Picking winners? Investment consultants' recommendations of fund managers

Tim Jenkinson, Howard Jones, and Jose Vicente Martinez*

Abstract

Investment consultants advise plan sponsors how to invest $25 trillion worldwide. Using survey data, we analyze what drives consultants' recommendations of institutional funds, what impact these recommendations have on flows, and how much value they add to plan sponsors. We examine the aggregate recommendations of consultants with a share of 90% of the consulting market. We find that consultants' recommendations of funds are driven largely by soft factors, rather than the funds' past performance, and that their recommendations have a very significant effect on fund flows. However, we find no evidence that these recommendations add value to plan sponsors.

December 2013

Key words:

Asset management, investment consultants, institutional investors, fund performance

JEL classification: G23, G11

* Jenkinson (tim.jenkinson@sbs.ox.ac.uk) is at the Saïd Business School, University of Oxford, Oxford OX1 1HP, UK and CEPR; Jones (howard.jones@sbs.ox.ac.uk) is at the Saïd Business School, University of Oxford; Martinez (jose.martinez@sbs.ox.ac.uk) is at the Saïd Business School, University of Oxford and at the Oxford-Man Institute of Quantitative Finance. We are grateful to Andrew Lo, Ludovic Phalippou, Tarun Ramadorai and Peter Tufano for valuable comments. We should like to thank Greenwich Associates and eVestment for making available and explaining their databases, for assistance in building the combined dataset, and for valuable insights on the analysis. We also gratefully acknowledge the help of research assistant Oliver Jones.

1. Introduction

Investment consultants are important intermediaries in institutional asset management. Many retirement plans, foundations, university and other endowments, and other so-called *plan sponsors*[1] engage investment consultants to provide a range of investment services. These include asset/liability modeling, strategic asset allocation, benchmark selection, fund manager selection, and performance monitoring. It has been estimated that, as at June 2011, almost $25 trillion of institutional assets worldwide were advised on by investment consultants, including over $13 trillion in U.S. tax-exempt assets (Pensions and Investments (2011)). Goyal and Wahal (2008) estimate that 82% of U.S. public plan sponsors use investment consultants, as do 50% of corporate sponsors. Furthermore, in some countries plan sponsors are required by law to consult investment consultants before making their investment decisions.[2] From the perspective of asset managers, investment consultants are key "gatekeepers", whose opinions determine whether a plan sponsor will even consider a particular fund.

Investment consultants have largely avoided the attentions of academics, reflecting the fact that consultants have disclosed too little data to allow rigorous analysis of their activities. However, their role and influence have recently attracted interest from various quarters. The "pay to play" scandals involving some large U.S. pension schemes have revealed that some investment

[1] In this paper we use the term 'plan sponsors' instead of 'institutional investors' in order to distinguish them clearly from fund managers.

[2] For instance, U.K. pension fund trustees must "obtain and consider the written advice of a person who is reasonably believed by the trustees to be qualified by his ability in and practical experience of financial matters and to have the appropriate knowledge and experience of the management of investments" (The Occupational Pension Schemes (Investment) Regulations 2005, regulation 2(2a)).

consultants receive compensation, or kick-backs, for recommending certain asset managers (Siedle (2013)). An earlier study by the SEC (2005) investigated the potential conflicts of interest facing investment consultants. While this report does not find conclusive evidence that consultants skew their recommendations to favor certain asset managers, it does note the failure of consultants to disclose the potential conflicts of interest they face. The New York state Department of Financial Services has recently started an investigation into the role of investment consultants to the New York pension funds (Kelleher (2013)). In the U.K., a consultation on the fiduciary duties of investment intermediaries is currently on-going, which is considering whether the role and regulation of investment consultants needs to be reviewed (U.K. Law Commission (2013)). In this paper we use a unique data set to explore the role, influence and performance of investment consultants in one of the key services they provide to plan sponsors, the recommendation of funds for investment.

Although many small firms, or even individuals, act as investment consultants, in terms of assets under advisement the investment consulting industry is highly concentrated: the top ten consultants have an 81% share of the U.S. market, and the top ten worldwide are estimated to have a market share of 82% (Pensions and Investments (2011)). The five leading investment consultants worldwide by 'assets under advisement' in 2011 were Hewitt EnnisKnupp ($4.4 trillion), Mercer ($4.0 trillion), Cambridge Associates ($2.5 trillion), Russell Investments ($2.4 trillion) and Towers Watson ($2.1 trillion). It is no surprise, therefore, that institutional asset managers view being highly rated by these major investment consultants as crucial to their success.

Using thirteen years of survey data, we focus on a key service provided by investment consultants: fund selection. We examine three questions relating to consultants' recommendations of fund products.[3] First, what drives consultants' recommendations? Second, are capital flows affected by consultants' recommendations, i.e. do consultants have substantial influence on the asset allocation decisions of plan sponsors? And, third – the main focus of this paper – do these recommendations successfully predict superior performance?

The funds that are rated by investment consultants are products aimed at institutional, rather than retail, investors. Clearly, a large literature exists on retail mutual funds, and the accuracy of ratings produced by intermediaries such as Morningstar (Blake and Morey (2000), Khorana and Nelling (1998)). There is also a recent literature exploring the benefits to retail fund investors of using professional brokerage firms: Bergstresser et al. (2009) examine these benefits in terms of fund selection, while Gennaioli et al. (2013) analyze other services of financial advisers, notably the confidence these firms give to invest in financial assets at all. A large literature also exists on the accuracy of analyst recommendations for individual stocks (see, for example, Womack (1996), Barber et al. (2001), Jegadeesh et al. (2004)). On the institutional side, previous authors have analyzed the performance of investment products (in particular Lakonishok, et al. (1992), Coggin et al. (1993), Ferson and Khang (2002) and Busse, Goyal and Wahal (2010)) and the relationship between performance and the hiring and firing of investment management firms (Goyal and Wahal (2008)). However, as far as we are aware, no paper has

[3] Although the terms 'asset manager selection' and 'manager selection' are widely used in the industry, it is in fact particular funds that are recommended. We shall refer to 'funds', 'products' and 'fund products' interchangeably when referring to what investment consultants recommend. We refer to the managers of these products as 'fund managers'.

analyzed the formation, impact, and accuracy of investment consultants' recommendations of institutional funds.

The main data we use in this study is provided by Greenwich Associates, which has conducted surveys of investment consultants since 1988. The data reflect the recommendations of U.S. equity products by investment consultants for the period 1999-2011. During this period, an average of 29 investment consultants answered the survey each year. As of 2011, the consultants in the survey had a 90% share of the consulting market worldwide and 91% of the U.S. consulting market.[4] The data provided by Greenwich Associates tell us how many consultants recommend each asset management firm's products in a particular size/style category. For example, in a year in which 28 consultants responded to the survey, the database might tell us that XYZ fund manager's Large Cap Growth products received recommendations from fourteen consultants. i.e. 50% of the total for that year. As investment consultants do not publicly disclose their own past recommendations, it is not possible to analyze the recommendations of individual consultants. We also use data on the actual performance of the products rated by consultants, their relevant benchmarks, and the capital allocated to each fund.

We first analyze what drives consultants' recommendations. We do so by comparing recommendations with various attributes of fund managers that are evaluated by consultants in the Greenwich Associates surveys. These attributes include the past performance of the funds, but they also include non-performance factors; these are divided into Soft Investment Factors (i.e. factors which relate to the investment process) and Service Factors (i.e. factors which relate to service delivery). We find that consultants' recommendations correlate partly with the past

[4] Based on the Pensions & Investments (2011) survey.

performance of fund managers, but more with non-performance factors, suggesting that consultants' recommendations do not merely represent a return-chasing strategy.

Next we compare consultants' recommendations of funds with fund flows. We find very significant flows of funds into, and out of, products following changes in recommendations by investment consultants. For instance, we estimate that attracting (or losing) recommendations from one-third of the investment consultants results, on average, in an increase (decrease) of around 10% or $0.8 billion in the size of the investment product within one year. This finding shows that consultants' recommendations have a large influence on investor allocation decisions and confirms survey data which reports that manager selection is one of the most highly valued services offered by consultants.

Finally we assess the performance of the funds recommended by investment consultants. To measure fund performance we use standard one-, three- and four-factor pricing models. We also examine returns relative to benchmarks chosen to match the size and style of each product. We measure performance gross of the fees of both fund managers and the investment consultants themselves. Starting with returns relative to benchmark, on a value-weighted basis we find no evidence that recommended products significantly outperform other products, although on some measures there is an insignificant outperformance. However, on an equally-weighted basis, we find that average returns of recommended products are actually around 1% lower than those of other products. This result is confirmed using one-, three- and four-factor pricing models, and the differences using returns against benchmark and factor models are in every case statistically significant. We also measure the performance of products in the one- and two-year period after they have experienced a net increase or decrease in the number of recommendations they receive; we do so in order to test for the possibility that consultants' recommendations add value in the short term, but then become stale and fail to make a contribution. However, there is no evidence

that the net increase or decrease in the number of recommendations predicts superior or inferior performance respectively.

Given that the more highly rated funds attract more capital, and larger funds may themselves underperform owing to diseconomies of scale (as found in the mutual fund sector by Chen et al. (2004)), we investigate whether the underperformance of the recommended funds persists having controlled for assets under management. We do this because the choice of larger products may not be a free one: investment consultants may be forced to recommend products of a certain size, perhaps due to concerns about capacity constraints among small products. We find a significant negative impact of fund scale on performance, which seems fully to explain the underperformance of recommended products. Even controlling for fund size, however, there is no evidence that the recommendations of investment consultants for these U.S. equity products enabled investors to outperform their benchmarks or generate alpha.

Our analysis focuses on one asset class, U.S. active equity, which may be more efficient than other asset classes, and it is possible that elsewhere the recommendations of investment consultants are more prescient. However, U.S. active equity is a major asset class for plan sponsors, and our analysis of flows indicates that consultants' recommendations in this asset class are highly influential. This raises the question why plan sponsors engage investment consultants to help select fund managers without evidence that they add value. We identify three possible reasons. First, in keeping with the hypothesis of Lakonishok, et al. (1992), plan sponsors may value the hand-holding service provided by consultants. To use the analogy of Gennaioli et al. (2013), investment consultants are 'money doctors' with whom investors develop a relationship of trust, and this in turn gives them confidence when they select fund managers.

The second possible reason is that investment consultants provide a shield that plan sponsors can use to defend their manager selection decisions. This is in keeping with the finding

of Goyal and Wahal (2008) that plan sponsors most sensitive to 'headline risk' are most likely to use investment consultants, and also with the conclusions of Jones and Martinez (2013) that plan sponsors ignore their own expectations of fund managers' performance in favor of the recommendations of investment consultants.

The third possible reason why plan sponsors follow investment consultants' recommendations is that, owing to the consultants' lack of transparency and their own naivety, plan sponsors misunderstand the utility of these recommendations (cf. Inderst and Ottaviani, (2012) for a perspective on retail financial advice in this vein). While consultants insist on full transparency in the performance of the fund managers they rate, they do not disclose their past recommendations to allow evaluation of their own performance. Our own analysis uses the recommendations of investment consultants that comprise 90% of the market. Of course some consultants will be more accurate than others, but plan sponsors do not have sufficient information to evaluate their individual performance. Certain consultants show their 'value added' by comparing the performance of a portfolio of their recommended funds with that of a chosen benchmark. However, they do not disclose the underlying data which would allow plan sponsors to make the more informative comparison between the recommended and the non-recommended funds. As our industry-wide analysis shows, the non-recommended funds perform at least as well as the recommended funds.

