06%
Actual weights	8.80%	7.78%	3.90%	3.21%	10.60%	10.04%	10.68%	10.60%	10.67%	10.24%
S.D.	3.92%	3.54%	5.16%	7.30%	4.17%	3.75%	4.55%	4.17%	4.65%	3.67%
Shuffled weight changes (10,000 trials)	8.69%	7.68%	3.84%	3.17%	10.58%	10.07%	10.65%	10.58%	10.54%	10.18%
Avg. S.D.	3.96%	3.55%	5.32%	7.32%	4.21%	3.86%	4.67%	4.21%	4.61%	3.64%
Times actual weights outperform:	6601	6689	5828	5420	5884	4484	5265	5476	7867	8526
Times SR_{act} exceeds SR_{chg}:	7077	6882	6545	5433	7044	5833	6184	6253	6427	5775
Equity corr. with consensus weights	0.47	0.01	0.54	0.61	0.06	0.46	0.41	0.19	0.29	0.23
Bond corr. with consensus weights	0.26	0.27	0.22	0.24	0.40	0.17	−0.04	0.11	0.45	0.16

Portfolio returns are net of transaction costs. Each panel contains data on the risk-less interest rate, return performance for a Robot-weight strategy, return performance using House-specified weights, and return performance from shuffling House-specified weight changes 10,000 times. Returns used are those for the countries that appear in the heading over each House. Each panel also contains the correlation between House-specified weights for equities and bonds with the overall weights for equities and bonds.

outperform individual House weight recommendations. In particular, returns from the Robot weights were about 200 to 400 basis points higher than the returns from any Individual House (the exception being Daiwa and Nikko in the April 1992 to October 1997 holding period). However, risk was also higher for the Robot weights than for the actual portfolios. By contrast, the returns and risks for the consensus weight recommendations were similar to the Robot weights.

In Table 4, we also report the correlation coefficient between equity and bond weight recommendations for individual Houses and the average, or consensus, equity and bond weight recommendations. For equities, the correlation coefficient ranges from 0.01 to 0.82. For bonds, the range is −0.13 to 0.81. By inspection, it appears that individual House weight recommendations are probably not highly correlated.

As in Table 3, a comparison of the standard deviations of the actual portfolios and the average standard deviations of the shuffled weight-change benchmarks demonstrates the risk matching property of the shuffled weight-change benchmark. This occurs because the bench-mark preserves the "style" by matching the assets in the portfolio and by matching the size of the allocation changes made to the portfolio over the measurement period. In addition, this benchmark preserves the average of the weights, their inertia, and their trends through time better than the other specified benchmarks.

7. Conditional weight measures of portfolio performance

In this section, we investigate whether the Houses have superior information using the weight-based measures developed by Ferson and Khang (2002). Because the shuffled weight-change method is a new way to examine performance, it is interesting to see whether the conclusions change when using an alternative weight-based measure. The weight-based approach of Ferson and Khang (2002) provides an interesting comparison because it produces a weight-based measure that incorporates conditioning information. Ferson and Khang (2002) find that their conditional measure is more precise than their unconditional measure. Their approach also reveals whether the Houses use conditioning information in setting their recommendations. Whether the shuffled weight-change method and the conditional weight-based method generate similar conclusions, and whether the Houses use conditioning information in generating their recommendations are the empirical questions addressed in this section.

7.1. An unconditional weight measure

Cornell (1979) was one of the first to suggest that portfolio weights should be used to measure subsequent portfolio performance. By using portfolio weights to measure performance, we can directly examine the individual choices made by portfolio managers. In the context of asset allocation, a weight-based measure examines whether the manager increases the portfolio's exposure to those asset classes that subsequently perform well, and reduces exposure to those asset classes that later perform poorly.

