1-2020

Sell-side analysts' benchmarks

Ohad KADAN
Washington University in St. Louis

Leonardo MADUREIRA
Case Western Reserve University

Rong WANG
Singapore Management University, rongwang@smu.edu.sg

Tzachi ZACH
Ohio State University - Main Campus

Published in Accounting Review, January 2020, Vol. 95, No. 1, pp. 211-232.
https://doi.org/10.2308/accr-52446

Accepted version

Sell-Side Analysts' Benchmarks

Ohad Kadan
Washington University in St. Louis
kadan@wustl.edu

Leonardo Madureira
Case Western Reserve University
leonardo.madureira@case.edu

Rong Wang
Singapore Management University
rongwang@smu.edu.sg

Tzachi Zach
Ohio State University
Zach.7@osu.edu

March 2019

We thank two anonymous reviewers, David Aboody, Mark Bradshaw (the editor), Roby Lehavy, Joshua Livnat, Roni Michaely, Brian Rountree, Kent Womack, Pei-Fang Hsieh, Roger Loh, Jialing Yu, Xiaoyun Yu, seminar participants at Case Western Reserve University, New York University, Washington University in St. Louis, Ohio State University, and conference audience at the 2012 IDC Summer Conference, the 2012 FARS Conference, the 2012 FIRS Conference, the 2012 Asian Finance Conference, the 2012 FMA Conference, the 2013 AFA Conference, the 2013 Financial Economics and Accounting Conference, the 2013 CICF conference, the 2013 SMU summer camp, the 2013 EFMA, the 2014 WFA Conference, and the 2014 EFA Conference for helpful comment and suggestions.

Sell-Side Analysts' Benchmarks

ABSTRACT

Sell-side analysts employ different benchmarks when defining their recommendations. A buy for some brokers means the stock is expected to outperform its industry, while for other brokers it means the stock is expected to outperform the market, or some return threshold. We show that these stated benchmarks have implications for the distribution of recommendations, price reactions to recommendations, and the investment value of recommendations. We conclude that, depending on the question, academics may need to account for the benchmarks when studying analysts' outputs, and investors may find the benchmarks beneficial in interpreting analysts' advice.

Keywords: Analysts; Benchmarks; Recommendations.

JEL Classifications: G10, G24

I. INTRODUCTION

A voluminous literature examines sell-side analysts' stock recommendations: how they are created, how they relate to other analysts' outputs, and whether they are useful to investors. However, the literal meaning of recommendations (buy, hold or sell), and its relation to the actual investment advice (to buy, hold, or sell a stock), have largely been implicitly assumed uniform across different brokers and analysts. Such an assumption may be an oversimplification. Inspections of the disclosures in which brokers describe the meaning of their recommendations reveal that different brokers assign different meanings to their recommendations and refer to different benchmarks when issuing them. For example, a buy at one broker means the stock is expected to outperform its industry peers (we call such a broker an "industry benchmarker"); at another a buy means the stock is expected to outperform the market ("market benchmarker"); and at yet another, a buy means that the stock is expected to earn a return above some pre-determined threshold such as ten percent ("total benchmarker").

The different literal meanings suggested by these benchmarks may carry implications on how the recommendations should be interpreted. For example, at face value, a buy recommendation from an industry benchmarker is a direct and clear signal of the within-industry prospects of the firm, as opposed to a buy recommendation from a market benchmarker, which carries with it information on both the firm and its industry. As a result, recommendations from industry benchmarkers may be better predictors of a firm's industry-adjusted returns. From an investment perspective, the ability to discern winners and losers in an industry from a set of recommendations may depend on the benchmark adopted by the broker. Thus, ignoring the benchmarks may hinder the ability to fully capture the information conveyed by recommendations. Accordingly, the objective of this paper is to examine empirically whether the various benchmarks used by different brokerage houses have implications for the information provided by analysts' stock recommendations. Specifically, we study the impact of different benchmarks on the distribution of recommendations, price reactions to recommendations, and the investment value of recommendations.

Large-scale data on the benchmarks has not been readily available. Before 2002, brokers were not required to disclose the detailed meaning of their recommendations. While nowadays a recommendation report must document the meaning of the recommendation advice, aggregators and disseminators of recommendation data (Dow Jones Newswires, Bloomberg, Finance Yahoo, etc.) still ignore such fine print.[1] For academics, the usual databases they rely upon (IBES and FirstCall) do not carry detailed information about the literal meaning of recommendations.

It can be argued that the benchmarks *should* be ignored if they do not affect the way recommendations are generated. This would be the case if benchmarks are irrelevant, non-binding, statements that in reality do not drive the way analysts work and have no implications for the way investors should interpret the recommendations. It is ultimately an empirical question whether and how the benchmarks are relevant for interpreting recommendations, and how they interact with other analysts' outputs.

Beginning in September of 2002, and following Rule NASD 2711, Rule NYSE 472, and the Global Settlement, which attempted to mitigate conflicts of interests in analysts' research, brokers are required to define in each report the literal meaning of their recommendations. This definition must include the benchmark used when interpreting the recommendation advice. To examine our research questions we hand-collect the meaning of recommendations for 195 brokers accounting for over 95 percent of all recommendations issued during our sample period

---

[1] An exception is the investment website MarketWatch (http://www.marketwatch.com/tools/guide.asp), which provides its readers with a detailed guide, including benchmarks, for properly interpreting recommendations issued by different brokers.

(September 2002-December 2014). We find that the most prevalent benchmarks adopted by brokers are the industry (accounting for 24 percent of the recommendations), market (15 percent of recommendations), and total benchmarks (20 percent of the recommendations). Other brokers typically use either combinations or risk-adjusted versions of these three benchmarks. Given their popularity and the simplicity of their meaning, we focus our empirical analysis on brokers employing these three benchmarks exclusively.

In our first analysis we test whether the distribution of stock recommendations issued by analysts is related to the benchmarks adopted by their brokerage house. We find that an analyst's tendency to issue more bullish recommendations is strongly associated with the benchmark. Specifically, the distribution of recommendations issued by industry benchmarkers is less bullish as compared to the distribution of recommendations issued by market and total benchmarkers. These patterns do not necessarily imply an inherent difference in optimism across analysts employing different benchmarks—given that each benchmark carries a different meaning for the recommendation advice. Nevertheless, they suggest that, when studying the optimism of analysts through their recommendations, one should account for the heterogeneity of the recommendation distribution by controlling for the benchmark type.

In our next analysis, we explore price reactions to recommendations. In our setup it is important to distinguish among different types of price changes as they relate to the benchmarks. When considering recommendations issued by industry benchmarkers it is natural to study industry-adjusted returns. Similarly, when considering recommendations issued by market (total) benchmarkers it is natural to study market-adjusted (unadjusted) returns. Following this approach we find that industry-adjusted price reactions to recommendations are more prominent for recommendations issued by industry benchmarkers compared to recommendation issued by market and total benchmarkers. By contrast, we do not find evidence that recommendations issued by market benchmarkers garner significantly different market-adjusted price reactions. Interestingly, unadjusted price reactions are significantly smaller in magnitude for recommendations issued by total benchmarkers.

These results are consistent with the flavor of the complexity argument introduced by Clement (1999). Clement argues that analysts facing more complex tasks issue less accurate forecasts. In similar vein, complexity may impair the quality of a recommendation signal, and thus investors should react most strongly to recommendations issued by analysts dealing with the least complex task. In the context of recommendations, total benchmarkers face the most challenging task, as they are required to provide insights on the expected performance of a given stock, its industry, and the market as a whole (as these are all embedded in the estimation of raw returns). Market and industry benchmarkers face less complex tasks as they are only required to provide insights on expected returns of a given industry and/or a given stock relative to its industry.

Each benchmark is associated with a specific objective for a recommendation. Industry benchmarkers' buy recommendations are aiming at beating industry peers; for market benchmarkers the objective is beating the market; and the objective of total benchmarkers is to beat an absolute return threshold. To evaluate analysts' ability to achieve their stated objectives and whether achieving the objectives varies by benchmark, we collect for each broker the target return associated with its benchmark. For example, a target return for a buy recommendation issued by an industry (market) benchmarker specifies by how much the recommended firm is expected to beat the industry (market). Similarly, a target return for a buy recommendation issued by a total benchmarker specifies an absolute return such as ten percent.[2] We then examine whether and by how much the return of a recommended firm meets or beats its stated objective.

[2] The stated objective of a recommendation thus includes both a benchmark *category* (whether the firm's stock is expected to beat industry peers, the market, or an absolute return) and a benchmark *target* (how much the

Not surprisingly, we find that industry and market benchmarkers are more successful in meeting their objectives compared to total benchmarkers. For example, about 51 percent of buy recommendations issued by industry or market benchmarkers meet or beat their objective, compared to 44 percent of buy recommendations issued by total benchmarkers. These results reflect the fact that meeting the objective for total benchmarkers is a relatively more difficult task. Indeed, total benchmarkers are expected to predict firm-specific returns, industry returns, and market returns, and they tend to adopt the most stringent target thresholds.