In the light of our findings, a natural response by plan sponsors (or regulators) would be to require investment consultants to provide the same level of disclosure as that which is

provided by fund managers on their performance, or the same level of disclosure provided by research analysts on their stock recommendations.[5]

The remainder of this paper is structured as follows. In the next section we describe the role of investment consultants in selecting institutional funds. In section 3 we describe our data set. In section 4 we describe our methodology and we analyze the drivers of consultants' recommendations, the capital flows that they bring about, and the performance difference between recommended and non-recommended products. We also conduct robustness checks on our main results. Section 5 concludes.

2. The role of investment consultants

Recommending fund managers is only one of the services provided by investment consultants to plan sponsors. In the sequence of decision making, it comes after a number of other important advisory services. The first, which is carried out when a plan is formed, and periodically thereafter, is to help the sponsor determine and formulate the objectives of the plan, as these vary broadly between plan types (e.g. pension funds and endowments) but also between plans of the same type. For example, pension funds with a surplus will have different objectives from those of

[5] It is market practice for U.S. institutional fund managers to comply with the detailed requirements of the Global Investment Performance Standards (GIPS) when disclosing their performance. As for research analysts, FINRA Rule 2711(h)(5) requires brokers to disclose the aggregate percentage of Buy, Hold and Sell recommendations for the set of all the companies they cover for the previous twelve months. However, market practice is for brokers also to disclose their actual recommendations of individual stocks for the previous three to five years and, in some cases, even longer.

an underfunded plan. In this role the investment consultant works closely with the actuary to the plan. Next, the consultant advises the plan sponsor on an investment strategy to accomplish those objectives, including the choice of benchmarks, the broad allocation of assets between asset classes, and agreed bands within which this allocation may vary. The next stage is manager selection. Here the first step is to decide whether to opt for an active or a passive approach within each asset class and the next is to select managers for whichever approach is adopted. Once fund managers are selected, the consultant monitors their performance, and may make recommendations for termination and replacement.

Turning in detail to investment consultants' role in manager selection, plan sponsors use consultants' recommendations both when they first hire, and when they replace, managers within an asset class. As part of the hiring procedure, the consultant typically draws up a shortlist of its most highly recommended fund managers, and the sponsor, advised by the consultant, makes the final decision. Even if the plan sponsor, for fiduciary reasons, makes the final selection alone, the shortlist itself is typically drawn up by the investment consultant from its own list of recommended fund managers. The consultant services outlined above are paid for by a retainer if they are recurrent (e.g. manager monitoring) or under a schedule of charges for *ad hoc* work (e.g. manager termination and selection).

Why do plan sponsors employ investment consultants for manager selection? In many cases, ultimate fiduciary responsibility for the performance of the assets rests with trustees who are non-specialists and require independent and specialist advice. Day-to-day management of the assets is typically carried out by investment professionals employed by the plan (often headed by a Chief Investment Officer assisted by a Treasury Office), but the trustees are ultimately responsible for hiring, monitoring and firing the investment professionals and fund managers employed by the plan, as well as strategic asset allocation decisions. Investment consultants

report directly to the trustees and provide them with information and advice to allow them to discharge their responsibilities.[6]

In practice, the scope of the advice sought from investment consultants depends on the professional skills of the trustees as well as the complexity of the investment strategy they follow. For instance, if an index-tracking strategy is followed for equity investments, the scope of advice from consultants may be limited, as the choice between index-tracking funds requires little value judgment, and passive managers require little monitoring. However, if a portion of the portfolio is allocated to active managers, investment consultants are often asked to provide a list of recommended funds and to monitor those which are appointed. For this purpose, they perform due diligence on fund managers, which involves both qualitative and quantitative assessments.

Qualitative assessment looks at aspects of a manager's investment process and business, and seeks to assess whether or not the manager is likely to have some form of sustainable competitive edge. Criteria assessed typically include idea generation, the ability to implement ideas in constructing portfolios, and business factors such as quality of staff and incentive arrangements. Quantitative assessment is used to support or refine the qualitative assessment, and to establish whether portfolio construction and performance patterns tally with the professed strategy of the fund manager.

The result of this research is that funds are categorized into broad groups, based on their future expected performance relative to appropriate benchmarks. Terminology differs between investment consultants, but the following system of classification is typical. The Buy list

[6] Goyal and Wahal (2008) report that larger plan sponsors are less likely to retain consultants to assist the trustees in the selection and monitoring of investment managers, presumably as they are able to employ more experienced and better resourced investment professionals themselves.

comprises funds that the consultant recommends to plan sponsors looking to engage a new manager; funds on the Hold list are not recommended for new mandates but, if already mandated, are not recommended for termination; and funds on the Sell list are recommended for termination. Outside of these three categories are funds which are not given any rating by consultants; these tend to be funds which fall outside of the investable universe of the plan sponsors, for example owing to their small size or short track record. The recommendations on which this paper is based are those on investment consultants' Buy lists. It is worth noting that whereas recommendations of retail funds are often provided free of charge, or, in the case of equity analysts, such research is bundled together with brokerage services, the advice of investment consultants on fund managers is paid for directly by the plan sponsors. The largest number, and longest series, of recommendations available in the Greenwich Associates database relate to equity products investing mainly in U.S. companies. Such funds also constitute a large majority of the invested assets in our data, and so we focus our analysis on U.S. equity products.

The importance of investment consultants to plan sponsors is reflected in survey data. The *Pensions and Investments* (2011) survey of plan sponsors found that 23% of respondents rated investment consultants as "crucial" to the operation and success of their plans, with a further 40% rating consultants as "very important", and 26% rating them "somewhat important". As noted earlier, investment consultants offer a range of services, but when plan sponsors were asked in what area they felt their consultants added the most value, the most frequent responses were "money manager search/selection" (27%), "asset allocation development" (27%) and "performance measurement/reporting" (23%). Similarly, on the sell-side, investment management firms view investment consultants as *the* key gatekeeper to plan sponsors.

Investment consultants do not disclose their past recommendations in a way that would allow their accuracy to be measured. Some consultants show their 'value added' by comparing

the performance of a portfolio of their recommended funds with that of a chosen benchmark. However, they do not generally compare this performance with the performance of institutional funds which they do *not* recommend, nor do they make available the underlying data for scrutiny by third parties.

A final role that, in recent years, investment consultants have started to offer is 'fiduciary management' or an 'outsourced CIO' service. In this function, consultants implement the entire investment strategy once it has been agreed with the plan sponsors, including the choice of asset classes, asset allocation between classes, the selection of benchmarks, the choice of an active or passive approach, and the selection and monitoring of fund managers. For this service investment consultants tend to be paid a fraction of the assets outsourced to them. This is still a minor part of the business of investment consultants in terms of the volume of assets involved; for example, Towers Watson had 'full discretion' over $21.6 billion of assets in 2012 (AICIO (2013)), around 1% of the assets they advised. However, this business is growing fast: according to the AICIO survey, 'full discretion' mandates rose from $91 billion in 2007 to $443 billion in 2012. Since many fund management firms also provide this service, consultants and fund managers are in competition; this contrasts with consultants' manager selection business, where consultants act as gatekeepers between plan sponsors and fund managers. This has led to concerns regarding conflicts of interest on the part of some asset managers – particularly fund of funds managers – who are subjected to the detailed due diligence of investment consultants, and may end up being dis-intermediated as investment consultants offer similar products.

3. Data

Our main source of data is a series of surveys conducted by Greenwich Associates. In these surveys investment consultants are asked to rate fund managers on various measures of performance and service, and also to state the names of the fund managers they recommend to their clients for each of a number of investment styles.

Greenwich Associates (GA) has been conducting these surveys annually since 1988. We draw on the surveys between 1999 and 2011. For the period before 1999 the GA survey does not contain information on investment consultants' recommended products. Each year the survey was carried out over a two-to-four month period starting between late November of one year and early January of the next. The GA surveys cover most of the investment consultant industry. According to *Pensions and Investments* (2011), investment consultants advised plan sponsors worldwide with almost $25 trillion assets, including over $13 trillion in U.S. tax-exempt assets. Based on these figures, the consultants responding to the GA survey in 2011 had a 90% share of the consultant market worldwide and 91% of the U.S. market. The responses by individual investment consultants to the GA questionnaires are not disclosed in the survey results, but rather the aggregate responses; thus, the survey results show what fraction of all the consultants surveyed in a given year recommended a particular fund manager in a given size/style category.

The main piece of information we obtain from these surveys is an annual list of fund managers showing, in each size/style category, the percentage of the consultants surveyed who recommended that fund manager. According to GA, consultants are asked to recommend between four and six fund managers for each of a number of categories (e.g. Large Cap Growth, Small Cap Value, etc.), but they are free to go beyond or fall short of this target. The consultants' recommendations which are captured in the GA survey data correspond to the 'Buy' ratings

referred to in Section 2; for example, if XYZ fund manager scores 50% for Large Cap value, this means that, in the year in question, 50% of the investment consultants polled recommended that fund manager in that size/style category. If a consultant does not recommend a fund manager in a size/style category, this means either that the consultant rated fund manager a 'Hold' or 'Sell' in that size/style category, or that the consultant did not rate the manager in that category at all.

In this study we limit our analysis to U.S. long-only active equity products. Within this asset class we work with recommendations on seven different market-capitalization-style sub-categories: Large Cap Growth, Large Cap Value, Small Cap Growth, Small Cap Value, Mid Cap Growth, Mid Cap Value and Domestic Equity Core. Recommendation data on the first four of these categories are available throughout the entire sample period; Mid Cap recommendations are available only from 2001 and Domestic Equity Core recommendations only from 2003. If a fund manager manages more than one product in a given market capitalization-style category we aggregate those products into a single one, to make it correspond to the GA classification. The GA surveys also include a few other categories we exclude from our analysis because they are only available for a few years or because they are not clearly defined in a way that would allow a clean matching with our return data (e.g. U.S. Equity Active Quantitative products). Since we obtain our data from original documents we are confident that all recommendations are included in the database even if a product ceases to exist, or if returns are no longer reported, and so the recommendations data are free from survivorship and backfill bias.

In addition to providing a shortlist of recommended products, consultants responding to the survey are asked to express their opinions on a fund manager in an entire asset class, (e.g. U.S. Equity, International Fixed Income) rather than on individual products or groups of similar products within that class. The headings under which respondents are asked to rate fund managers are divided by GA into two sets: investment factors and service factors. Three of the

investment factors, which we call "soft" investment factors, are: clear decision making, capable portfolio management, and consistent investment philosophy. The remaining investment factor is the future performance the respondent expects from the fund manager. The service factor category includes the capabilities of relationship professionals, usefulness of reports prepared by the fund manager, effective presentations to consultants, as well as some other factors that vary from year to year. For each factor the respondent is invited to rate a fund manager's performance in each asset class on a five-point scale. Under each factor GA then aggregates the responses into a single score for each fund manager in each asset class surveyed using the Rasch model (see Andrich (1978)). In this study we work with a modified version of these scores, the fractional rank, obtained by ranking the fund manager's scores for each variable into percentiles and dividing them by 100 to arrive at a factor for each manager between zero and one.