Studies by Grinblatt and Titman (1989), Grinblatt et al. (1995), Wermers (1999), Zheng (1999), and Anneart et al. (2005) all use the Grinblatt and Titman (1993) unconditional weight-based measure to measure portfolio performance. Grinblatt and Titman (1993) base their measure on the covariance between the weights of a portfolio and subsequent returns. Ferson and Khang (2002) modify the Grinblatt and Titman (1993) measure to create their unconditional weight measure (UWM):

$$\mathrm{UWM} = E\left\{\sum_{j=1}^{N}\left[\tilde{\omega}_{j,t} - \omega^{B}_{j,t}\right]\left[\tilde{r}_{j,t+1} - E\left(r_j\right)\right]\right\}, \tag{12}$$

where

$\tilde{\omega}_{j,t}$	Recommended weight for asset j at time t
$\omega^{B}_{j,t}$	*Benchmark weight for asset j at time t*
$\tilde{r}_{j,t+1}$	The return for asset j over the period t to $t+1$
$E(r_j)$	The unconditional mean return for asset j over the sample period

The benchmark weight, $\omega^{B}_{j,t}$, is the average weight over the measurement period. The assumption is that the average weight indicates the long-run "target" weight for the asset class. Essentially, the UWM is an average covariance over the sample. For each monthly observation, we can use the UWM to compute the covariance between recommended deviations from benchmark weights and mean-adjusted returns across asset classes. Thus, if we consider bonds, the measure captures whether the weight on bonds at time t is higher than average when abnormal returns on bonds at time $t+1$ are also high. In other words, the measure captures the covariance between the weights and subsequent abnormal returns.

The null hypothesis for the Ferson–Khang (2002) unconditional return measure is

$$H_0 : \mathrm{UWM} = 0. \tag{13}$$

If the null is rejected in Eq. (13), we could conclude that the manager in question displays skill in asset allocation.

7.2. The conditioning, or public, information set

Recent research suggests returns in international equity markets are time-varying and fairly predictable using international instrumental variables.[12] Other papers that focus on international bond markets find that bond returns are also influenced by time-varying global risk factors.[13]

Following the literature on conditional performance evaluation, we estimate conditional expected returns by regressing portfolio returns on the public information set. The conditioning information set includes the lagged one-month U.S. Treasury bill yield; the lagged ten-year bond return; the lagged dividend yield on the MSCI world stock market index; the lagged U.S. term spread between the yield on the constant maturity ten-year U.S. Treasury bond and the three-month U.S. Treasury bill yield; the lagged U.S. default spread between the Moody's BAA rated corporate bond yield and the AAA rated corporate bond yield; the lagged return on the MSCI world stock index; and a January indicator.

7.3. A conditional weight measure

Ferson and Schadt (1996) argue that a manager who uses only publicly available information in his trading strategy could be judged to have had superior performance when using an unconditional returns-based measure. Ferson and Khang (2002) demonstrate that unconditional weight-based measures also suffer from this potential bias. To eliminate this bias, Ferson and Khang (2002) develop a Conditional Weight-based Measure (CWM):

$$\mathrm{UWM} = E\left\{\sum_{j=1}^{N}\left[\widetilde{\omega}_{j,t} - \omega_{j,t}^{B}\right]\left[\widetilde{r}_{j,t+1} - E\left(r_{j,t+1}|I_t\right)\right]|I_t\right\}. \tag{14}$$

In Eq. (14), $\tilde{\omega}_{j,t}$, $\omega_{j,t}^{B}$, *and* $\tilde{r}_{j,t+1}$ are the recommended weight for asset j at time t, the benchmark weight for asset j at time t, and the return for asset j over the period t to $t+1$, respectively. $E(r_{j,t+1}|I_t)$ is the expected return for asset j over the period t to $t+1$ conditional on the public information set, I_t. We estimate the UWM and the CWM using the Generalized Methods of Moments (GMM) system and standard errors specified in Ferson and Khang (2002). We test for significance using asymptotic standard errors.

7.4. Conditional weight measure results

Table 5 contains the results for the unconditional and conditional weight measures, denoted UWM and CWM, respectively. Overall, there appears to be little evidence of performance for either the UWM or the CWM in any sub-period. All of the UWM and CWM values are insignificant. These values confirm the results from the shuffled weight-change benchmark in that there is little evidence of performance. Further, these results show that using conditioning information does not alter our conclusions.