When comparing a recommendation return with its stated objective, we are evaluating the analyst performance based on the literal meaning of her recommendation advice. But performance of a recommendation may simply derive from the risk profile of the recommended stock. To account for such risk, we use a propensity score methodology matching each actual recommendation to a control unit (some other firm at another point in time) with a similar risk profile. We find that for all types of benchmarks, firms for which analysts issue buy (sell) recommendations perform better (worse) than firms with similar risk characteristics that did not receive such recommendations. Thus, our results show that, regardless of the benchmarks, analysts perform better than a "naïve" strategy that simply picks stocks based on their risk characteristics.

In our final analysis we study how the different benchmarks affect the value of analysts' advice in identifying winners and losers within a particular industry (stock picking). The benchmarks can either help or hinder the stock picking ability of analysts. Indeed, the complexity hypothesis (discussed above) suggests that the stock-picking skills of market and total benchmarkers would be hindered as compared to industry benchmarkers. Furthermore, the benchmarks can affect the ability of analysts to convey their stock-picking opinion through their

---

recommended firm is expected to surpass its benchmark category). Throughout the paper, we employ the word benchmark to refer to a benchmark category rather than a benchmark target.

recommendations. Industry benchmarkers assert to rely on stock picking alone, so their buys and sells are expected to be direct signals of winners and losers within an industry. By contrast, recommendations issued by market and total benchmarkers do not provide as clear a signal on within-industry winners and losers. For example, a good stock in a very poor industry may be a buy for an industry benchmarker, but a hold or even a sell for a market or total benchmarker. Thus, once again, recommendations issued by industry benchmarkers are expected to provide a superior signal for picking stocks within an industry.

To test this hypothesis we use the methodology introduced in Boni and Womack (2006), which forms industry- and market-neutral portfolios such that the only source of abnormal performance is within-industry stock picking. We then calculate the Fama-French four-factor alphas of these portfolios separately for each benchmark. We find that recommendations issued by industry benchmarkers generate a statistically significant and economically large (0.44 percent per month) alpha associated with the stock picking information contained in their recommendations. By contrast, we find no significant alpha for either market or total benchmarkers. This indicates that the more elaborate definitions associated with the latter's benchmarks hinder their ability to provide useful stock picking information to investors.

This paper is the first to focus exclusively on sell-side benchmarks, for which we provide the first large-scale and comprehensive analysis. Our analysis is consistent with Bradshaw's (2012) assessment that these benchmarks are important for the study of sell-side research. We contribute to the literature by documenting how these benchmarks are related to the distribution of analysts' recommendations, studying how they affect different outputs of analysts, and investigating their relation to the investment value of stock recommendations. Kadan, Madureira, Wang, and Zach (2012) provide some preliminary discussion of the sell-side benchmarks, within their study of industry recommendations. They point out the existence of the benchmarks and use a small sample of disclosures from 20 brokers to study the relation between firm and industry recommendations. Thus, the scale and focus of these two papers are different.

II. DATA AND PRELIMINARY ANALYSIS

Data

We focus on analysts' stock recommendations and other types of outputs of all U.S. firms in the period of September 2002 to December 2014. The source for analysts' data is the IBES database. The data on firm characteristics are from COMPUSTAT. We obtain stock returns from CRSP, and equity offerings data from SDC. Industry membership is inferred through the industry classification defined by the Global Industry Classification Standard (GICS) obtained from COMPUSTAT. The GICS system is widely adopted by financial practitioners as an industry classification system, and has been increasingly used in academic studies (e.g., Bhojraj, Lee, and Oler 2003, Boni and Womack 2006; and Kadan et al. 2012). Following Kadan, Madureira, Wang, and Zach (2009), we remove the recommendations issued during the events of rating changes in 2002, as these are merely technical.

We manually collect data on the benchmarks used by brokers that issued at least 200 recommendations during our sample period. There are 374,807 recommendations issued by all brokers during our sample period for U.S. firms, out of which 355,317 are issued by brokers with at least 200 recommendations. The threshold of 200 recommendations enables us to collect benchmark data of large brokers without a significant loss of recommendation data.

Under NASD Rule 2711 and NYSE Rule 472 adopted in mid-2002, brokers are required to disclose the literal meaning of their recommendations inside their reports. We collect these disclosures from three sources. First, we retrieve information from full-text research reports in the Investext database when available. Second, for brokerage houses not appearing in Investext, we collect data from the Investars website,[3] which contains the ratings' definitions of some brokers. Finally, if necessary, we obtain data directly from brokers' websites.

< Insert Table 1 here >

Based on the disclosures, we categorize brokers into ten different types according to the benchmarks they use. Table 1 describes these benchmarks and provides general summary statistics. The three most basic benchmarks involve determining recommendations according to the expected performance of the covered stock compared to the performance of industry peers, the market, or some fixed return threshold. More formally, we classify brokers as industry benchmarkers if they state that their stock recommendations are benchmarked against industry performance. For example, Smith Barney's analysts rate stocks based on the "stock's performance vs. the analyst's industry coverage for the coming 12-18 months." We classify brokers as market benchmarkers if they state that their stock recommendations are benchmarked against market performance. For example, Wachovia's analysts rate a stock based on its expected performance "relative to the market over the next 12 months." Finally, we classify brokers as total benchmarkers if they issue recommendations based on a stock's expected total return. This is the case, for example, with Deutsche Bank, where a buy recommendation means that the stock's price is "expected to appreciate ten percent or more over a 12-month period."

Occasionally brokers determine their recommendations using some combination of these three basic benchmarks. We identify four such combinations. For example, Dougherty & Co combines features of market and industry benchmarks, so that its buy means a stock is "expected to outperform the broader market and/or its sector." We categorize this broker as a market/industry benchmarker. Other hybrids we identify are market/total, industry/total, and market/industry/total.

---

[3] http://www.investars.com

Other brokers refine the basic benchmarks by adding a risk-adjustment feature. For example, Morgan Stanley establishes its recommendations based on the "stock's total return vs. analyst's coverage on a risk-adjusted basis." Notably, the nature of the adjustment for risk is often vague. To highlight this feature, we add a new category and classify Morgan Stanley as an industry/risk benchmarker. Similarly, we classify a broker as market/risk (total/risk) when the benchmark involves comparing the stock's expected performance to the market (a total return threshold) on some type of risk-adjusted measure.

Some brokers changed their benchmarks during our sample period. For example, Merrill Lynch used a total/risk benchmark between September 2002 and May 2008, and an industry/total benchmark since June 2008. In this case, we classify Merrill Lynch as a total/risk benchmarker between September 2002 and May 2008, and as an industry/total benchmarker between June 2008 and December 2014. However, for some brokers, we failed to identify the exact date of the change. We classify such instances as a "Changes" category. Finally, some brokers could not be classified in any of the above categories, either because we could not find any data on their analysts' disclosures or because their disclosures did not fall into any of the above categories.

Table 1 reports that 38 brokers use the industry benchmark during our sample period, and the number of recommendations issued by such brokers accounts for about 24.1 percent of all recommendations. There are 52 brokers that base their recommendations on a total benchmark, and as a group they issued about 19.8 percent of all recommendations. The number of brokers relying on a market benchmark is 29, accounting for 14.8 percent of all recommendations.[4]

[4] For each broker, if available, we download research reports every year to check the consistency of the ratings definition during the sample period. If there is a change in ratings definition, we download more reports to identify the date of change. For some brokers, reports are only available in some of the years, and we then extrapolate the definition to the rest of the sample period. Similarly, when we collect ratings definitions from Investars or the broker' website, we assume that the broker uses the same ratings definition for the entire sample period. Overall, during the sample period there are 195 brokers with at least 200 recommendations, of which 20 change their benchmarks once, and one broker changed its benchmark twice. Therefore, the total number of brokers in Table 1 is 217.

Brokers using risk-adjusted benchmarks are usually large, as revealed by the average number of recommendations issued in each category (Morgan Stanley is one such case), but there are relatively few of them. Therefore, as a group they account for just under 16 percent of recommendations. Similarly, there are few brokers combining the basic benchmarks. Finally, we fail to collect data on benchmarks for 39 brokers, but these brokers are relatively small (with an average of 566 recommendations during the sample period), and as a group they issued about 6.2 percent of recommendations in our sample.

In this paper we focus our attention on the three basic benchmarks. Two reasons drive our choice. First, we want to address a set of benchmarks that is representative of the universe of brokers. Industry, market, and total benchmarkers thoroughly satisfy this requirement: Together they account for about 59 percent of the recommendations in our sample period, and they are adopted by 12 of the 20 largest brokers. Second, we need to address benchmarks that have a straightforward interpretation, so that clear testable hypotheses can be developed. This requirement again favors the three basic benchmarks, as they are the most precisely defined, particularly when compared to the risk-adjusted benchmarks (brokers adopting these risk-adjusted benchmarks do not properly document the meaning of their risk-adjustment feature) or to the benchmarks that combine more than one basic benchmark.[5] Therefore, the sample used in this paper, encompassing data from industry, market and total benchmarkers, includes 208,674 recommendations (59 percent of the overall sample).