We combine this data from investment consultants with the returns of institutional U.S. equity funds, and their assets under management, for the same period, provided by eVestment (a provider of analytic services for the institutional fund management industry). We draw data from two eVestment databases: one which tracks the monthly returns of institutional funds, and one which tracks those funds' assets under management (at an annual, and sometimes quarterly, frequency). The returns we obtain for the products in the eVestment database are "composite" returns. The individual returns earned by each client may deviate from these composite returns, but deviations are typically small.[7] Composite returns are net of trading costs, but gross of investment management fees. The data are self-reported by the fund managers, but constant

[7] For example, some investors may require that their part of the overall portfolio is purged of, say, the influence of tobacco companies or arms manufacturers.

scrutiny from clients using this data guarantees a high degree of accuracy. The return data are free from survivorship bias. For each product, the databases also provide cross-sectional information (as of June 2012) on investment style and capitalization bracket, manager-designated benchmark, and the latest fees. See Jones and Martinez (2013) for further information on the eVestment database.

In order to make the eVestment data correspond to the GA asset class of U.S. long-only equity, from where our sample of shortlisted products is drawn, we first eliminate index funds (including enhanced index funds), hedge funds, REITs and retail funds. Given that the GA survey collects recommendation data only for Large Cap Growth, Large Cap Value, Small Cap Growth, Small Cap Value, Mid Cap Growth, Mid Cap Value and Domestic Equity Core products, we also eliminate products that do not match these capitalization and style sub-categories.[8] We also drop observations for Mid Cap Growth and Value products before 2001 and for U.S. Equity Core products before 2003, as the GA survey did not ask for recommendations on these products before these dates.

Table I provides descriptive statistics for the final sample. In the upper part of Table I, the first two columns show the number of investment consultants surveyed each year from 1999 to 2011, and the total number of recommendations collected in each survey. The table also shows the total number of U.S. Equity products in our sample each year, recommended and not recommended, as well as the mean assets under management (in millions of U.S. dollars) for those products which reported this figure. The average number of available products during the

[8] We match U.S. Equity Core products in the GA surveys to Large Cap Core, Mid-Large Cap Core, Mid Cap Core, and All Cap Core products in the eVestment databases.

sample period is 1,919. The number of products is smaller in the early years of the sample because there is no recommendation data on products in the Mid Cap (Value and Growth) and Equity Core products categories in these years. Approximately 21% of the products received at least one recommendation each year, a percentage that remained relatively stable during the entire sample period (other than for the first year). During the same period, an average of 29 investment consultants answered the survey each year, each of them making, on average, 55 recommendations per year across the seven equity categories in our sample, which means that each recommended product received an average of 4 recommendations. End of year asset data are available for approximately 68% of the total sample. Average assets per product during the sample period are $1.5 billion. Recommended products are substantially larger than non-recommended ones: average assets managed by rated product are $4 billion, whereas the average size of unrated products is only $0.8 billion. At the end of our sample period the total assets under management of (recommended and non-recommended) products in the sample are $3 trillion. Once we account for our sample exclusions, our descriptive statistics are similar to those reported by Busse et al. (2010), which suggests that the coverage of the database is similarly comprehensive.

The lower part of Table I presents similar information for each market capitalization-style category. The majority of products (841) are in the Large Cap (Value and Growth) segments, and the smallest number (280) is in Mid Cap (Value and Growth) segments. Large Cap products are also the ones that attract the highest number of recommendations each year (892), but, on average, the percentage of Large Cap products receiving at least one recommendation is not different from that of any of the other categories. Unsurprisingly, average portfolio sizes are biggest for Large Cap products ($6 billion and $5.6 billion for Growth and Value products respectively) and smallest for small cap ($1.3 billion and $1.5 billion for Growth and Value products).

4. Methodology and results

The first part of this section investigates the factors which drive investment consultants' recommendations of institutional investment products; the factors we analyze include not only the past performance of the product, but also the non-performance attributes which the consultant identifies in the fund manager, that is, the Soft Investment Factors and Service Factors. In the second part we assess the impact of investment consultants' recommendations on flows into and out of these investment products, showing the extent to which investment consultants' recommendations are actually followed by plan sponsors. In the third part we examine whether investment consultants' recommendations of fund managers add value to plan sponsors, by comparing the performance of recommended and non-recommended products. In the final part of this section we test the robustness of our performance findings by allowing for possible back-fill bias, and we explore whether differences in size between the products recommended by consultants and other products can explain the observed underperformance of the first.

A. Drivers of Recommendations

In this section we explore what drives investment consultants' recommendations. In particular we explore whether recommendation decisions are a function of past product performance, a common phenomenon in financial analysts' decisions to recommend stocks (Altinkilic and Hansen, (2009)), or of other less tangible factors, namely the Soft Investment Factors and Service Quality Factors that consultants evaluate in fund managers.

We proceed by estimating a Poisson model, with the standard exponential mean parameterization, on pooled yearly data. The pooled Poisson estimator assumes that the number of recommendations received by a product i at time t ($Recs_{i,t}$) is Poisson distributed with a mean of:

$$E\left(Recs_{i,t}\middle|x_{i,t}\right) = Consultants_t \times exp\left(\alpha + \beta_1 Past\ Perf_{i,t} + \beta_2 Soft\ Inv.Factors_{i,t} + \beta_3 Service\ Factors_{i,t} + \beta_4 Return\ Vol_{i,t}\right) \quad (1)$$

Consultants$_t$ is the number of investment consultants surveyed in year *t,* which we treat as an exposure variable (as we would expect the number of recommendations received by a given product on a given year to be proportional to the number of consultants surveyed that year, other things equal). *Past Perf*$_{i,t}$ is either the return or Fama-French three factor alpha fractional rank of product *i* in relation to the other products in the same market capitalization and style category over the two year period immediately preceding the survey from which recommendations are collected. *Soft Inv. Factors*$_{i,t}$ is the fractional rank at time *t* of a set of Soft Investment Factors of fund manager *i*'s U.S. equity team (i.e., Clear Decision Making, Capable Portfolio Manager, and Consistent Investment Philosophy). *Service Factors*$_{i,t}$ is the fractional rank at time *t* of a set of Service Factors of fund manager *i*'s U.S. equity team (i.e., Capabilities of Relationship Professionals, Usefulness of Reports prepared by the fund manager, Effective Presentations to consultants). Finally, *Return Vol*$_{i,t}$ is a measure of product *i*'s return volatility over the two year period preceding the survey.

When estimating this model we use robust standard errors clustered at the product level. The Poisson Quasi-MLE estimator we use retains consistency if the count is not actually Poisson distributed, provided that the conditional mean function is correctly specified (see Wooldridge, (2010)).

Table II shows the results of this exercise. The table contains both coefficient estimates and average marginal effects; the latter are computed as the product of the coefficient estimate times the average number of recommendations received by the products in our sample. Models I and II use the soft investment and service quality indexes as regressors whereas Models III and

IV replace them with (some of) its components.[9] Results suggest that investment consultant recommendations are at least partly driven by past good performance. Moving from the bottom to the top percentile of past performance leads, on average, to an increase of half a recommendation in the average number of recommendations received by a product (out of an average maximum of 29 recommendations per year, the average number of consultants issuing recommendations in our sample). But the most important driver of recommendations seems to be, not past product performance, but the Soft Investment Factors which consultants identify in fund managers, notably Capable Portfolio Manager and Consistent Investment Philosophy. Estimates in Models I and II indicate that moving from the bottom to the top of the Soft Investment Factors ranking is rewarded, on average, with slightly more than six extra recommendations per year. Service Factors, and in particular Capabilities of Relationship Professionals and Usefulness of Reports, also appear to be important drivers of recommendations. Estimates in Models I and II suggest that improvements in service quality, from the bottom percentile to the top one, lead to 1.2 extra recommendations received per product per year.

Although an analysis of the distribution of the number of recommendations per product indicates that the data is over-dispersed, thus violating the Poisson variance assumption, the Quasi-MLE estimator used to estimate Models I to IV is robust to this problem. For further robustness, however, Models V to VIII in Table II repeat the estimation using the Negative Binomial (NB2) of Cameron and Trivedi (1986), a particular parameterization of the negative

[9] We exclude Service Factors which were part of the survey for only part of the sample period (to avoid losing too many observations in our estimations).

binomial distribution that allows for over-dispersion, instead of the Poisson distribution.[10] As shown in columns V to VIII results are nearly identical to those of the Poisson model, suggesting that both models provide similar fit for the conditional mean.

To summarize the results in this section, we find that investment consultants' recommendations of funds are a function of the past performance of those funds, but especially of the two sets of non-performance factors (Soft Investment Factors and Service Factors) which consultants identify in fund managers. In particular, Soft Investment Factors appear to have a far more powerful effect on consultants' recommendations than past performance. So although consultants' recommendations to some extent reflect a return-chasing strategy, they seem to be more heavily influenced by the consultants' qualitative judgment of intangible factors.

B. Flows

Jones and Martinez (2013) show that plan sponsors follow investment consultants' recommendations of fund products. Their analysis focuses on the incremental effect on flows of consultants' recommendations, over and above the effect of such recommendations which is channelled through the other variables of interest, notably through the expectations of plan sponsors themselves.[11] In this section we explore how asset flows respond to consultants' recommendation changes at the level of individual products (Large Cap Growth, Small Cap

[10] Unlike the Poisson Quasi-MLE this estimator is not consistent if the variance specification is not correct. It is however very flexible and tends to fit probabilities better than the Poisson model when data is over-dispersed.

[11] The flow analysis in Jones and Martinez (2013) is at the level of asset class, not at the level of individual products, and the set of explanatory variables includes plan sponsors' expectations of future performance (which is affected by the advice received from consultants).

Value, etc.), rather than at the asset class level. We also capture the full effect of consultants' recommendation changes, rather than just their incremental effect. We do this by expanding a typical flow-performance regression (see, for instance, Ippolito (1992), Chevalier and Ellison (1997) and Sirri and Tufano (1998)) to include a recommendation change variable as regressor.

We consider two flow measures. The Dollar Flow in and out of an investment product is defined as the yearly change in the total net assets minus appreciation:

$$\$Flow_{i,t} = TNA_{i,t} - TNA_{i,t-1} * \left(1 + r_{i,t}\right) \qquad (2)$$

where $TNA_{i,t}$ is the total net assets for product i at date t, and $r_{i,t}$ is the return on product i between dates t-1 and t. This measure reflects the growth of a fund in excess of the growth that would have occurred if no new funds had flowed in but dividends had been reinvested.