In Table 5 we also report a value for the difference between the UWM and CWM, the Z-adjusted Weight Measure (ZWM), where Z represents the public information set. The ZWM represents mechanical trading based on publicly available information. To interpret the ZWM, we consider the following example. Suppose the Houses have no private information. Instead, they mechanically follow a momentum trading strategy. Further, we assume that momentum in index returns is true in that high returns are followed by more high returns. In this case, the Houses would have increased equity weights after equity displayed high excess returns. This action would yield a positive, significant

[12] For example see Harvey (1991), Ferson and Harvey (1991, 1993, 1994), Solnik (1993), Dumas (1994), and Harvey (1995).

[13] See Campbell and Hamao (1992), Dumas and Solnik (1995), Ilmanen (1995, 1996), and Harvey et al. (2002).

Table 5
Portfolio performance of the recommended weights using unconditional and conditional weight measures for equities, bonds, and cash

Portfolio consists of: world equity, world Gov't bonds, and 1-mo. LIBOR												
Category	Long sample:			Freq. rec. sample:			Bull market sample:			Boom and bust sample:		
	April 1982–October 1997			April 1982–March 1989			April 1992–October 1997			January 1999–October 2005		
	UWM	CWM	ZWM	UWM	CWM	ZWM	UWM	CWM	ZWM	UWM	CWM	ZWM
All houses	0.00	0.03	−0.02	0.03	−0.01	0.04***	−0.02	−0.01	0.00	−0.01	0.00	−0.01*
	(0.08)	(0.67)	(−1.60)	(1.03)	(−0.39)	(2.81)	(−1.21)	(−0.91)	(−0.57)	(−0.72)	(0.15)	(−1.82)
High-Equity houses	0.00	0.03	−0.03	0.02	−0.01	0.03**	−0.02	−0.01	−0.01	0.01	0.00	0.01*
	(−0.04)	(0.61)	(−1.54)	(0.94)	(−0.32)	(2.18)	(−0.81)	(−0.48)	(−0.53)	(0.97)	(0.06)	(1.92)
Low-Equity houses	0.00	0.01	−0.01	0.04	0.00	0.03**	−0.01	0.00	−0.01	−0.03	−0.01	−0.03***
	(−0.11)	(0.21)	(−0.74)	(1.22)	(0.06)	(2.26)	(−0.85)	(−0.21)	(−1.41)	(−1.55)	(−0.28)	(−2.63)

This table shows the results for the asset allocation choices among equities, bonds, and cash over various measurement periods for all investment houses that have sufficient data and a portfolio that averages the weights across the available houses. UWM is unconditional weight-based measure, and CWM is the conditional weight-based measure. ZWM is UWM-CWM. The benchmark weight in each case is the average weight over the measurement period. We measure the performance at monthly intervals. The conditioning information lag is one month and consists of a January binary variable, the lagged one-month U.S. T-bill yield, the lagged U.S. default spread, the lagged dividend yield on the MSCI world index, the lagged bond return, and the lagged one-month return on the MSCI world index. The *t*-statistics (in parentheses) are White–Hansen. Significance at the 0.10, 0.05, and 0.01 levels are indicated by *,**, and ***, respectively.

UWM; an insignificant CWM; and a positive ZWM. Thus, ZWM represents how changes in the weight recommendations by the Houses covary with the portion of subsequent returns that could have been predicted using past returns.

The results in Table 5 support the notion that the Houses use conditioning information in setting their asset allocation weights, but that the importance of this information in their decision-making process varies over time. For example, there is evidence consistent with this notion over the April 1982 to March 1989 and the January 1999 to October 2005 periods, but none over the April 1992 to October 1997 period. Further, in the January 1999 to October 2005 period, the High-Equity Houses and Low-Equity Houses appear to use the information in different ways. The Low-Equity Houses appear to be contrarians over this later period.