< Insert Table 2 here >

Besides the benchmark, the recommendation's stated objective (or, its literal meaning) carries a target threshold as well. For example, in the case of a buy, a broker may adopt a target

[5] In unreported results, we conducted all the relevant analyses in the paper including separate categories for industry, market and total benchmarkers who adjust for risk. None of our conclusions is affected by this inclusion while the results for the risk-adjusted categories are mixed and inconclusive. This is perhaps due to the lack of specifics regarding the nature of the risk-adjustments used by brokers.

threshold of ten percent, such that the literal meaning of the recommendation is that the stock is expected to surpass the benchmark return by ten percent.[6] Table 2 presents summary statistics of the target thresholds used by brokers in our sample. Panel A shows the thresholds used by industry benchmarkers. The most frequent target is zero, saying that a typical buy (sell) recommendation issued by an industry benchmarker means that the recommended stock's return will exceed (be lower than) the industry return over the forecast horizon. Panel B shows that for market benchmarkers the most common threshold is also zero, corresponding to the expectation that the stock's return of a buy (sell) recommendation exceeds (is lower than) the market return over the forecast horizon. Finally, Panel C presents the threshold distribution for total benchmarkers. Here, the most prevalent thresholds for buy recommendations are ten percent and 15 percent, and for sell recommendations are zero percent, ten percent, and 15 percent.[7] Notice that brokers may use different thresholds for buy and sell recommendations, especially for total benchmarkers.

It is important to clarify that the benchmarks and their associated target thresholds are related to but different from the target prices issued by analysts. To see the relation between the two, note that given a 12-month ahead target price and a current stock price, one can derive the analyst's expected return implied by the target price (ERTP) as

$$ERTP=(TP_0 - P_{-1})/ P_{-1},$$

where TP_0 is the target price, and P_{-1} is the stock price one day before the target price is issued.[8] Thus, the target price gives rise to an implied expected return, while the target threshold associated

[6] The literal meaning of the recommendation also includes the forecast horizon: how long should it take for the recommendation prediction to materialize. In our data, the vast majority of brokers consider a 12-month horizon.
[7] Some brokers also change the target thresholds, which explains the different total number of brokers for "Target (Buy)" and "Target (Sell)" in panel C of Table 2.
[8] For simplicity, this definition of expected return inferred from target price focuses on expected price appreciation only, ignoring dividend yield. When we use ERTP in our analysis, we also consider versions that do account for the dividend yield.

with the benchmark specifies a bound on the expected return relative to some benchmark return. We further discuss the relation between the benchmarks and the target prices later in this section.

Benchmark Determinants

The analysts' disclosures demonstrate that different brokerage houses rely on different benchmarks. One obvious question is why. In fact, it may be surprising that some brokers are *not* industry benchmarkers. Using an industry benchmark fits well with the structure of research departments in brokerage houses, where analysts work in industry groups and are deemed industry specialists (e.g., Boni and Womack, 2006; Kadan et al., 2012). On the other hand, being an expert in one industry is not enough for an analyst working under a market or total benchmark. These analysts may also need to have knowledge across other industries— for them to be able to predict a stock's performance relative to the market—or even to have an overall knowledge of the market— to be able to predict a stock's raw return.

Analysts we have interviewed hinted at a tension about which benchmark should be used. Analysts working for an industry benchmarker emphasized the natural fit between being an industry specialist and using an industry benchmark. They also pointed out that ranking firms within an industry arises directly from application of techniques such as comparables. Others expressed preference towards a total benchmark, given that a total return expectation is a direct product of applying a discounted cash flow (DCF) methodology. They also argued that an expectation about total return is the most useful output from the perspective of investors. Finally, some argued that the market benchmark makes sense as well, since it is common practice to evaluate each equity asset relative to the market (or a popular index such as the S&P 500).

To add to this anecdotal evidence and provide some large sample results on the determinants of the benchmarks, we explore their possible association with brokers' characteristics. In unreported analysis, we show that two variables emerge as strong determinants of the choice of benchmark. The first is broker size, measured by the number of recommendations issued by a broker as a fraction of all recommendations issued during the year. Larger brokers are more likely to adopt an industry benchmark as opposed to either market or total benchmarks. It may be that large brokers that employ a large number of analysts can allow analysts to focus on a select group of firms in one particular industry, leading to more industry specialization and thereby to industry benchmarking. The second determinant is the number of industries covered. A larger number of covered industries is associated with a lower likelihood of industry benchmarking and a higher likelihood of adopting a market benchmark. It may be that brokers that follow many industries have a better perspective of the market, and thereby are more capable of benchmarking their recommendations to it.

Do Analysts Abide by the Benchmarks?

Before turning to our main empirical analysis of benchmarks' implications, it is interesting to ask whether analysts abide by their broker's stated benchmark. Recommendations from analysts employing an industry benchmark are statements about the analysts' expectations on how stocks will perform relative to their industry peers; that is, these analysts rely on ranking firms within the industry. By contrast, market and total benchmarkers may determine their recommendations by considering their expectations of both the intra-industry ranking of the firms and the industry performance relative to the market. We thus expect that industry benchmarkers would primarily use within-industry information about firm fundamentals, while market and total benchmarkers would also rely on across-industry information.

To examine this hypothesis we relate stock recommendations to firms' fundamentals as expressed by analysts' other outputs: earnings and long-term growth forecasts. Our analysis (unreported, available upon request) shows that all types of benchmarkers rely similarly on within-industry information, but that market and total benchmarkers indeed place more weight on across-industry information than industry benchmarkers when forming their recommendations.

We also examine the implication of adoption of different benchmarks on the relation between recommendations and target prices. For total benchmarkers, we expect a close alignment between recommendations and target prices, because both rely on forecasts of raw returns. For example, if a buy is issued and the literal meaning of the buy is that the stock should yield at least 10%, then one should see the ERTP of the contemporaneous target price to be at least 10%. However, less consistency is expected from industry and market benchmarkers.

In unreported analysis we find evidence consistent with this hypothesis: The fraction of optimistic recommendations that appear with strictly positive (negative) ERTPs is significantly higher, and the average ERTP is more positive (negative) for optimistic (pessimistic) recommendations of total benchmarkers, compared to industry and market benchmarkers.

Overall, these results are consistent with analysts indeed abiding by their benchmarks, suggesting that the benchmarks matter in the way recommendations are created.[9] In the next section we present our main results, studying whether and how benchmarks should be accounted for when interpreting analysts' outputs.

III. RESULTS

In this section, we explore how several aspects of analyst recommendations are affected by the adoption of different benchmarks. We study how the distribution of analysts' recommendations relates to the benchmarks. We then examine how investors react to recommendations based on different benchmarks. Next, we study how the benchmarks are associated with analysts' ability to

[9] One possibility is that analysts are following the benchmarks determined by their brokerage house. An alternative is that brokerage houses simply adopted benchmarks reflecting the common practice of their own analysts.

deliver long-term returns through their recommendations. We finish by investigating how the different benchmarks affect the value of analysts' stock picking advice.

Analysts' Benchmarks and the Distribution of Recommendations

We examine whether the choice of the benchmark is associated with the characteristics of the recommendations issued by a broker. The literature has largely ignored the fact that analysts use a variety of benchmarks and has treated recommendations as homogeneous. Figure 1 plots the distribution of recommendations broken down by benchmark type. In the classification of the recommendation levels, we denote strong buys and buys as buy recommendations and sells and strong sells as sell recommendations. The figure demonstrates an important and salient feature that distinguishes the recommendation distribution of industry benchmarkers from that of market and total benchmarkers: Recommendations issued by industry benchmarkers tend to be less bullish than those issued by market or total benchmarkers. It appears that the distribution of recommendations is strongly related to benchmark choice.

< Insert Figure 1 here >

Table 3 explores the relation between benchmark choice and the distribution of recommendations in a multivariate setting. We use firm fixed-effects logistic regressions in which the dependent variable is an indicator equal to one when the recommendation is a buy or a strong buy in model (1) and sell or strong sell in model (2). Our main explanatory variables are indicators equal to one if the broker is an industry (market) benchmarker and zero otherwise.

< Insert Table 3 here >

The inclusion of firm fixed-effects frees us from having to control for firm characteristics that are not varying over time. Instead, we focus on broker characteristics and time-varying aspects that have been shown to affect the optimism of analysts. There is long literature relating conflicts of interest stemming from the relationship between investment banking and sell-side research to the optimism in analyst recommendations (e.g., Lin and McNichols 1998; Michaely and Womack 1999). We use a broker affiliation dummy to proxy for such conflicts of interest. The affiliation dummy variable is equal to one if the broker issuing the recommendation was a lead underwriter or a co-manager in an equity offering for the firm in the 24 months before the recommendation announcement date. We also control for past market and firm performance, based on the evidence that analysts chase momentum (Jegadeesh, Kim, Krische, and Lee 2004), and for broker and analyst characteristics. We further include indicators for whether the recommendation is issued by an analyst who is employed by a brokerage house that was sanctioned during the Global Settlement (Barber, Lehavy, McNichols, and Trueman 2006; Kadan et al. 2009), and for whether the brokerage house uses a three-tier recommendation grid at the time the recommendation is issued (Kadan et al. 2009). Finally, we control for analyst experience, measured as the number of days the analyst has appeared in IBES.