The second measure is the Percentage Flow relative to the total net assets invested in the product as of the end of the previous year:

$$\%Flow_{i,t} = {}^{\$Flow_{i,t}}\Big/_{TNA_{i,t-1}} \qquad (3)$$

The bivariate relationship between recommendation changes and Dollar and Percentage Flows is shown in Figure 1.[12] The graph plots show the results of estimating kernel weighted local linear regressions of Dollar Flows (Panel A) and Percentage Flows (Panel B) on lagged changes in consultants' recommendations. The results indicate a positive relationship between the change in consultants' recommendations (measured as the change in the percentage of recommendations

[12] To reduce the effect of outliers on the coefficient estimates, we winsorize the percentage flows variable at the 95th percentile (as in Barber et al. (2005)). Similar results obtain if we remove small products from the sample instead.

received by a product over the total possible) and subsequent flows. We are interested, however, in measuring how flows respond to recommendation changes, controlling for publicly-available measures of past performance, as well as for other product attributes known to affect flows and which could also affect recommendations (namely past performance, fund size, and return volatility). We estimate the response of flows to recommendation changes using the following regression on yearly data:

$$Flow_{i,t} = \alpha_t + \beta_1 \Delta Consultants\ Recs_{i,t-1} + \beta_2 Past\ Perf_{i,t-1} + \delta' Controls_{i,t-1} + \epsilon_{i,t} \quad (4)$$

$Flow_{i,t}$ is either the Dollar Flow or the Percentage Flow. $\Delta Consultants\ Recs_{i,t-1}$ is the change in the number of recommendations product *i* received, as a fraction of the highest possible number of recommendations which that product could have received from all of the consultants in our sample, between time *t-2* and *t-1*. $Past\ Perf_{i,t-1}$ is a set of performance measures of product *i* at date *t-1* (the one year return and Fama-French three factor alpha rankings of a product in relation to the other products in the same market capitalization and style category). The regressions also include the total net assets for product *i* at date *t-1* (in the Percentage Flow regression we use log assets as the regressors; see Del Guercio and Tkac (2002)), a measure of product return volatility over the previous two years, and a full set of time dummies (one per each year in the sample) as additional controls.

Table III reports the results of estimating this regression using pooled time-series cross-sectional data and Dollar (in millions) and Percentage Flows as the dependent variable. Each column in this table represents a separate regression. The coefficients of the variables capturing the effect of lagged changes in recommendations on absolute flows are positive and statistically significant. This suggests that plan sponsors respond to the investment consultant

recommendation changes by moving money in the direction implied by the recommendation change. The estimate in column I indicates that moving from a situation where no consultant recommends the product to another when all of the consultants in our sample recommend it leads to an extra inflow of assets of $2.4 billion. Qualitatively similar results obtain if we use percentage flows as the dependent variable. Estimates in column IV suggest that a product shortlisted by all the consultants in the sample, in the previous year, receives, on average, extra inflows equal to 29% of the assets managed by that product in the previous year, compared to a similar product that is not shortlisted by any consultant. In all cases t-statistics are based on clustered standard errors, which are White heteroskedastic-consistent standard errors corrected for possible correlation across observations of a given investment product (White (1980) and Rogers (1993)). This method seems to be the most sensible given the size of our panel (see Petersen (2009)).

The difference between, on the one hand, columns I and IV and, on the other, columns II and V is that II and V include also lagged recommendation levels (as opposed to changes) as regressors while I and IV do not. The economic reason for including lagged recommendation levels is that institutional money may be slow to react. This regressor reflects the extent to which the level of recommendations explains flows in a steady state, independently of any change in those recommendations, and helps account for flows arriving more than one year after the recommendation change. The estimates in models II and V indicate that the bulk of the effect of recommendation changes on flows happens in the year after the recommendation change, with relatively minor effects visible later on.

Estimates in columns III and VI of Table III suggest that the relation between flows and past recommendation changes is nearly linear. In these models we replace the recommendation change variable with two variables capturing positive and negative recommendation changes.

Although the coefficients attached to the negative change variables seem slightly higher than the regression coefficients on positive recommendation changes, differences between the two slopes are not statistically significant.

Our regressions also indicate that previous performance has a large and significant impact in asset flows (a result that is very much in line with the previous literature on the topic; see, for example, Del Guercio and Tkac (2002)). In these regressions the figures against fractional performance ranks are the dollar or percentage change in assets in the current year that reflect the difference between the bottom percentile and the top percentile of raw returns or three factor alphas in the previous year.[13] Taking into account both measures of performance in combination, estimates in models I and IV indicate that moving from the 25th percentile of performance to the 75th percentile is rewarded with a 18% increase in assets (inflows of around $235m). Interestingly, the estimates in columns III and VI indicate that the performance measure that most strongly affects assets flows is the product's excess returns over the chosen benchmark and not its risk-adjusted returns. This is consistent with the widespread use in the industry of this measure.

To summarize, we find that changes in investment consultants' recommendations have a large and significant effect on flows into institutional investment products.

C. Performance

We measure the performance of consultants' recommended products by estimating factor models using time-series regressions. To generate aggregate measures of performance, we create equal- and value-weighted portfolio returns of recommended and not recommended products available

[13] A portfolio's fractional performance rank represents its percentile performance relative to other U.S. equity investment products in the same equity category and period and ranges from 0 to 1.

in each month. The weight used for value-weighting is based on the assets in each product at the end of December of the prior year. With these returns, we estimate:

$$r_{p,t} = \alpha_p + \beta_p' f_t + \varepsilon_{p,t} \qquad (5)$$

where r_p is the portfolio excess return (over the risk-free return), f_t is a vector of excess returns on benchmark factors, and α_p is the abnormal performance measure of interest. We use three established factor models: CAPM (Sharpe (1964), Jensen (1968)), the Fama-French (1993) three factor model and the Fama-French-Carhart four factor model (Carhart (1997)). We obtain these four factors from Kenneth French's web site.

In addition to these measures we also report the average returns of the products in excess of a selected benchmark. The benchmarks we use are listed in Table A1 of the appendix and are standard in the investment industry. Viewed through the academic lens, such performance measures are clearly highly restrictive versions of more general asset pricing models, but they are, nonetheless, key metrics that are used in practice by plan sponsors, fund managers and investment consultants to evaluate performance. Moreover, from the practitioner's point of view, academic models, especially three- and four-factor models, may seem highly demanding: fund sponsors are often happy to give credit to fund managers for allocation decisions which are not reflected in alpha under such models. For example, although practitioners use 'style' benchmarks such as 'small caps', investing in very small stocks could enhance a fund's performance against this benchmark, but this 'outperformance' would be factored out in academic models.[14]

[14] This is not the only source of differences between the approaches. Cremers et al. (2012) note that the practitioner and academic approaches can yield very different results, as the standard Fama-French and Carhart factor models may assign non-zero alphas even to passive benchmark indices such as the S&P 500 and Russell 2000. At the same

Table IV shows the average returns, excess returns over a portfolio of chosen benchmarks and one-, three- and four-factor alphas obtained by the portfolio of U.S. equity products recommended by the investment consultants in our sample. It also shows the performance of institutional products not recommended by any of the consultants. All reported figures are gross of management fees. In this analysis the recommended portfolio includes each fund as many times as it is recommended. Results in this table indicate that, in the 13-year period of our study, and on an equal-weighted basis, the portfolio of all products recommended by investment consultants delivered average returns before management fees of 7.13% per year. These returns are lower than the returns obtained by other products available to plan sponsors, which are not recommended by consultants. The annualized difference in returns between recommended products on the one hand and products which are not recommended on the other is -1.00%. Since there is substantial heterogeneity in the products in our sample in terms of size of the stocks they invest in, and in their value/growth characteristics, raw returns may not be very appropriate for performance comparisons. When we risk-adjust returns using the three- and four-factor models, recommended products obtain an alpha of 1.14% per year, once annualized. This is still significantly lower than the alpha obtained by non-recommended products (the difference between recommended and non-recommended products is statistically significant at -0.86% per year). Risk-adjusting returns using benchmarks chosen to match the products' style and market capitalization delivers almost identical results. Results in Table IV also indicate that the

---

time, the classification system on which benchmark-adjusted returns are based is ambiguous and leaves much room for interpretation resulting in frequent misclassifications (some of them perhaps deliberate). For example, in a study of U.S. mutual funds, diBartolomeo and Witkowski (1997) find that almost 40% of all U.S. equity funds are misclassified.

underperformance of the recommended products on an equal-weighted basis is most notable among the most frequently recommended products (the portfolio of the top 50% most recommended products).

When we perform the same analysis on a value-weighted basis recommended products still obtain lower returns (or CAPM alphas) than those obtained by non-recommended products, but outperform them based on a three- or four-factor model. None of these differences is statistically significant, however. Value-weighted returns and alphas are consistently lower, suggesting that smaller products perform relatively better.[15]

In all cases the returns we work with are gross of asset management fees. From the fund sponsor's point of view a more relevant comparison would be between net of fees returns. However, evidence from eVestment 2012 fee records indicates that intra-style variation in fees for active U.S. Domestic Equity managers is extremely small (see also Busse et al. (2010)). This means that our performance comparisons do not lose much by reflecting gross rather than net returns. Our returns do not take into account the impact of the fees of the investment consultants themselves either.

The funds we are analyzing manage equities according to different styles: value, growth, small cap, large cap, etc. In Table V we separately consider the subcategories within our sample: Large Cap Value, Large Cap Growth, Mid Cap Value, Mid Cap Growth, Small Cap Value, Small Cap Growth and Domestic Equity Core products. The table compares the performance of recommended products with the performance of all other products in the sample in each of these

[15] Our equal- and value-weighted figures, once aggregated across recommendation categories, are generally in line with those reported by Busse et al. (2010) for their sample of institutional products.

seven categories. The results we obtain broadly confirm those presented in Table IV: recommended products underperform other products in all the categories studied when returns are equal-weighted. On a value-weighted basis results are mixed.

Our results so far suggest that investment consultants are not able consistently to add value by selecting superior investment products. This is particularly true when we compare recommended products with non-recommended products on an equal-weighted basis. This is a surprising result, as we should expect recommended products to outperform other products by a margin sufficient to cover the cost of hiring the investment consultant. How can we explain this result? Why do the investment consultants select products that appear to underperform other products significantly on an equal-weighted basis, but not on a value-weighted basis? We turn our attention to these questions in the next section.

D. Performance, Robustness and Further Analysis

Results in the previous section suggest that products recommended by investment consultants do not outperform other investment products and may in fact underperform them. In this section we explore those results in more detail. First we explore the possibility that backfill bias may affect our results and, in particular, that it may be responsible for the relatively good performance of non-recommended products, which would therefore constitute an unfair benchmark for recommended products. Second, we examine the impact of investment product size on performance. This allows us to assess the relative performance of recommended and the (generally) smaller non-recommended products controlling for the potential effect of product size on performance. Finally, we investigate the short term performance of products that experience a net increase (decrease) in the number of recommendations they receive, as it is possible that consultants' failure to add value by their recommendations is due, not to their inability to identify

outperforming products, but to the fact that they continue to recommend those products for too long.