Bange and Miller (2004) and Annaert et al. (2005) also examine the Houses and find that momentum may play a role in the setting of these recommendations and their performance within an asset class (i.e. allocation between US, UK, Japan, and German equity allocations). However, they do not find evidence that momentum is important for the equities–bonds–cash asset allocation itself. However, it is possible that changes in the serial correlations in returns explain the variation in our ZWM results. We investigate this issue in two ways. First, we estimate the correlations in the returns series for the different sub-periods. For the equity returns, we estimate the following equation:

$$r_{e,t} = \alpha_e + \beta_e r_{e,t-1} + \beta_b r_{b,t-1} + \beta_c r_{c,t-1} + \epsilon_t. \tag{15}$$

We do the same for bonds and cash. We find that most of the partial correlations are insignificant and do not vary across sub-periods.[14] Hence, it appears unlikely that correlations in and across returns can help explain our ZWM results.

In the second analysis, we look at the differences in the ZWM that occur if we change the conditioning information set. First, we examine what happens when we exclude the lagged equity and bond returns from the conditioning information set. When we do this, we find results that are similar to those in Table 5, except the positive ZWM for the High-Equity Houses in the January 1999 to October 2005 period is no longer significant. Next, we estimate the weight measures if lagged equity and bond returns are the only conditioning information included. When we do this, we find the significance of the ZWM disappears. The (unreported) results for the two analyses show that correlation in returns does not appear to be important in the setting the equities-bonds-cash allocations. Rather, the other variable in the information set appear to drive the results and the importance of these as factors in setting the allocations varies over time.

[14] Detailed results available upon request.

8. Conclusion

In this study, we introduce two new performance evaluation methods that use randomization to generate a benchmark return distribution. In these randomization methods, the recommended portfolio weights, as well as the recommended portfolio weight changes, are repeatedly shuffled to generate two distributions of portfolio returns. We also introduce a method to rebalance portfolio weights to targeted weights after accounting for transaction costs. We use the randomization method and a conditional weight-based method to examine the unconditional and conditional performance of recommended portfolio weights to equities, bonds, and cash by a panel of international investment houses (the "Houses") from 1982 through 2005. The recommendations come from surveys published in the Financial Report, a confidential newsletter purchased by The Economist Newspaper, Ltd., in 1989. In the surveys, the Houses recommended portfolio allocations among equities, bonds, and cash, as well as equity holdings across countries.

We examine the performance of portfolios constructed from the average recommendation of the participating Houses at each survey date. We also look for style performance by separating the Houses into equity champions, fixed-income champions, and by examining individual House performance. We find some differences between the performance of Houses that tend to favor equities and Houses that tend to favor bonds. Houses that favor equity outperform several benchmarks toward the end of the sample period. However, the performance of the two styles is similar. We conclude that as a group or individually, the Houses do not exhibit much, if any, skill in shifting asset allocations among equities, bonds, and cash.

Blake et al. (1999) point out that "most (pension) funds would have been better off with their strategic asset allocations placed in passive index funds, and yet they purchased active management services." If investment managers decide to index, they are still left with the question of how best to allocate portfolio dollars among equity, bond, and cash index funds. Our results suggest active management services for asset allocation among equities, bonds, and cash might also be of little value to investors.

A valid comparison between Sharpe Ratios (as well as returns) is to compare an investment portfolio to a benchmark with a similar ex-ante risk. A benchmark portfolio that matches a manager's style should therefore result in a benchmark that also matches the ex-ante risk of the manager's portfolio. The shuffled-weight change benchmark portfolio exhibits a robust "style-preserving" and risk matching property. That is, the average portfolio return standard deviation is nearly equal to the portfolio return standard deviation from the actual recommended weights. This property is robust across time periods and return series. It is also robust to whether the recommendations come from individual Houses, equity champions, fixed-income champions, or an overall consensus. Further research on the properties of this benchmark are necessary. But, because a pre-specified benchmark is not needed, it is possible that this benchmark could prove useful in judging the performance of many types of money managers, including hedge funds. Additionally, inference is based on an empirical distribution so this benchmark can be used to judge the significance level of portfolio performance.