The results confirm the univariate inferences in Figure 1, showing that the adopted benchmark is strongly associated with the bullishness of recommendations. Columns (1) and (2) show that industry benchmarkers are less likely to issue buy and strong buy recommendations and more likely to issue sell and strong sell recommendations as compared to market and total benchmarkers. Columns (3) and (4) show in addition that market benchmarkers are less likely to issue bullish recommendations compared to total benchmarkers. Columns (5) to (8) repeat the analysis controlling for threshold levels presented in Table 2. The results are similar.[10,11]

[10] A possible explanation for the difference in the distribution of recommendation is that the coverage universe varies across the different benchmarkers. To explore this possibility, in an unreported analysis we restrict attention only to firms that are covered by all three types of benchmarkers. The results are not materially affected, reinforcing the link between benchmark choice and the distribution of recommendations.

[11] One way to reinforce the association between the benchmarks and the distribution of recommendations, is to look at instances in which a broker changes its benchmarks. Our data contain five such changes, from either total or market to industry benchmarks. In three of them, no significant shift in the distribution of recommendations follows the change in benchmark. In the other two, though, there is a significant increase in the fraction of sell recommendations around the event: a jump from 5% to 12% following a change from market to industry benchmarks, and from 3% to 17% following a change from total to industry benchmarks.

We emphasize that while the distribution of recommendations issued by industry benchmarkers is clearly tilted toward less bullish recommendations compared to that of market and total benchmarkers, one cannot simply conclude that industry benchmarkers are fundamentally less optimistic than market or total benchmarkers. The reason is that the different benchmarks imply different meanings (or adjustments) of the recommendations. This is akin to raw, market-adjusted, or industry-adjusted returns, which cannot be directly compared. Instead, the conclusion of this analysis should be that when studying the optimism of analysts through their recommendations, one should control for (or take into account) the benchmark type.[12]

Analysts' Benchmarks and Price Reactions to Recommendations

Next we evaluate the extent to which investors take into account the benchmarks when responding to stock recommendations. The complexity of generating recommendations' signals varies with the benchmark. Total benchmarkers are facing the most challenging task, since predicting whether a given stock will exceed a return threshold requires them to provide insights on the expected performance of a stock, its industry, and the market as a whole. Market benchmarkers have a less demanding task, since predicting whether a given stock will outperform the market requires ranking firms within their industries and ranking industries, but no specific outlook is required for the market as a whole. Finally, industry benchmarkers are facing the least challenging task of just ranking firms within their industries.

To this end, we study the price reactions around the issuance of analyst recommendations. We distinguish between three types of price reactions: industry-adjusted, market-adjusted, and unadjusted. A buy recommendation from an industry benchmarker implies the recommended stock is expected to outperform its industry peers. Therefore, the recommendation is a statement about the firm's expected industry-adjusted returns. If this view is shared by investors, an optimistic

[12] We thank a referee for pointing out this issue to us.

(pessimistic) recommendation should result in a positive (negative) industry-adjusted return. Similarly, if investors believe a recommendation from a market benchmarker reflects analysts' expectations of a stock's future performance relative to the market, then the price response to an optimistic (pessimistic) recommendation should be positive (negative) after adjusting for the market return. Finally, unadjusted (or raw returns) should be the natural price reaction to recommendations from total benchmarkers.

We should also see differential price reactions when recommendations are pooled across different types of benchmarkers. A buy recommendation from an industry benchmarker is a more clear statement regarding expected industry-adjusted returns than a buy from a market or total benchmarker. For a market benchmarker, for example, a buy suggests the recommended stock is expected to beat the market, but this outcome may be due to the performance of the industry rather than of the firm, so this buy does not necessarily imply positive industry-adjusted returns for a particular stock. Thus, industry-adjusted price reactions should be stronger for recommendations from industry benchmarkers compared to recommendations from other benchmarkers. Similarly, market-adjusted price reactions should be more aligned with recommendations from market benchmarkers and unadjusted price reactions to recommendations from total benchmarkers.

Panel A of Table 4 presents results from regressing the three types of price reactions to recommendation upgrades on indicators for the types of benchmarks and several control variables.[13] We focus on recommendation revisions instead of levels because prior research finds that revisions are more informative (Boni and Womack 2006; Jegadeesh and Kim 2010). The stark result from this analysis is that industry-adjusted price reactions are significantly more positive following upgrades of recommendations issued by industry benchmarkers compared to

[13] To avoid the confounding effects of price reactions to earnings, we remove recommendations issued within a three-day window of earnings announcements. We also remove days on which the IBES universe records multiple analysts issuing recommendations for the firm.

recommendations issued by market and total benchmarkers. This result is consistent with industry benchmarkers providing more useful insights in terms of ranking firms within their industries, and with investors understanding this. By contrast, we do not find evidence that recommendations issued by market benchmarkers garner significantly higher market-adjusted price reactions. Furthermore, unadjusted price reactions to total benchmarkers' upgrades are smaller than reactions to upgrades from other benchmarkers. This is consistent with the complexity hypothesis, suggesting that investors do not find recommendations issued by total benchmarkers more valuable in predicting unadjusted returns.

< Insert Table 4 here >

Due to the close relation between target prices and benchmarks, it is possible that the information in target prices subsumes much of the information in the recommendation benchmarks and their associated targets. To further evaluate the role of the benchmarks in generating price reactions, in columns (5)-(8) we control for the expected returns implied by target prices as defined earlier.[14] The sample size here is substantially reduced since we restrict attention to recommendations that are accompanied by target prices only. Note first that the expected return from target prices is a strong determinant of price reactions to recommendation upgrades. Intuitively, investors respond more positively to stock recommendations that are associated with higher expected returns. Regarding the benchmarks, the results in columns (5)-(8) are qualitatively similar to those in columns (1)-(4). In particular, industry-adjusted price reactions to upgrades issued by industry benchmarkers are significantly positive, consistent with industry benchmarkers being able to provide useful insights in terms of ranking firms within their industries.

[14] We repeat the analysis with an alternative definition of expected returns after taking into account the expected dividend yield. We also repeat the analysis controlling for changes in expected returns (rather than levels) and for earnings-forecasts revisions. The results are similar.

In Panel B we study price reactions to downgrades and we find parallel results. The only difference is that once we control for the expected return from target prices, the coefficient on the industry benchmark becomes insignificant. This, however, may well be attributed to the smaller sample.[15] To explore whether the price reaction results are driven by differences in coverage across different benchmarks, in Panel C we repeat the analysis restricting attention to firms that are covered by all three benchmark types. The conclusions are similar with the exception of somewhat weaker results for downgrades. Finally, in Panel D we control for the target thresholds documented in Table 2. The conclusions are not affected materially.

Analysts' Benchmarks and the Investment Value in Recommendations

Different benchmarks imply different objectives for recommendations. For industry benchmarkers the objective is to beat the industry peers; for market benchmarkers it means beating the market; and for total benchmarkers it means beating an absolute return threshold. In this section, we analyze the performance of analysts based on whether the recommended stocks behave "as promised" in the analysts' disclosures, meeting or beating their declared objective.

< Insert Table 5 here >

To evaluate whether the recommendation's objective has been achieved we follow two approaches. First, we compare the cumulative stock return associated with the recommendation to its stated objective as indicated by the benchmark plus the target threshold (Table 2) over a period of 12 months or until the recommendation is revoked (if earlier than 12 months).[16] Under this approach, we follow the literal meaning of the recommendation's stated objective, without

[15] In unreported results we estimate the model for downgrades on the subsample with target prices but without controlling for target price expected returns. We find similar results to those reported in columns (5)-(8) of Panel B of Table 4, suggesting that the reduction in significance in columns (5)-(8) may be attributed to the smaller sample rather than the additional controls.
[16] In other words, a recommendation is evaluated throughout its stated life span as long as its advice is still outstanding. This definition of the life span of a recommendation is similar to the approach used in the literature when examining the investment value of recommendations. See, for example, Barber, Lehavy, McNichols, and Trueman (2006) and Barber, Lehavy, and Trueman (2007).

accounting for risk. This is consistent with how the analysts' employers and institutional investors most often judge recommendations' performance.[17]

In the second approach, we also consider the risk profile of stock recommendations. We want to isolate any performance that is associated with loadings on risk factors, and only measure performance that is due to insights offered by the analysts. For this purpose we match each recommendation (a firm i that receives a buy or sell at time t) to a control unit (another firm i_c and another time period t_c) such that firm i at time t and firm i_c at time t_c have a similar risk profile based on the four Fama-French factors: beta, size, book-to-market and momentum. The matching procedure is based on the nearest neighbor matching of propensity scores (Rosembaum and Rubin 1983). This procedure solves the problem of the "curse of dimensionality" that appears when matches over multiple dimensions are required, and has been used in many different corporate finance settings (e.g., Bharath, Dahiya, Saunders, and Srinavasan 2011; Drucker and Puri 2005; Villalonga 2004; Colak and Whited 2007; Hellman, Lindsey and Puri 2008).