Because we work with self-reported returns a natural concern in the measurement of performance is that the data might not accurately reflect the performance of worse performing products. Since managers may have a greater incentive to volunteer information to eVestment after a period of good performance, products may be subject to 'instant history' or backfill bias, as described by Fung and Hsieh (2000). Moreover, this problem could be more serious for smaller non-recommended products if they are subject to less scrutiny than their recommended counterparts. It is unlikely that this problem can ever be completely eliminated, but we follow the approach in Jagannathan et al. (2010) to determine the potential impact on our results. We eliminate the first one, two and three years of returns for each product and re-run our main performance regressions. By requiring three years of return history to show fund performance, this procedure also addresses another concern with our performance comparison: most pension sponsors and consultants require the existence of a three-year performance track record to be considered in the initial phases of a manager search (Del Guercio and Tkac, 2002).

Results reported in Table VI reveal that performance figures may be affected by backfill bias, but the evidence we find suggests that it is not enough to affect our main results. After eliminating the first three years of return history, on an equal-weighted basis the three-factor alphas of non-recommended products decline by only 0.21% (from 2.00% to 1.79%), and those of recommended products decline by 0.15% (1.14% to 0.99%). The difference is significantly smaller than the actual gap in performance between recommended and non-recommended products reported in Table IV, -0.85% per year for equally-weighted portfolios.

Evidence from value-weighted returns indicates that the performance of recommended and non-recommended products as reported is very similar to the performance figures obtained

after eliminating the first one to three years of reported history (regardless of the model used to measure performance). This suggests that backfill bias is not a problem among those products which also report assets under management or, to put it differently, that those fund managers that have backfilled data have probably backfilled only the return data.

A second concern is the impact of product size on returns. Table I presented evidence suggesting that investment consultants tend to concentrate on larger products. There is research showing that funds that manage more assets perform worse (see Chen et al. (2004)), a finding that is consistent with results in Table IV showing that funds recommended by consultants perform worse (in comparison with non-recommended products) on an equal-weighted than on a value-weighted basis. We therefore re-assess the investment performance of recommended products in light of this tendency. The choice of larger products may reflect the investment consultants' optimization of their own research and monitoring efforts, or a belief that recommending a large well-known fund manager will be easier to justify after the event. However, the preference for recommending large products may not be entirely free, for consultants may be required by plan sponsors to recommend products of a certain size, perhaps because of doubts about the ability of small products to handle a larger pool of assets.

To control for the impact of product size on performance, we use a regression-based generalization of the calendar time portfolio approach (see Hoechle et al. (2009) and Dahlquist et al. (2011)). This generalization relies on estimating, at the investment product level, a pooled linear regression model with Driscoll and Kraay (1998) standard errors. Driscoll-Kraay standard errors are robust to heteroskedasticity and general forms of cross-sectional and temporal dependence. The advantage of the regression-based approach is that, in a regression framework, it is straightforward to include continuous and multivariate explanatory variables, and so to control for product size.

The pooled linear regression model we estimate has the following panel structure:

$$r_{i,t} = \left(\alpha + \alpha_z' z_{i,t}\right) + \left(\beta + \beta_z' z_{i,t}\right)' f_t + \varepsilon_{i,t} \quad (6)$$

where $r_{i,t}$ is the excess return for product i in period t, and where we condition on time-varying product characteristics $z_{i,t}$. Product characteristics include investment consultant ratings, captured by a full set of dummy variables, and de-trended log assets under management at the end of the previous period.[16] The risk-adjusted performance of recommended products is computed using a four-factor model represented in equation (6) by the vector f_t.

Table VII provides the results of estimating different specifications of equation (6). These specifications differ according to whether we control for lagged assets under management, and whether we include products with no information available on assets under management. The coefficients of interest are those of α_p, since they inform us if recommendations and the other control variables have predictive power over abnormal returns. In this table the constant captures the expected monthly excess return or alpha of a non-recommended product (of average size in the models that include lagged de-trended assets under management as controls) and the coefficient associated with the recommended dummy indicates the expected extra performance delivered by recommended products. Models I and IV are the closest to the equal-weighted calendar time specifications showed in Table IV. A difference between the results in Table IV and Models I and IV of Table VII is that the panel we estimate in equation (6) is unbalanced and

[16] We use de-trended log assets under management (de-trended by subtracting from the log of assets under management the mean of this variable across all period t observations) to address the possible non-stationarity of log assets under management. However, almost identical results obtain if we replace de-trended log assets under management with standard log assets under management.

therefore the weight attached to each month in the sample is not the same. The panel gives the same weight to each individual observation whereas the portfolio method employed in Table IV gives the same weight to each monthly observation. Another source of difference between these results is that Table IV shows results for portfolios that include each product as many times as they have been recommended whereas Table VII does not. If we modify the weights in the panel regression to account for these differences, as indicated in Hoechle et al. (2009), both results coincide.[17] In model I, where performance is assessed using industry-standard excess returns over benchmark indices, recommended products underperform non-recommended products by 0.88% per year; less than the 1.10% per year reported in Table IV but still highly statistically significant.[18] This difference shrinks if we exclude from the sample products that do not report assets under management (Model II) and completely disappears when we control for the de-trended natural logarithm of assets under management in each product at the end of the previous year (Model III).

Similar conclusions can be drawn by looking at four-factor alpha models (Models IV to VI), adjusted and unadjusted for product size. In model IV, recommended products underperform non-recommended products by 0.36% per year but the difference is not statistically significant this time. Moreover, once we control for lagged product size (Model VI), the underperformance

[17] Hoechle et al. (2009), following Loughran and Ritter (2000), argue however that weighting all observations equally makes more sense.

[18] Annualized differences are computed by multiplying monthly coefficients times 12. Notice that since monthly returns are linear in the (same) set of variables, it does not matter where the difference is calculated (i.e., for which values of the lagged assets under management, the variable used as control).

of recommended products disappears or even turns into a small, but still not statistically significant, outperformance (0.32% per year).

Finally, we look at the performance of portfolios of US equity products that experience a net increase (decrease) in the number of recommendations they receive. This analysis has two objectives: first, to explore whether consultants' failure to add value by their recommendations is due to their inability to identify outperforming products or to the fact that they keep them on their shortlist of recommended products for too long; and, second, to provide an alternative benchmark in the performance analysis, by concentrating on products that are unquestionably in the choice set of (at least some of) the consultants we study.

We proceed by forming two different portfolios. The first portfolio includes all investment products that experience a net increase in the number of recommendations received; these are products that, in net terms, are being added to the shortlist of recommended products. The second portfolio includes all investment products that experience a net decrease in the number of recommendations received; these are products that, in net terms, are being dropped from the shortlist of recommended products. Products are included in these portfolios in the month in which they experience the increase/decrease in the number of recommendations and kept there for 12 or 24 months. Each product is included in these portfolios as many times as it is newly recommended/de-recommended, thus giving more weight to products that experience a larger increase/decrease in the number of recommendations received.[19]

[19] It is possible that some of the portfolios that we consider to have experienced an increase/decrease in the number of recommendations did in fact receive the same number of recommendations as in the previous year, and that we capture instead the effect of the changing composition of the survey. However, because the coverage of the sample is large and relatively stable through time, this problem, which may add noise to our estimates, is likely to be limited.

Table VIII compares the performance of these portfolios. Results indicate that products that experience a net increase in the number of recommendations do no better than products that are on average being de-recommended by consultants. In fact they do worse, according to all our measures, although not significantly so. The difference in performance goes from a few basis points per year to more than three percent per year in some specifications, but it is never statistically significant (the number of products in each portfolio is considerably smaller than in previous tables).

One of the objections that could be made against the results we presented in the previous section is that some of the non-recommended products included in the analysis are effectively off-limits to the consultants (because they are below the size threshold of plan sponsors, because they lack sufficient longevity, or for some other reason). However, results in Table VIII suggest that consultants fail to identify superior future performers, even among the set of products that are on their radar. We conclude that the underperformance of recommended products can be explained by consultants' tendency to recommend relatively large products. This may reflect the fact plan sponsors and consultants are generally concerned about liquidity risk, and plan sponsors avoid products that are small relative to their holding size. We understand, for example, that plan sponsors would typically avoid holding more than 10% of the total assets under management of a product. At the same time investment consultants may be inclined to focus on products that can be used broadly across their client base; this would economize the costs of monitoring client fund holdings, and would guard against a liquidity shortage in a product if, following a rating downgrade by the consultant, many plan sponsors withdrew their assets simultaneously.

However, even allowing for this constraint, recommended products still fail consistently to outperform other products in our sample on a gross basis. And they would need to deliver higher gross performance to offset consultant fees.

5. Conclusions

Using survey data from consultants with a combined share of 90% of the investment consulting market worldwide and 91% of the U.S. consulting market, we analyze these consultants' recommendations of U.S. active equity products over a 13 year period, and we examine the drivers, the impact, and the accuracy of these recommendations.

We find first that, while consultants' recommendations of fund products are partly a function of the past performance of those products, it is mainly the fund managers' non-performance attributes which drive recommendations. Consultants are not merely 'return-chasing' when they form their recommendations. Second, we find that investment consultants' recommendations have a large and significant effect on institutional asset allocation. Third, we find no evidence that consultants' recommendations add value to plan sponsors.

On an equal-weighted basis, the performance of recommended funds is significantly worse than that of non-recommended funds, while on a value-weighted basis the performance is mixed, and the recommended and non-recommended products do not perform significantly differently from each other. When we measure the performance of funds in the one- and two-year period after a net increase or decrease in the number of recommendations they receive there is, similarly, no evidence that consultants add value by their recommendations. The underperformance of recommended products on an equal-weighted basis can be explained by the tendency of consultants to recommend large products which perform worse. When we adjust for the different sizes of recommended and non-recommended products, we find that recommended products still fail consistently to outperform non-recommended products. The same result holds when we adjust for possible backfill bias.

Our analysis has shown performance in gross terms because, in the U.S. long-only active equity asset class we study, fees are relatively consistent across products. Fund manager fees vary by product and size of mandate, but are typically around 50 basis points per annum (and, of course, investment consultants charge for their services as well). When the performance of recommended funds is shown net of such fees, their failure to add value becomes more pronounced. In light of the evidence presented, what is striking is that fund sponsors follow such recommendations to the extent and at the expense that they do.

Plan sponsors have other reasons to engage investment consultants other than their fund manager recommendations: asset/liability modeling, strategic asset allocation, etc. It is also possible that the payments which plan sponsors make for consultants' manager selection services in fact reward other benefits provided by consultants, which are not directly rewarded. For example, consultants may provide plan sponsors with a 'narrative' to explain their actions, which provide comfort to plan sponsors, and allow plan sponsors to explain their actions to their own stakeholders. This would be an example of the hand-holding service described by Lakonishok et al. (1992), or of the trust-based relationship that investors have with their adviser, as patients have with their doctor (see Gennaioli et al. (2013)).

However, while these reasons might explain why plan sponsors engage investment consultants in general, they do not tell us why they follow consultants' recommendations when they are apparently not rewarded for doing so. Assuming plan sponsors know that they are not being rewarded for following consultants' recommendations, one possible reason for doing so is that plan sponsors hide behind consultants' recommendations when they have to account for their decisions. As Goyal and Wahal (2008) find, plan sponsors that are more likely to be sensitive to adverse publicity ('headline risk') are more liable to use investment consultants. Jones and Martinez (2013) put forward evidence that plan sponsors follow consultants' recommendations

even against their own judgment. According to this explanation, a tendency for plan sponsors to follow investment consultants, in the knowledge that their recommendations do not add value, would be evidence of an agency problem.