References

Annaert, J., De Ceuster, M.J.K., Van Hyfte, W., 2005. The value of asset allocation advice: Evidence from The Economist's quarterly poll. Journal of Banking and Finance 29, 661–680.

Arshanapalli, B., Coggin, D., Nelson, W., 2001. Is fixed-weight asset allocation really better? Journal of Portfolio Management 27, 27–34.

Badrinath, S., Wahal, S., 2002. Momentum trading by institutions. Journal of Finance 57, 2449–2478.

Bange, M., Miller, T., 2004. Return momentum and global portfolio allocations. Journal of Empirical Finance 11, 429–459.

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

Brinson, G., Hood, L., Beebower, G., 1986. Determinants of portfolio performance. Financial Analysts Journal 39–44 July/August.

Brinson, G., Singer, B., Beebower, G., 1991. Determinants of portfolio performance II: an update. Financial Analysts Journal 40–48 May/June.

Campbell, J., Hamao, Y., 1992. Predictable stock returns in the United States and Japan: a study of long-term capital market integration. Journal of Finance 46, 43–69.

Chakravarty, S., Sarkar, A., 1999. Liquidity in U.S. fixed income markets: a comparison of the bid-ask spread in corporate, government and municipal bond markets. Staff Reports 73, Federal Reserve Bank of New York.

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

Dumas, B., 1994. A test of the international CAPM using business cycles indicators as instrumental variables. In: Frankel, J.A. (Ed.), The Internationalization of Equity Markets. NBER, Chicago, pp. 23–58.

Dumas, B., Solnik, B., 1995. The world price of foreign exchange risk. Journal of Finance 50, 445–479.

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

Ferson, W., Harvey, C., 1993. The risk and predictability of international equity returns. Review of Financial Studies 6, 527–566.

Ferson, W., Harvey, C., 1994. Sources of risk and expected returns in global equity markets. Journal of Banking and Finance 18, 775–803.
Ferson, W., Khang, K., 2002. Conditional performance measurement using portfolio weights: evidence for pension funds. Journal of Financial Economics 65, 249–282.
Ferson, W., Schadt, R., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51, 425–461.
Goetzmann, W., Ingersoll, J., Ivkovic, Z., 2000. Monthly measurement of daily timers. Journal of Financial Quantitative Analysis 35, 257–290.
Grinblatt, M., Titman, S., 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 393–416.
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.
Harvey, C., 1991. The world price of covariance risk. Journal of Finance 46, 111–157.
Harvey, C., 1995. Predictable risk and returns in emerging markets. Review of Financial Studies 8, 773–816.
Harvey, C., Solnik, B., Zhou, G., 2002. What determines expected international asset returns? Annals of Economics and Finance 3, 83–127.
Ilmanen, A., 1995. Time-varying expected returns in international bond markets. Journal of Finance 50, 481–506.
Ilmanen, A., 1996. When do bond markets reward investors for interest rate risk? Journal of Portfolio Management 52–64. Winter.
Jones, C., 2002. A century of stock market liquidity and trading costs. Working Paper, Columbia University.
Lakonishok, J., Shleifer, A., Vishny, R., 1992. The impact of institutional trading on stock prices. Journal of Financial Economics 32, 23–43.
Liu, J., Peng, K., Wang, S., 2003. Time varying prediction of UK asset returns. Working Paper, University of Sheffield.
Noreen, E., 1989. Computer intensive methods for testing hypotheses. John Wiley & Sons, New York.
Pesaran, M., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. Journal of Finance 50, 1201–1228.
Schultz, P., 2001. Corporate bond trading costs: A peek behind the curtain. Journal of Finance 56, 677–698.
Solnik, B., 1993. The performance of international asset allocation strategies using conditioning information. Journal of Empirical Finance 1, 33–55.
Wermers, R., 1999. Mutual fund herding and the impact on stock prices. Journal of Finance 54, 581–622.
Zheng, L., 1999. Is money smart?: a study of mutual fund investors' fund selection ability. Journal of Finance 54, 901–933.