Panel A of Table 5 presents the fraction of buy/sell recommendations that meet their stated objectives. We report this success rate broken down by the three different benchmarks, and separately for the actual recommendations and for their control units. The results indicate that about 51 percent of buy recommendations issued by industry and market benchmarkers meet or beat their objective. By contrast, 43.8 percent of buy recommendations issued by total benchmarkers do so. The results for sell recommendations paint a similar picture. These results seem plausible, as meeting the objective for total benchmarkers is arguably a harder task. Indeed, total benchmarkers need to base their advice on predictions related to firm-specific returns, industry returns, and market returns. Note also that common targets used by total benchmarkers

[17] Conversations with sell-side analysts indicated that the benchmarks are in fact used internally by the brokers when assessing the performance of their analysts. A press article related to the benchmarkers reinforces the view that analysts do want their recommendations to be interpreted relative to the adopted benchmarks. See "Credit Suisse: These Downgrades Aren't Personal," *The Wall Street Journal*, October 2nd, 2012.

are ten percent and 15 percent (see Table 2). These quite high thresholds could also contribute to total benchmarkers' lower success rate in hitting their targets.[18]

Next, we consider whether the success rate for buy recommendations is driven by their risk characteristics. To do that, we compare the success rates between the actual recommendations and the control units. For all types of benchmarks, firms for which analysts issue buy recommendations perform better than firms with similar risk characteristics that did not receive such recommendations.[19] For example, 51.0 percent of buy recommendations issued by industry benchmarkers hit their targets, compared with 45.6 percent of control units. Additionally, while recommendations issued by total benchmarkers are less likely to beat the target when compared to those issued by market or industry benchmarkers, they perform better than their control units (43.8 percent vs. 38.4 percent). The difference in success rates between actual recommendations and their control units for sell recommendations is even larger than for buy recommendations.

Panel B of Table 5 considers the magnitudes by which analysts beat (or miss) their stated objectives. The table reports the average, as well as the median, difference between the realized return and the stated objective for each recommendation in our sample as well as for the control units. The results are consistent with those in Panel A. Indeed, industry and market benchmarkers significantly beat their stated objective for both buy and sell recommendations. For example, a buy recommendation from an industry benchmarker on average yields a return that exceeds the stated objective by 456 basis points. By contrast, total benchmarkers on average miss their stated objective. For example, a sell recommendation issued by a total benchmarker misses the target by 925 basis points. To evaluate the performance of recommendations relative to the performance of

[18] The ability of total benchmarkers to meet their stated objectives is closely tied to the performance of the market. In an unreported analysis we examine the performance of total benchmarkers separately based on ex-post annual market returns. Indeed, we find that total benchmarkers meet their buy (sell) targets more (less) frequently during periods of superior (inferior) market returns.

[19] For the control unit, we measure performance over a period with the same length and using the same stated objective as for its corresponding actual recommendation.

stocks with similar risk characteristics and stated objective, we consider the control units. We find that for all types of benchmarkers, the mean and median excess returns over the stated objectives for buy (sell) recommendations are better (worse) than those of the control units.

In sum, Table 5 reveals that for all types of benchmarks buy (sell) recommendations outperform (underperform) stocks with similar risk profiles and subject to the same investment objective. Analysts using industry or market benchmarks are more likely to achieve the stated recommendation objectives than analysts using total benchmarks. The lower success rate of total benchmarkers is likely a result of a more stringent stated objective.

Analysts' Benchmarks and Stock Picking

Institutional investors consistently rank industry knowledge as the most important research attribute of sell-side analysts (Bradley, Gokkaya, and Liu, 2017), and analysts agree with this assessment (Brown, Call, Clement, and Sharp, 2015). Accordingly, analysts' ability to identify the best (winners) and worst (losers) prospects within an industry—to which we refer as stock picking—should be a valuable skill to investors. In this section, we analyze the implications of adopting different benchmarks on the usefulness of recommendations as signals for winners and losers within industries.

Boni and Womack (2006) show that analysts create value to investors primarily through their ability to rank stocks within industries. Boni and Womack's analysis, however, does not account for the benchmarks. The adoption of a specific benchmark may have two inter-related effects on the value analysts deliver through their stock picking. First, the benchmark can either help or hinder the stock-picking ability of the analysts, and second, the benchmark can affect the ability of analysts to convey their opinion through their recommendations.

To see the first issue, note that an industry benchmarkers' main focus is on ranking stocks within an industry, while both market and total benchmarks attend to additional tasks. Consistent with the complexity hypothesis discussed earlier, one may expect the stock-picking skills of the latter would be reduced. Second, even if all benchmarkers possess stock-picking skills, different benchmarks can help or hinder an analyst's ability to effectively communicate her opinion. Indeed, buys and sells issued by industry benchmarkers are expected to be direct signals of winners and losers within an industry. In contrast, since market (total) benchmarkers add industry (industry and market) information to the formation of their recommendations, these recommendations may not be as clear an indicator of within-industry winners and losers. For example, a good stock in a very poor industry may be a buy for an industry benchmarker, but a hold or even a sell for a market benchmarker. Good prospects for a stock might also be offset by poor market prospects: a total benchmarker may issue fewer buys if she forecasts low future market performance, so that some within-industry winners may not receive a buy signal from that broker.

The discussion thus far suggests that recommendations issued by industry benchmarkers possess superior value to investors in terms of both their stock-picking performance, and in terms of their ability to convey this information. We test this prediction by applying the trading strategy devised in Boni and Womack (2006). The strategy forms within-industry portfolios which are long in upgraded stocks and short in downgraded stocks in a way that the performance of the resulting portfolios is both industry- and market-neutral. The null hypothesis is that the performance of such portfolios is identical regardless of the type of benchmark adopted by the brokers issuing the recommendations. The alternative hypothesis is that such portfolios based on recommendations issued by industry benchmarkers outperform those of market and total benchmarkers, reflecting the former's superior stock picking and/or better ability to convey information.

To illustrate the methodology, consider a portfolio based on recommendations from industry benchmarkers. For each month and each stock, we count the number of recommendation upgrades minus the number of recommendation downgrades from industry benchmarkers. If this measure is positive (negative), the stock is considered net upgraded (net downgraded). The strategy takes long (short) positions in net upgraded (net downgraded) stocks and keeps them open over the following month. In each month, we only consider stocks from industries that have at least one net upgraded and one net downgraded stock. We derive portfolio returns by first computing the weighted-average returns for the long-short positions within each industry, then equal-weighting the results across the industries. This yields a time series of monthly portfolio returns. We repeat the process to create portfolios based on recommendations from other types of benchmarkers.

< Insert Table 6 here >

Panel A of Table 6 shows mean returns, Sharpe ratios, and the Fama-French 4-factor alphas for different portfolios depending on the type of benchmark adopted by brokers. By every measure, the portfolio based on recommendations from industry benchmarkers presents the best performance, with a higher average monthly return and a higher Sharpe ratio. More importantly, while the portfolio based on recommendations from market and total benchmarkers yields an insignificant monthly risk-adjusted return (alpha) of 0.124 percent, the portfolio based on recommendations from industry benchmarkers yields a significantly positive monthly risk-adjusted return of 0.40 percent. These alphas are significantly different, as suggested by the alpha of a difference portfolio (last row of the table) that goes long on the portfolio based on market and total benchmarkers and short on the portfolio based on industry benchmarkers.

The attempt to add industry and market signals to the recommendation advice may blur the stock-picking signals from market and total benchmarkers. If there are subsets of firms for which stock picking signals are stronger, these signals may be sufficient to overcome the noise brought by the inclusion of industry and market signals. Barber, Lehavy, McNichols, and Trueman (2001) propose and confirm that the value of investment strategies depend on characteristics of the firm receiving the recommendation advice—in their case, firm size. This follows Womack's (1996) evidence that post-recommendation price drift is higher for smaller firms. Barber et al. (2001) reason that the value of analysts' advice could be heightened for smaller firms, for which there is less information publicly available, and whose price discrepancies are more difficult to be arbitraged away by investors. In a similar vein, firms with low analyst coverage are the ones for which the opportunity to unlock value is the highest.