However, it is hard to assume that plan sponsors know whether investment consultants add value or not. While fund managers testify to the rigor with which investment consultants scrutinize their performance, and measure the effectiveness of their decisions, investment consultants themselves are shy of disclosing the sort of information which would allow plan sponsors, or any outsider, to measure *their own* performance. Among the consultants whose aggregate recommendations we have analyzed, some will do better than others, and a knowledge of differential performance would inform a plan sponsor's decision about which consultant to appoint. Without this knowledge, plan sponsors are making appointments partly uninformed, and some may be naïve about the actual utility of consultants' recommendations. An obvious policy response by regulators, or a market response by plan sponsors, is to require full disclosure of consultants' past recommendations so that such decisions are better informed and, as a consequence, their assets more efficiently allocated.

We have analyzed the influence and performance of consultants in one key area, the recommendation of investment products, and found no evidence that their recommendations add value. Recent attention from regulators and the press has focused not only on the performance of investment consultants but also on the conflicts of interest to which they are exposed. An opportunity for further research, which is beyond the scope of this paper, is to explore whether conflicts of interest induce consultants to recommend products that fail to perform; or, more generally, whether investment consultants tend to recommend an active U.S. equity strategy, which involves more complexity, more switching, and more consultancy work, rather than suggesting passive products where the consultants' role is minimal.

References

AICIO, 2013, Outsourced Chief Investment Office Survey, February, 45-57.

Altınkılıç, Oya, and Robert S. Hansen, 2009, On the information role of stock recommendation revisions, *Journal of Accounting and Economics* 48 (1), 17–36.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman, 2001, Can investors profit from the prophets? Security analyst recommendations and stock returns, *Journal of Finance* 56, 531–563.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, Out of Sight, Out of Mind: The Effects of expenses on Mutual Fund Flows, *The Journal of Business* 78 (6), 2095–2120.

Bergstresser, Daniel, John M. R. Chalmers, and Peter Tufano, 2009, Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry, *Review of Financial Studies* 22, 4129 – 4156.

Blake, Christopher and Matthew Morey, 2000, Morningstar ratings and mutual fund performance, *Journal of Financial and Quantitative Analysis* 35, 451–483.

Busse, Jeffrey, Amit Goyal and Sunil Wahal, 2010, Performance and persistence in institutional investment management, *Journal of Finance* 65, 765–790.

Cameron, A. Colin, and Pravin K. Trivedi, 1986, Econometric models based on count data. Comparisons and applications of some estimators and tests, *Journal of Applied Econometrics* 1(1), 29–53.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chen, Joseph, Harrison Hong, Ming Huang and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Coggin, T. Daniel, Frank Fabozzi, and Shafiqur Rahman, 1993, The investment performance of U.S. equity pension fund managers: An empirical investigation, *Journal of Finance* 48, 1039–1055.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2012, Should benchmark indices have alpha? Revisiting performance evaluation. NBER Working Paper No. 18050.

Dahlquist, Magnus, Jose Vicente Martinez and Paul Söderlind, 2011, Individual investor activity and performance. Working paper.

Del Guercio, Diane, and Paula Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds, *Journal of Financial and Quantitative Analysis* 37, 523–557.

diBartolomeo, Dan, and Erik Witkowski, 1997, Mutual fund misclassification: Evidence based on style analysis, *Financial Analysts Journal* 53, 32–43.

Driscoll, John C. and Aart C. Kraay, 1998, Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data, *The Review of Economics and Statistics*, MIT Press, vol. 80(4), pages 549–560, November.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds *Journal of Financial Economics* 33, 3–56.

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

Fung, William, and David A. Hsieh, 2000, Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, *Journal of Financial and Quantitative Analysis* 35, 291–307.

Gennaioli, Nicola, Andrei Shleifer and Robert Vishny, 2013, Money Doctors. NBER Working Paper No. 18174.

Goyal, Amit, and Sunil Wahal, 2008, The selection and termination of investment management firms by plan sponsors, *Journal of Finance* 63, 1805–1847.

Hoechle, D., Schmid, M., Zimmermann, H., 2009, A generalization of the calendar time portfolio approach and the performance of private investors. Working Paper University of Basel.

Inderst, Roman and Marco Ottaviani, 2012, Financial Advice, *Journal of Economic Literature* 50(2), 494-512.

Ippolito, R., 1992, Consumer reaction to measures of poor quality, *Journal of Law and Economics* 35, 45–70.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov, 2010, Do hot hands persist among hedge fund managers? An empirical evaluation, *Journal of Finance* 65, 217–255.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan Krische, and Charles Lee, 2004, Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59, 1083–1124.

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

Jones, Howard, and Jose Martinez, 2013, How Institutional Investors Form and Ignore Their Own Expectations. Available at SSRN 2252122.

Kelleher, Ellen, 2013, New York public pensions investigated, *Financial Times*, 10 November.

Khorana, Ajay and Edward Nelling, 1998, The determinants and predictive ability of mutual fund ratings, *Journal of Investing* 7, 61–66.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1992, The structure and performance of the money management industry, Brookings Papers: Microeconomics, 339–391.

Loughran, Tim, and Jay R. Ritter, 2000, Uniformly least powerful tests of market efficiency, *Journal of Financial Economics* 55(3), 361-389.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pensions and Investments, 2011, Consultant Directory, November 28.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, pages 435–480.

Rogers, W., 1993, Regression standard errors in clustered samples, *Stata Technical Bulletin*, 13, 19–23.

Securities and Exchange Commission, 2005, *Staff Report Concerning Examinations of Select Pension Consultants*, The Office of Compliance Inspections and Examinations, U.S. Securities and Exchange Commission, May 16.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Siedle, Edward, 2013, Rhode Island State Pension Admits History Of 'Pay-To-Play' And SEC Inquiry, *Forbes Investing*, 29 April.

Sirri, Erik, and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

U.K. Law Commission, 2013, Fiduciary Responsibilities of Investment Intermediaries, Law Commission Law Commission Consultation Paper No 215.

White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity, *Econometrica* 48, 817–838.

Womack, Kent L., 1996, Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51, 137–167.

Wooldridge, Jeffrey M., 2010, Econometric analysis of cross section and panel data. Second Edition, MIT Press.

Figure I

Flow-recommendation Relation

This figure shows the results of estimating kernel weighted local linear regressions of Dollar flows (Panel A) and percentage flows (Panel B) on lagged changes in consultants' recommendations. Dollar flows are expressed in millions of Dollars. Percentage flows are computed as the ratio of Dollar flows to total assets under management at the end of the previous year. The change in consultants' recommendations is the change in the percentage of short list recommendations received by a product over the total possible. The figures are produced using the Epanechnikov kernel and a window width of 0.5 and include 90% confidence bands.

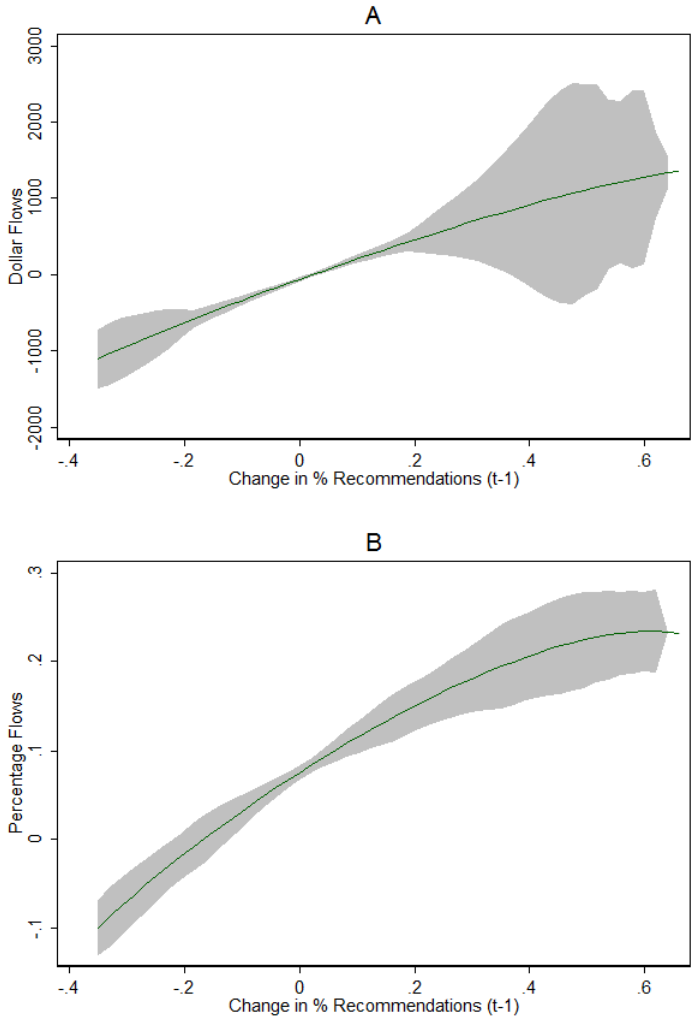

Table I
Descriptive Statistics

The table presents descriptive statistics on the sample of investment consultants and institutional investment products used in our study. Average asset size is in millions of US dollars. The market cap-style based statistics in the second part of the table are averages over the 13 year period covered by the sample (1999 to 2011).

	Number of Investment Consultants	Number of Recommendations	Number of Products			Average Product Asset Size		
			Recommended	Not Recommended	Total	Recommended	Not Recommended	Total
1999	25	459	116	849	965	7,911	560	1,871
2000	36	1398	241	856	1,097	5,624	737	2,140
2001	27	966	230	993	1,223	4,168	828	1,659
2002	32	1434	314	1,266	1,580	2,757	632	1,150
2003	30	1444	357	1,306	1,663	3,244	721	1,382
2004	30	1745	409	1,913	2,322	4,056	1,079	1,709
2005	29	1940	452	1,959	2,411	3,925	994	1,641
2006	28	2107	503	1,930	2,433	4,198	984	1,733
2007	29	2297	526	1,909	2,435	3,836	1,108	1,749
2008	30	2164	557	1,842	2,399	2,611	650	1,138
2009	29	1887	533	1,742	2,275	2,982	655	1,219
2010	27	1608	476	1,672	2,148	3,481	798	1,414
2011	28	1501	454	1,537	1,991	3,549	864	1,490
Large Cap Growth	29	437	90	345	435	6,045	927	2,185
Large Cap Value	29	455	91	315	406	5,610	1,257	2,334
Mid Cap Growth	24	150	38	121	159	2,216	513	958
Mid Cap Value	25	108	29	92	121	2,753	564	1,169
Small Cap Growth	26	160	50	204	254	1,309	483	664
Small Cap Value	27	167	51	191	242	1,519	499	746
Core	23	316	104	491	595	3,147	902	1,332

Table II
What Drives Consultants' Recommendations?

This table reports the results of pooled time-series cross-sectional Poisson and Negative Binomial regressions of the number of investment consultants' recommendations received by a product on past performance measures and variables capturing soft investment and service characteristics of the asset managers as perceived by the consultants. Soft investment factors, and service factors are expressed using the fractional rank of each asset manager in the sample. An asset manager's fractional rank, for a given variable, represents its percentile rank relative to other asset managers in the same period, and ranges from 0 to 1. Past performance measures (excess returns over benchmarks and three factor alphas) are expressed using the fractional rank of each product in its investment category. All regressions also include a lagged measure of return volatility and the log of the number of consultants surveyed each year (with a coefficient constrained to one). Each column represents a separate regression. z-scores based on standard errors clustered at the product level are included in parenthesis. The second part of the table displays model implied average marginal effects and the bottom part shows the squared correlation between observed recommendations and model predicted ones together with the number of observations in each regression.

	I	II	III	IV	V	VI	VII	VIII
	(Poisson)	(Poisson)	(Poisson)	(Poisson)	(NB)	(NB)	(NB)	(NB)
Soft Investment Factors (t)	2.61	2.62			2.72	2.72		
	(12.78)***	(12.78)***			(15.64)***	(15.65)***		
- Consistent Inv. Philosophy (t)			1.38	1.39			1.41	1.42
			(9.85)***	(9.95)***			(9.47)***	(9.56)***
- Clear Decision Making (t)			0.07	0.08			0.12	0.12
			(0.52)	(0.60)			(0.81)	(0.84)
- Capable Inv. Professionals (t)			1.17	1.15			1.22	1.20
			(8.96)***	(8.79)***			(8.88)***	(8.73)***
Service Factors (t)	0.51	0.50			0.55	0.54		
	(2.58)**	(2.55)**			(3.36)***	(3.32)***		
- Relationship Manager (t)			0.23	0.23			0.30	0.30
			(2.32)**	(2.32)**			(3.03)***	(3.03)***
- Useful Reports (t)			0.01	0.01			0.05	0.05
			(0.13)	(0.09)			(0.56)	(0.52)
- Presentation to Consultants (t)			0.52	0.52			0.42	0.42
			(4.58)***	(4.60)***			(3.52)***	(3.51)***
Past Performance Rank - Return (t)	0.23		0.20		0.16		0.14	
	(3.13)***		(2.80)***		(2.18)**		(1.96)*	
Past Performance Rank - Alpha (t)		0.25		0.21		0.18		0.16
		(3.06)***		(3.14)***		(2.29)**		(2.35)**
Return Volatility (t-1)	0.29	0.19	0.17	0.09	0.85	0.81	0.72	0.68
	(0.60)	(0.40)	(0.46)	(0.24)	(1.79)*	(1.71)*	(1.92)*	(1.82)*
Constant	-4.57	-4.57	-4.67	-4.67	-4.68	-4.68	-4.79	-4.79
	(-46.48)***	(-45.97)***	(-65.58)***	(-67.37)***	(-48.12)***	(-47.74)***	(-66.34)***	(-67.25)***
				Average Marginal Effects				
Soft Investment Factors (t)	6.04***	6.06***			6.34***	6.34***		
- Consistent Inv. Philosophy (t)			3.20***	3.23***			3.30***	3.32***
- Clear Decision Making (t)			0.16	0.18			0.27	0.28
- Capable Inv. Professionals (t)			2.71***	2.66***			2.84***	2.80***
Service Factors (t)	1.18**	1.16**			1.27***	1.26***		
- Relationship Manager (t)			0.53**	0.53**			0.69***	0.70***
- Useful Reports (t)			0.03	0.02			0.12	0.11
- Presentation to Consultants (t)			1.19***	1.20***			0.98***	0.98***
Past Performance Rank - Return (t)	0.53***		0.45***		0.37**		0.32*	
Past Performance Rank - Alpha (t)		0.57***		0.49***		0.43**		0.38**
Return Volatility (t-1)	0.68	0.44	0.39	0.21	1.98*	1.88*	1.67*	1.59*
Squared Corr $(y; \hat{y})$	0.31	0.31	0.33	0.33	0.30	0.30	0.33	0.33
Number of observations	4,238	4,238	4,230	4,230	4,238	4,238	4,230	4,230

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table III
Regressions of Asset Flows on Past Consultants' Recommendations

This table reports the results of pooled time-series cross-sectional regressions of yearly Dollar and percentage asset flows on past consultants' recommendations changes (and levels). Dollar flows are expressed in millions of Dollars. Percentage flows are computed as the ratio of Dollar flows to total assets under management at the end of the previous year. Each column represents a separate regression. All regressions also include lagged measures of the fractional performance rank of the investment products in the sample, lagged assets under management (log assets in the percentage flow regressions), lagged return volatility, an intercept and a full set of time dummies (which are not reported in the table). The change in consultants' recommendations is the change in the percentage of short list recommendations received by a product over the total possible. A portfolio's fractional rank represents its percentile performance relative to other equity funds in the same category and period, and ranges from 0 to 1. Fractional ranks are defined on the basis of a fund's one-year excess returns over its benchmark and three factor alphas. t-statistics based on standard errors clustered at the product level are included in parenthesis.

	Dollar Flow			Percentage Flows		
	I	II	III	IV	V	VI
Chg. in Recommendations (t-1)	2,403.90	2,509.54		0.29	0.24	
	(2.75)***	(2.66)***		(4.35)***	(3.53)***	
Chg. in Recs * I (Chg. >0) (t-1)			1,987.42			0.28
			(1.38)			(2.61)***
Chg. In Recs * I (Chg. ≤0) (t-1)			3,023.69			0.29
			(2.85)***			(2.38)**
Recommendations (t-1)		-190.57			0.08	
		(-0.24)			(1.47)	
Performance Rank - Return (t-1)	438.23	437.66	436.65	0.24	0.24	0.24
	(7.18)***	(7.11)***	(7.12)***	(12.04)***	(12.05)***	(12.05)***
Performance Rank - Alpha (t-1)	32.91	32.30	32.19	0.11	0.11	0.11
	(0.52)	(0.51)	(0.51)	(5.14)***	(5.15)***	(5.14)***
Total Net Assets (t-1)	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
	(-3.33)***	(-2.80)***	(-3.10)***	(-16.18)***	(-14.91)***	(-15.32)***
Return Volatility (t-1)	-2,151.19	-2,122.88	-2,107.92	-1.08	-1.08	-1.08
	(-4.42)***	(-4.17)***	(-4.03)***	(-5.94)***	(-6.00)***	(-5.93)***
Year Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.08	0.08	0.08	0.12	0.12	0.12
Number of observations	9,094	9,094	9,094	9,094	9,094	9,094

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table IV
Performance of Recommended and Not Recommended Products

The table shows the performance of portfolios of US equity actively managed products recommended by the investment advisors in our sample during the 1999 to 2011 period. Portfolios analyzed are the portfolio of all recommended products, the portfolio of the top 50% most recommended products and the portfolio of the bottom 50% recommended products. The table also shows the performance of institutional products not recommended by any of the advisors. Performance is measured using raw returns, returns in excess of a benchmark chosen to match the product style and market capitalization, and one, three and four factor alphas (corresponding to CAPM, the Fama-French three factor model and Fama-French-Carhart model). Excess returns and alphas are express in % per year. These statistics are computed on monthly returns and annualized by multiplying returns and alphas times twelve. All reported figures are gross of fees. The first part of the table shows the results for equally weighted portfolios of products whereas the second part of the table shows the same statistics for portfolios of products weighted using total net assets at the end of the previous year. t-statistics based on standard errors, robust to conditional heteroscedasticity and serial correlation of up to two lags as in Newey and West (1987), are reported in parentheses.

		Avg. Returns	Avg. Excess Ret. over Benchmark	One Factor Alpha	Three Factor Alpha	Four Factor Alpha
Equally Weighted	Recommended Products	7.13% (1.40)	1.25% (2.14)**	2.43% (2.63)***	1.14% (1.42)	1.14% (1.36)
	Top Recommended Products	6.82% (1.30)	1.15% (1.97)**	2.13% (2.21)**	0.90% (1.15)	0.91% (1.13)
	Bottom Recommended Products	8.58% (1.58)	1.70% (2.70)***	3.85% (3.07)***	2.31% (2.22)**	2.25% (2.12)**
	Not Recommended Products	8.13% (1.59)	2.35% (3.19)***	3.52% (3.30)***	2.00% (2.33)**	2.00% (2.33)**
	Recommended - Not Recommended Products	-1.00% (2.01)**	-1.10% (-3.03)***	-1.09% (2.49)**	-0.85% (2.31)**	-0.86% (2.33)**
Value Weighted	Recommended Products	4.90% (0.92)	0.96% (1.26)	0.18% (0.22)	0.39% (0.48)	0.39% (0.48)
	Top Recommended Products	4.88% (0.91)	0.98% (1.24)	0.16% (0.19)	0.39% (0.46)	0.38% (0.46)
	Bottom Recommended Products	5.68% (1.10)	1.22% (1.50)	1.03% (1.33)	0.55% (0.65)	0.60% (0.71)
	Not Recommended Products	5.16% (1.02)	0.57% (0.73)	0.55% (0.55)	-0.32% (-0.41)	-0.23% (-0.31)
	Recommended - Not Recommended Products	-0.26% (-0.20)	0.40% (0.51)	-0.37% (-0.29)	0.72% (0.73)	0.62% (0.68)

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table V
Performance Difference by Market Capitalization and Style Category

The table shows, for each market capitalization and style category in our sample, the difference in performance between products recommended by the investment advisors and all other products during the 1999 to 2011 period. Performance is measured using raw returns, returns in excess of a benchmark chosen to match the product style and market capitalization, and one, three and four factor alphas (corresponding to CAPM, the Fama-French three factor model and Fama-French-Carhart model). Excess returns and alphas are express in % per year. These statistics are computed on monthly returns and annualized by multiplying monthly returns and alphas times twelve. The first part of the table shows the results for equally weighted portfolios of products whereas the second part of the table shows the same statistics for portfolios of products weighted using total net assets at the end of the previous year. t-statistics based on standard errors, robust to conditional heteroscedasticity and serial correlation of up to two lags as in Newey and West (1987), are reported in parentheses.