To this end, we reexamine the manifestation of stock picking-skills on subsets of firms in which these skills are expected to be stronger. For these firms, even market and total benchmarkers may show stock picking skills. We repeat the Boni and Womack (2006) algorithm for small vs. large firms, and low- vs. high-coverage firms. We define small (large) firms as the ones whose market value of equity, measured in the month before the recommendation announcement date, is below (above) the median value of that measure. Similarly, low-coverage (high-coverage) firms are those for which the number of analysts in the previous year is below (above) the median.

Results appear in Panels B and C of Table 6. A few patterns stand out. First, as in Barber et al. (2001), the value of investment strategies are confined to the samples in which analysts' abilities are more ripe for manifestation: Fama-French 4-factor alphas for portfolios from industry benchmarkers are highly significant for the subsamples of small or low-coverage stocks, but no longer significant for the subsamples of large or high-coverage stocks. Besides, for small or low-coverage stocks, even portfolios aggregating all types of brokers demonstrate stock picking skills. On the other hand, portfolios based solely on recommendations from market and total benchmarkers never show evidence of stock picking skills, even for the subsamples of small and low-coverage stocks.[20] Evidence of stock picking skills from market and total benchmarkers is missing, even for subsamples where such skills could be manifested at their best.

[20] The result that alphas from portfolios based on all types of benchmarkers, but not from portfolios based on market and total benchmarkers alone, are significantly positive may come from an improvement of alphas from industry benchmarkers. The alpha from portfolio of industry benchmarkers in the subsample of low-coverage stocks, of 0.723, is significantly bigger than the corresponding alpha of 0.400 for the overall sample of recommendations.

Overall, the results suggest that recommendations from industry benchmarkers are indeed better indicators of within-industry winners and losers: Investors can gain more valuable within-industry information from industry benchmarkers, and more so for small and low-coverage firms.

IV. IMPLICATION FOR FUTURE RESEARCH AND CONCLUDING REMARKS

In this paper, we examine the literal meaning of sell-side analysts' stock recommendations. We document that different brokers rely on different benchmarks with respect to which the investment advice in each recommendation should be interpreted. Our empirical analyses demonstrate that benchmarks have implications on how analysts' outputs should be interpreted and used by investors.

Our results suggest that future research should consider analysts' benchmarks when studying analysts' behavior and outputs. A large stream of papers has focused on analysts' optimism and how it relates to conflicts of interests and other analysts' characteristics. We find that the benchmarks are an important determinant of the distribution of analysts' recommendations. Industry benchmarkers are less likely to issue bullish recommendations compared to market and total benchmarkers. We caution that such differences cannot be interpreted simply as systematic differences in optimism among analysts given the different meanings of recommendations. Instead, when studying analysts' optimism, one needs to control for benchmarks, essentially comparing optimism only within analysts using the same benchmark.

More generally, when considering variations in analysts' recommendations, researchers need to account for the fact that, in forming recommendations, analysts deal with within-industry and across-industry information differently, depending on their adopted benchmark. Thus, differences in opinions (disagreement) between analysts could be attributed to the benchmarks they use, a factor that might be used as a control variable in some settings.

When studying target prices, the literature has often focused on the implied expected return. This implied expected return is most consistent with stock recommendations from total benchmarkers, which is natural given the objectives defined by their benchmarks. Therefore, future research might account for the type of benchmark when studying certain aspects of target prices.

Our analysis also highlights new aspects of Clement's (1999) complexity hypothesis. We show that industry-adjusted price reactions to recommendation changes issued by industry benchmarkers are larger, reflecting the credence investors attribute to their outputs. By contrast, price reactions to upgrades issued by total benchmarkers, whose task is the most complex, are smaller, reflecting the lower confidence with which the market views their output. Future research studying market response to analysts' outputs may well account for the benchmarks.

There has been a lot of interest in the literature in whether analysts' recommendations provide investment value. Our paper sheds new light on this issue and offers a new method to evaluate investment value, accounting for risk, while simultaneously considering the different analysts' benchmarks. We also find that analysts' stock picking skills and their ability to communicate their views depend on the benchmarks. Thus, future studies investigating the value in analyst recommendations should take the different benchmarks into account.

In summary, our study suggests that both academics and investors should pay more attention to the declared objective of analysts' stock recommendations. In particular, the fact that different recommendations carry different meanings can be used to shed new light on a range of empirical questions. Ramnath, Rock, and Shane (2008), for example, advocate the need for a better understanding of how analysts operate. The different benchmarks employed by brokers suggest that information shocks would affect recommendations differently depending on the broker's benchmark, e.g., with industry shocks affecting more the recommendations from market and total benchmarkers. Other potential areas worth of a second look are the roles of incentives and bias in determining recommendations (e.g., Lin and McNichols 1998; Michaely and Womack 1999) and ranking analysts based on their stock-picking ability (e.g., Mikhail, Walther, and Willis 2004). These come naturally once one recognizes that performance is a comparison between the return of the recommended stock and the stated objective. These are left as avenues for future research.

References

Barber, B., R. Lehavy, M. McNichols, and B. Trueman. 2001. Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance* 56: 531-563.

Barber, B., R. Lehavy, M. McNichols, and B. Trueman. 2006. Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. *Journal of Accounting and Economics* 41: 87-117.

Barber, B., R. Lehavy, and B. Trueman. 2007. Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics* 85: 490-517.

Bharath, S., S. Dahiya, A. Saunders, and A. Srinavasan, 2011, Lending relationships and loan contract terms. *Review of Financial Studies* 24, 1141-1203.

Bhojraj, S., C. M. Lee, and D. K. Oler. 2003. What's my line? A comparison of industry classification schemes for capital market research. *Journal of Accounting Research* 41: 745-774.

Boni, L., and K. L. Womack. 2006. Analysts, industries, and price momentum. *Journal of Financial and Quantitative Analysis* 41: 85-109.

Bradley, D., S. Gokkaya, and X. Liu. 2017. Before an analyst becomes an analyst: Does industry experience matter? *Journal of Finance* 72: 751-792.

Bradshaw, M. 2012. Discussion of "Analysts' Industry Expertise". *Journal of Accounting and Economics* 54: 95-120.

Brown, L., A. Call, M. Clement, and N. Sharp. 2015. Inside the "Black Box" of sell-side financial analysts. *Journal of Accounting Research* 53: 1-47.

Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285-303.

Colak, G., and T. Whited. 2007. Spin-offs, divestitures, and conglomerate investments. *Review of Financial Studies* 20: 557-595.

Drucker, S., and M. Puri. 2005. On the benefits of concurrent lending and underwriting. *Journal of Finance* 60: 2763–99.

Hellman, T., L. Lindsey, and M. Puri. 2008. Build relationships early: Banks in venture capital. *Review of Financial Studies* 21: 513-541.

Jegadeesh, N., J. Kim, S. D. Krische, and C. M. Lee. 2004. Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59: 1083-1124.

Jegadeesh, N., and J. Kim. 2010. Do analysts herd? An analysis of recommendations and market reactions. *Review of Financial Studies* 23: 901-37.

Kadan, O., L. Madureira, R. Wang, and T. Zach. 2009. Conflicts of interest and stock recommendations: The effect of the global settlement and related regulations. *Review of Financial Studies* 22: 4189-4217.

Kadan, O., L. Madureira, R. Wang, and T. Zach. 2012. Analysts' industry expertise. *Journal of Accounting and Economics* 54: 95-120.

Lin, H., and M. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101-127.

35: 167-199.

Michaely, R. and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12: 653-686.

Mikhail, M., B. Walther, and R. Willis. 2004. Do security analysts exhibit persistent differences in stock picking ability? *Journal of Financial Economics* 74: 67-91.

Ramnath S., S. Rock, and P. B. Shane. 2008. Financial analysts' forecasts and stock recommendations: A review of the research. *Foundations and Trends in Finance* 2: 311-421.

Rosenbaum, P., and D. Rubin. 1983. The central role of the propensity Score in Obsevational Studies for Causal Effects. *Biometrika* 70: 373-401.

Womack, K. 1996. Do brokerage analysts' recommendations have investment value? *Journal of Finance* 54: 137-157.

TABLE 1

Summary Statistics

This table presents summary statistics on the different types of benchmarks and their description. Only brokerage houses which issued at least 200 recommendations to U.S. firms during our sample period (9/2002 – 12/2014) are included in the analysis. For each type of benchmark, we describe the meaning of the benchmark, report the number of brokers using it, the mean and median of the number of recommendations issued by each broker, the total number of recommendations issued by all brokers, the percentage out of the total number of recommendations, and the number of top-20 brokers (based on the number of issued recommendations) using it.