		Recommended - Not Recommended Products Performance				
		Avg. Returns	Avg. Excess Ret. over Benchmark	One Factor Alpha	Three Factor Alpha	Four Factor Alpha
Equally Weighted	Large Cap Growth	-1.40% (1.83)*	-1.40% (1.83)*	-1.59% (-2.47)**	-0.92% (-2.69)***	-0.87% (-2.58)**
	Large Cap Value	-0.53% (-0.77)	-0.53% (-0.77)	-0.64% (-1.01)	-0.93% (-1.94)*	-0.79% (-1.97)*
	Mid Cap Growth	-0.41% (-0.84)	-0.41% (-0.84)	-0.45% (-0.93)	-0.23% (-0.59)	-0.19% (-0.50)
	Mid Cap Value	-0.13% (-0.31)	-0.13% (-0.31)	-0.19% (-0.46)	-0.09% (-0.20)	-0.07% (-0.15)
	Small Cap Growth	-1.34% (-1.35)	-1.34% (-1.35)	-1.47% (-1.51)	-1.45% (-1.72)*	-1.65% (-1.85)*
	Small Cap Value	-0.34% (-0.85)	-0.34% (-0.85)	-0.37% (-0.93)	-0.76% (-2.11)**	-0.80% (-2.18)**
	Core	-0.27% (-0.72)	-0.45% (-1.18)	-0.42% (-1.55)	-0.40% (-1.64)	-0.41% (-1.70)*
Value Weighted	Large Cap Growth	0.20% (0.15)	0.20% (0.15)	0.06% (0.05)	0.72% (0.72)	0.55% (0.58)
	Large Cap Value	0.71% (0.79)	0.71% (0.79)	0.70% (0.77)	0.19% (0.23)	0.29% (0.38)
	Mid Cap Growth	1.10% (1.93)*	1.10% (1.93)*	1.08% (1.92)*	1.44% (2.78)***	1.50% (2.95)***
	Mid Cap Value	-0.66% (-0.83)	-0.66% (-0.83)	-0.68% (-0.86)	-0.55% (-0.74)	-0.48% (-0.73)
	Small Cap Growth	0.97% (1.27)	0.97% (1.27)	0.95% (1.24)	1.31% (1.64)	1.31% (1.64)
	Small Cap Value	-0.41% (-0.52)	-0.41% (-0.52)	-0.40% (-0.50)	-1.13% (-1.46)	-1.16% (-1.50)
	Core	0.26% (0.41)	0.27% (0.45)	0.41% (0.76)	0.47% (0.91)	0.42% (1.01)

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table VI
Backfill Bias in Reported Returns

The table shows the performance of portfolios of recommended and not recommended products as reported and eliminating the first one to four years of reported history for each product. Performance is measured using raw retirms, returns in excess of a benchmark chosen to match the product style and market capitalization, and one, three and four factor alphas (corresponding to CAPM, the Fama-French three factor model and Fama-French-Carhart model). Excess returns and alphas are express in % per year. These statistics are computed on monthly returns and annualized by multiplying monthly excess returns and alphas times twelve. All reported figures are gross of fees. The first part of the table shows the results for equally weighted portfolios of products whereas the second part of the table shows the same statistics for portfolios of products weighted using total net assets at the end of the previous year. t-statistics based on standard errors, robust to conditional heteroscedasticity and serial correlation of up to two lags as in Newey and West (1987), are reported in parentheses.

		Avg. Returns	Avg. Excess Ret. over Benchmark	One Factor Alpha	Three Factor Alpha	Four Factor Alpha
Equally Weighted	Recommended Products	7.13% (1.40)	1.25% (2.14)**	2.43% (2.63)***	1.14% (1.42)	1.14% (1.36)
	Rec. Products (with 1Y Backfill correction)	7.06% (1.33)	1.16% (2.06)**	2.35% (2.53)**	1.09% (1.42)	1.08% (1.36)
	Rec. Products (with 2Y Backfill correction)	6.94% (1.31)	1.08% (1.93)*	2.24% (2.40)**	0.98% (1.27)	0.98% (1.22)
	Rec. Products (with 3Y Backfill correction)	6.90% (1.31)	1.08% (1.91)*	2.21% (2.36)**	0.99% (1.26)	0.99% (1.23)
	Not Recommended Products	8.13% (1.59)	2.35% (3.19)***	3.52% (3.30)***	2.00% (2.33)**	2.00% (2.33)**
	Not Rec. Products (with 1Y Backfill correction)	7.91% (1.55)	2.14% (3.02)***	3.31% (3.16)***	1.87% (2.29)**	1.82% (2.16)**
	Not Rec. Products (with 2Y Backfill correction)	7.82% (1.54)	2.06% (2.94)***	3.22% (3.08)***	1.81% (2.21)**	1.76% (2.09)**
	Not Rec. Products (with 3Y Backfill correction)	7.78% (1.53)	1.98% (2.86)***	3.18% (3.02)***	1.79% (2.17)**	1.74% (2.04)**
Value Weighted	Recommended Products	4.90% (0.92)	0.96% (1.26)	0.18% (0.22)	0.39% (0.48)	0.39% (0.48)
	Rec. Products (with 1Y Backfill correction)	4.92% (0.92)	0.97% (1.27)	0.20% (0.24)	0.40% (0.48)	0.39% (0.48)
	Rec. Products (with 2Y Backfill correction)	4.94% (0.93)	1.00% (1.28)	0.22% (0.28)	0.44% (0.53)	0.44% (0.53)
	Rec. Products (with 3Y Backfill correction)	4.95% (0.93)	1.01% (1.29)	0.23% (0.29)	0.45% (0.54)	0.44% (0.53)
	Not Recommended Products	5.16% (1.02)	0.57% (0.73)	0.55% (0.55)	-0.32% (-0.41)	-0.23% (-0.31)
	Not Rec. Products (with 1Y Backfill correction)	5.15% (1.01)	0.55% (0.71)	0.54% (1.54)	-0.33% (-0.41)	-0.24% (-0.32)
	Not Rec. Products (with 2Y Backfill correction)	5.17% (1.02)	0.56% (0.71)	0.56% (0.57)	-0.31% (-0.39)	-0.21% (-0.29)
	Not Rec. Products (with 3Y Backfill correction)	5.23% (1.03)	0.58% (0.76)	0.62% (0.62)	-0.21% (-0.26)	-0.11% (-0.14)

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table VII
Product Size and Performance

This table reports the coefficient estimates and t-statistics (in parentheses) from pooled OLS regressions with Driscoll-Kraay standard errors. The standard error estimates are heteroscedasticity consistent and robust to both cross-sectional dependence and autocorrelation. In systems I, II and III, the dependent variable is the monthly gross return of the investment product in excess of a benchmark chosen to match the product style and market capitalization. In systems IV, V and VI the dependent variable is the monthly gross excess return of the investment product over the risk free rate and the explanatory variables are obtained by aid of a Kronecker expansion between the factors of a Carhart (1997) like performance measurement model on the one hand, and a recommendation dummy and the natural logarithm of product assets at the end of the previous year on the other. For brevity, the table does not present the estimation results for the factors or the interaction terms between product characteristics (including the recommendation dummy) and factors. We provide p-values of Wald tests of the difference in annualized returns between recommended products and non-recommended products in square brackets.

	I	**II**	**III**	**IV**	**V**	**VI**
	(B. Adj.)	(B. Adj.)	(B. Adj.)	(FFC)	(FFC)	(FFC)
Constant	0.00144	0.00071	0.00064	0.00134	0.00073	0.00063
	(2.98)***	(1.72)*	(1.58)	(2.26)**	(1.44)	(1.30)
Recommended	-0.00073	-0.00029	0.00005	-0.00030	-0.00021	0.00027
	(-2.65)***	(-1.11)	(0.23)	(-1.47)	(-1.17)	(1.18)
Log AUM			-0.00015			-0.00022
			(-3.62)***			(-3.83)***
Factors	No	No	No	4 Factors	4 Factors	4 Factors
Interactions	No	No	No	Yes	Yes	Yes
Sample	Full	Restricted	Restricted	Full	Restricted	Restricted
Observations	225,726	157,633	157,633	225,726	157,633	157,633
Groups	2,146	2,051	2,051	2,146	2,051	2,051
R-squared	0.00	0.00	0.00	0.77	0.84	0.84
			Annualized Differences in Performance			
Recommended - Not	-0.88%	-0.35%	0.06%	-0.36%	-0.25%	0.32%
Recommended	[0.00]***	[0.26]	[0.79]	[0.14]	[0.24]	[0.23]

***, **, * Statistically significant at the 1%, 5%, and 10% levels, respectively.

Table VIII
Net Change in Number of Recommendations and Performance

The table shows the performance of portfolios of US equity actively managed products that experience a net increase (decrease) in the number of recommendations in the twelve or twenty-four month period following the recommendation change. Performance is measured using raw returns, returns in excess of a benchmark chosen to match the product style and market capitalization, and one, three and four factor alphas (corresponding to CAPM, the Fama-French three factor model and Fama-French-Carhart model). Excess returns and alphas are expressed in % per year. All reported figures are gross of fees. The first part of the table shows the results for equally weighted portfolios of products whereas the second part of the table shows the same statistics for portfolios of products weighted using total net assets at the end of the previous year. t-statistics based on standard errors, robust to conditional heteroscedasticity and serial correlation of up to two lags as in Newey and West (1987), are reported in parentheses.

		12 Month Period Following Addition/Deletion				
		Avg. Returns	Avg. Excess Ret. over Benchmark	One Factor Alpha	Three Factor Alpha	Four Factor Alpha
Equally Weighted	Increase in Number of Recommendations	5.34% (0.95)	0.62% (1.02)	2.26% (2.31)**	0.99% (1.25)	0.94% (1.19)
	Decrease in Number of Recommendations	6.58% (1.20)	1.19% (2.13)**	3.55% (2.34)**	1.48% (1.21)	1.54% (1.32)
	Difference	-1.24% (-0.86)	-0.57% (-0.80)	-1.29% (-0.89)	-0.49% (-0.49)	-0.59% (-0.69)
Value Weighted	Increase in Number of Recommendations	2.12% (0.36)	-0.35% (-0.24)	-0.98% (-0.56)	-0.22% (-0.19)	-0.30% (-0.29)
	Decrease in Number of Recommendations	4.62% (0.90)	0.54% (0.75)	1.63% (1.09)	0.74% (0.67)	0.81% (0.78)
	Difference	-2.51% (-0.83)	-0.89% (-0.54)	-2.61% (-0.88)	-0.97% (-0.50)	-1.12% (-0.65)
		24 Month Period Following Addition/Deletion				
		Avg. Returns	Avg. Excess Ret. over Benchmark	One Factor Alpha	Three Factor Alpha	Four Factor Alpha
Equally Weighted	Increase in Number of Recommendations	5.27% (0.94)	0.55% (0.90)	1.92% (2.17)**	0.96% (1.24)	0.94% (1.21)
	Decrease in Number of Recommendations	6.86% (1.26)	1.33% (2.40)**	3.57% (2.46)**	1.78% (1.44)	1.81% (1.50)
	Difference	-1.59% (-1.16)	-0.78% (-1.17)	-1.65% (-1.22)	-0.82% (-0.89)	-0.87% (-1.05)
Value Weighted	Increase in Number of Recommendations	1.64% (0.28)	-0.58% (-0.40)	-1.74% (-1.03)	-0.74% (-0.64)	-0.77% (-0.73)
	Decrease in Number of Recommendations	4.93% (0.95)	0.50% (0.72)	1.69% (1.18)	1.11% (1.06)	1.15% (1.15)
	Difference	-3.29% (-1.14)	-1.07% (-0.68)	-3.43% (-1.21)	-1.85% (-1.01)	-1.92% (-1.16)

***, **, * Statistically significant at 1%, 5% and 10% levels respectively

Table A1
Investment Products and Benchmarks

Investment Product	Benchmark
Large Cap Growth	Russell 1000 Growth
Large Cap Value	Russell 1000 Value
Mid Cap Growth	Russell Midcap Growth
Mid Cap Value	Russell Midcap Value
Small Cap Growth	Russell 2000 Growth
Small Cap Value	Russell 2000 Value
Domestic Equity Core - All Cap	Russell 3000
Domestic Equity Core - Large Cap	Russell 1000
Domestic Equity Core - Mid Cap	Russell Midcap