			# of Recommendations Per Broker				
Benchmark	Description	No. of Brokers	Mean	Median	Total	% of all	Top 20 Brokers
Industry	Recommendation is benchmarked against performance of peers in the same sector	38	2257	904	85756	24.14%	7
Total	Recommendation is based on a stock's total return.	52	1352	779	70285	19.78%	3
Market	Recommendation is benchmarked against market performance	29	1815	928	52633	14.81%	2
Changes	A broker changes the benchmark during our sample period and we cannot identify when the broker made the change.	7	3723	2612	26061	7.33%	2
Sector/risk	Recommendation is based on a stock's risk-adjusted return relative to industry performance.	7	3425	1343	23977	6.75%	2
Total/risk	Recommendation is based on a stock's risk-adjusted return relative to sector performance.	12	1897	1814	22768	6.41%	1
No data	Cannot find data on the definition of ratings.	39	566	285	22065	6.21%	0
Not sure	Cannot identify which benchmark a broker uses.	11	1139	766	12531	3.53%	1
Market/Total	Recommendations is based on a stock's total return and/or benchmarked against market performance.	7	1781	451	12468	3.51%	1
Industry/Total	Recommendations is based on a stock's total return and/or benchmarked against industry performance.	5	2363	2215	11815	3.33%	1
Market/risk	Recommendation is based on a stock's risk-adjusted return relative to the market performance.	4	2506	2394	10025	2.82%	0
Market/Industry	Recommendation is benchmarked against market and/or industry performance.	4	830	693	3318	0.93%	0
Market/Industry/Total	Recommendations is based on a stock's total return and/or benchmarked against market and/or industry performance.	2	808	808	1615	0.45%	0
All		217			355317		

TABLE 2

Distribution of Target Thresholds

This table summarizes the distribution of buy and sell targets thresholds for industry, market and total benchmarkers in our sample. For an industry benchmarker, a buy (sell) recommendation target is defined as the 'x' percent return a stock is expected to outperform (underperform) the industry. For a market benchmarker, a buy (sell) recommendation target is defined as the 'x' percent return a stock is expected to outperform (underperform) the market. For a total benchmarker, a buy (sell) recommendation target is defined as the 'x' (negative 'x') percent total return a stock is expected to outperform (underperform).

Panel A: Industry Benchmarkers

Buy Recommendations				Sell Recommendations			
Target	# Brokers	# Recs	% Recs	Target	# Brokers	# Recs	% Recs
0	33	29296	86.8%	0	33	9183	90.9%
10%	2	1667	4.9%	10%	2	294	2.9%
11%	1	149	0.4%	11%	1	20	0.2%
20%	1	175	0.5%	20%	1	38	0.4%
N.A.	1	2455	7.3%	N.A.	1	566	5.6%
All	38	33742	100%	All	38	10101	100%

Panel B: Market Benchmarkers

Buy Recommendations				Sell Recommendations			
Target	# Brokers	# Recs	% Recs	Target	# Brokers	# Recs	% Recs
0	17	14186	61.2%	0	19	2142	54.4%
5%	5	2466	10.6%	5%	4	394	10.0%
6%	1	4433	19.1%	6%	1	1097	27.8%
10%	1	107	0.5%	15%	3	215	5.5%
15%	3	1201	5.2%	20%	1	20	0.5%
20%	1	124	0.5%				
N.A.	2	647	2.8%	N.A.	2	72	1.8%
All	30	23164	100%	All	30	3940	1.0

Panel C: Total Benchmarkers

Buy Recommendations				Sell Recommendations			
Target	# Brokers	# Recs	% Recs	Target	# Brokers	# Recs	% Recs
0	3	1193	3.3%	0	17	1031	22.2%
5%	1	231	0.6%	5%	1	24	0.5%
6%	1	83	0.2%	6%	1	4	0.1%
7%	1	260	0.7%	10%	16	1854	39.8%
10%	19	11642	32.5%	13%	1	3	0.1%
13%	1	208	0.6%	15%	14	836	18.0%
15%	24	12244	34.2%	20%	3	202	4.3%
20%	5	2994	8.4%				
25%	1	3412	9.5%				
30%	1	97	0.3%				
N.A.	3	3448	9.6%	N.A.	6	699	15.0%
All	60	35812	100%	All	59	4653	100%

TABLE 3

Determinants of Optimism/Pessimism in Recommendations

The table presents results of logistic regressions whose dependent variable equals 1 when a recommendation is either optimistic or pessimistic. Our sample period is between 9/2002 and 12/2014. All models use firm fixed effects. Optimistic recommendations are strong buy and buy, and pessimistic recommendations are sell and strong sell. **Industry** (**Market**) takes value of 1 if a broker uses an industry benchmark and 0 if a broker uses other benchmarks. **Aff** is an indicator variable equal to 1 if the broker issuing the recommendation was a lead underwriter or a co-manager in an equity offering for the firm in the 24 months before the recommendation announcement date. **Sanct** is an indicator variable equal to 1 if the recommendation is issued by an analyst who is employed by a sanctioned brokerage house. **Pastfirmperf** is the stock return over [-180, -2]. **Pastmkperf** is the market return over [-180, -2]. **Experience** is defined as the number of days the analyst has appeared in IBES. **Tier3** is an indicator variable for whether a brokerage house uses a three-tier recommendation grid at the time a recommendation is issued. **Target_buy** (**Target_sell**) is the buy (sell) targets thresholds for each broker. For detailed definition on targets thresholds, please see table 2. Robust standard errors (in parentheses) are clustered at the firm level. ***, **, * denote statistical significance at the 1percent, 5 percent, and 10 percent levels, respectively.

TABLE 3 *(continued)*

	(1) Prob(Rec=OPT)	(2) Prob(Rec=PESS)	(3) Prob(Rec=OPT)	(4) Prob(Rec=PESS)	(5) Prob(Rec=OPT)	(6) Prob(Rec=PESS)	(7) Prob(Rec=OPT)	(8) Prob(Rec=PESS)
Industry	-0.227***	0.410***	-0.300***	0.425***	-0.144***	0.410***	-0.276***	0.407***
	(0.0104)	(0.0215)	(0.0116)	(0.0260)	(0.0124)	(0.0255)	(0.0192)	(0.0356)
Market			-0.160***	0.0329			-0.155***	-0.00406
			(0.0117)	(0.0288)			(0.0179)	(0.0343)
Target_buy					0.868***		0.182*	
					(0.0732)		(0.104)	
Target_sell						0.874***		0.861***
						(0.229)		(0.257)
AFF	0.252***	-0.618***	0.257***	-0.619***	0.264***	-0.637***	0.266***	-0.637***
	(0.0199)	(0.0456)	(0.0199)	(0.0456)	(0.0209)	(0.0474)	(0.0209)	(0.0474)
PASTMKTPERF	0.292***	-0.347***	0.284***	-0.345***	0.249***	-0.280***	0.243***	-0.280***
	(0.0560)	(0.0974)	(0.0559)	(0.0973)	(0.0573)	(0.100)	(0.0573)	(0.100)
PASTFIRMPERF	0.157***	-0.237***	0.158***	-0.237***	0.157***	-0.239***	0.159***	-0.239***
	(0.0192)	(0.0400)	(0.0191)	(0.0400)	(0.0197)	(0.0400)	(0.0197)	(0.0400)
SANCT	-0.228***	0.366***	-0.214***	0.363***	-0.231***	0.385***	-0.223***	0.385***
	(0.0124)	(0.0242)	(0.0125)	(0.0243)	(0.0136)	(0.0257)	(0.0137)	(0.0260)
EXPERIENCE	-0.0220***	0.0255***	-0.0220***	0.0255***	-0.0205***	0.0184***	-0.0215***	0.0184***
	(0.00298)	(0.00602)	(0.00298)	(0.00602)	(0.00308)	(0.00619)	(0.00308)	(0.00619)
TIER3	-0.257***	-0.180***	-0.243***	-0.182***	-0.276***	-0.128***	-0.257***	-0.128***
	(0.0122)	(0.0264)	(0.0123)	(0.0263)	(0.0135)	(0.0297)	(0.0137)	(0.0300)
Observations	176,934	148,307	176,934	148,307	163,486	131,443	163,486	131,443
Pseudo R2	0.010	0.017	0.011	0.017	0.011	0.016	0.011	0.016

TABLE 4

Price Reactions to Recommendations

This table reports OLS regression results for the market reaction to analysts' recommendation upgrades/downgrades for the full sample. The dependent variable is the cumulative 3-day return (market adjusted or industry adjusted or raw) around the announcement of upgrades/downgrades by analysts. Our sample period is between 9/2002 and 12/2014. **Industry** (**Market or Total**) takes value of 1 if a broker uses an industry (Market or Total) benchmark and 0 if a broker uses other benchmarks. **Pastfirmperf** is the stock return over [-180, -2]. **Pastmkperf** is the market return over [-180, -2]. **Experience** is defined as the number of days the analyst has appeared in IBES. **Firmsize** is defined as the log of market value of equity 30 days prior to the recommendation date. **BE/ME** is defined as the book value of equity over market value of equity at the end of previous fiscal year. **ERTP** is the expected return implied from target price which is measured as $(PT_0 - P_{-1})/P_{-1}$, where PT_0 is the target price and P_{-1} is the closing stock price the day before the target price issuance. **Target_buy** (**Target_sell**) is the buy (sell) targets thresholds for each broker. For detailed definition on targets thresholds, please see table 2. Robust standard errors (in parentheses) are clustered at the firm level. ***, **, * denote statistical significance at the 1percent, 5 percent, and 10 percent levels, respectively.

Panel A: Upgrades

	Without Target Price				With Target Price			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ind. Adj	Mkt. Adj	Raw	Raw	Ind. Adj	Mkt. Adj	Raw	Raw
Industry	0.0039***				0.0024***			
	(0.0005)				(0.0005)			
Market		0.0002		-0.0025***		-0.0001		-0.0022***
		(0.0005)		(0.0007)		(0.0007)		(0.0008)
Total			-0.0033***	-0.0043***			-0.0017***	-0.0025***
			(0.0005)	(0.0006)			(0.0006)	(0.0007)
ERTP					0.0138***	0.0141***	0.0139***	0.0139***
					(0.0013)	(0.0014)	(0.0014)	(0.0014)
Pastfirmperf	-0.0022**	-0.0025**	-0.0020*	-0.0020*	0.0015	0.0015	0.0017	0.0017
	(0.0010)	(0.0011)	(0.0011)	(0.0011)	(0.0012)	(0.0013)	(0.0013)	(0.0013)
Pastmkperf	0.0101***	0.0107***	0.0078**	0.0076**	0.0133***	0.0106***	0.0089**	0.0086**
	(0.0027)	(0.0029)	(0.0033)	(0.0033)	(0.0032)	(0.0035)	(0.0039)	(0.0039)
Experience	0.0013***	0.0014***	0.0014***	0.0014***	0.0012***	0.0013***	0.0014***	0.0014***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Firmsize	-0.0056***	-0.0053***	-0.0054***	-0.0054***	-0.0059***	-0.0055***	-0.0056***	-0.0056***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Log(1+BM)	0.0043***	0.0052***	0.0060***	0.0060***	0.0056***	0.0063***	0.0073***	0.0073***
	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0015)	(0.0015)
Constant	0.0901***	0.0863***	0.0906***	0.0917***	0.0907***	0.0872***	0.0897***	0.0906***
	(0.0026)	(0.0027)	(0.0029)	(0.0029)	(0.0031)	(0.0032)	(0.0034)	(0.0035)
Observations	56,349	56,349	56,349	56,349	37,931	37,931	37,931	37,931
R-squared	0.0346	0.0276	0.0242	0.0244	0.0558	0.0472	0.0403	0.0405

TABLE 4 ***(continued)***

Panel B: Downgrades

	Without Target Price				With Target Price			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ind. Adj	Mkt. Adj	Raw	Raw	Ind. Adj	Mkt. Adj	Raw	Raw
Industry	-0.0019***				-0.0005			
	(0.0007)				(0.0011)			
Market		-0.0011		0.0005		-0.0014		-0.0005
		(0.0008)		(0.0009)		(0.0013)		(0.0015)
Total			0.0025***	0.0027***			0.0014	0.0013
			(0.0008)	(0.0009)			(0.0013)	(0.0014)
ERTP					0.0159***	0.0164***	0.0165***	0.0165***
					(0.0030)	(0.0031)	(0.0032)	(0.0032)
Pastfirmperf	0.0116***	0.0123***	0.0121***	0.0121***	0.0121***	0.0119***	0.0111***	0.0111***
	(0.0015)	(0.0016)	(0.0016)	(0.0016)	(0.0024)	(0.0025)	(0.0026)	(0.0026)
Pastmkperf	0.0188***	0.0197***	0.0243***	0.0243***	0.0230***	0.0253***	0.0306***	0.0305***
	(0.0040)	(0.0042)	(0.0046)	(0.0046)	(0.0056)	(0.0059)	(0.0065)	(0.0065)
Experience	-0.0010***	-0.0012***	-0.0013***	-0.0013***	-0.0014***	-0.0016***	-0.0017***	-0.0017***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Firmsize	0.0060***	0.0055***	0.0056***	0.0056***	0.0075***	0.0071***	0.0072***	0.0072***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Log(1+BM)	0.0105***	0.0105***	0.0118***	0.0119***	0.0112***	0.0123***	0.0131***	0.0131***
	(0.0018)	(0.0019)	(0.0020)	(0.0020)	(0.0026)	(0.0027)	(0.0029)	(0.0029)
Constant	-0.1094***	-0.1008***	-0.1028***	-0.1031***	-0.1340***	-0.1277***	-0.1284***	-0.1282***
	(0.0044)	(0.0046)	(0.0049)	(0.0049)	(0.0064)	(0.0067)	(0.0070)	(0.0070)
Observations	37,141	37,141	37,141	37,141	16,270	16,270	16,270	16,270
R-squared	0.0303	0.0252	0.0239	0.0239	0.0485	0.0419	0.0380	0.0380

TABLE 4 *(continued)*

Panel C: Subsample of Firms Which are Covered by All Three Benchmarkers in a Year

	Upgrades				Downgrades			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ind. Adj	Mkt. Adj	Raw	Raw	Ind. Adj	Mkt. Adj	Raw	Raw
Industry	0.0034***				-0.0014			
	(0.0005)				(0.0008)			
Market		0.0004		-0.0021**		-0.0011		0.0003
		(0.0006)		(0.0008)		(0.0010)		(0.0012)
Total			-0.0032***	-0.0041***			0.0018*	0.0019*
			(0.0007)	(0.0008)			(0.0010)	(0.0011)
Pastfirmperf	-0.0030**	-0.0035**	-0.0030*	-0.0031**	0.0079***	0.0092***	0.0090***	0.0090***
	(0.0014)	(0.0014)	(0.0016)	(0.0016)	(0.0019)	(0.0020)	(0.0020)	(0.0020)
Pastmkperf	0.0018	0.0009	-0.0016	-0.0017	0.0142***	0.0136***	0.0165***	0.0165***
	(0.0033)	(0.0037)	(0.0043)	(0.0043)	(0.0048)	(0.0050)	(0.0056)	(0.0056)
Experience	0.0010***	0.0010***	0.0010***	0.0010***	-0.0010***	-0.0013***	-0.0012***	-0.0012***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Firmsize	-0.0053***	-0.0050***	-0.0050***	-0.0050***	0.0062***	0.0055***	0.0056***	0.0056***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0004)	(0.0004)
Log(1+BM)	0.0042***	0.0057***	0.0067***	0.0067***	0.0107***	0.0109***	0.0128***	0.0128***
	(0.0015)	(0.0016)	(0.0017)	(0.0017)	(0.0020)	(0.0021)	(0.0023)	(0.0023)
Constant	0.0885***	0.0852***	0.0877***	0.0888***	-0.1120***	-0.1008***	-0.1024***	-0.1025***
	(0.0034)	(0.0036)	(0.0040)	(0.0040)	(0.0058)	(0.0061)	(0.0065)	(0.0065)
Observations	30,068	30,068	30,068	30,068	20,534	20,534	20,534	20,534
R-squared	0.0323	0.0243	0.0205	0.0208	0.0324	0.0248	0.0224	0.0224

TABLE 4 *(continued)*

Panel D: Controlling for the Targets Thresholds

	Upgrades				Downgrades			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ind Adj	Mkt. Adj	Raw	Raw	Ind Adj	Mkt. Adj	Raw	Raw
Industry	0.0025***				-0.0016*			
	(0.0006)				(0.0009)			
Market		-0.0013**		-0.0023***		-0.0003		0.0010
		(0.0006)		(0.0007)		(0.0008)		(0.0010)
Total			-0.0011	-0.0025**			0.0024**	0.0030**
			(0.0009)	(0.0010)			(0.0011)	(0.0013)
Target_buy	-0.0229***	-0.0320***	-0.0233***	-0.0203***				
	(0.0037)	(0.0033)	(0.0057)	(0.0059)				
Target_sell					-0.0194**	-0.0173**	-0.0325***	-0.0342***
					(0.0078)	(0.0072)	(0.0091)	(0.0094)
Pastfirmperf	-0.0020*	-0.0023**	-0.0017	-0.0017	0.0113***	0.0119***	0.0118***	0.0118***
	(0.0010)	(0.0011)	(0.0012)	(0.0012)	(0.0016)	(0.0017)	(0.0018)	(0.0018)
Pastmkperf	0.0103***	0.0099***	0.0077**	0.0076**	0.0175***	0.0181***	0.0215***	0.0215***
	(0.0028)	(0.0030)	(0.0034)	(0.0034)	(0.0043)	(0.0045)	(0.0049)	(0.0049)
Experience	0.0013***	0.0013***	0.0013***	0.0013***	-0.0009***	-0.0012***	-0.0012***	-0.0012***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Firmsize	-0.0057***	-0.0054***	-0.0054***	-0.0054***	0.0064***	0.0059***	0.0060***	0.0060***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Log(1+BM)	0.0041***	0.0052***	0.0061***	0.0060***	0.0118***	0.0120***	0.0135***	0.0135***
	(0.0012)	(0.0013)	(0.0014)	(0.0014)	(0.0019)	(0.0020)	(0.0021)	(0.0021)
Constant	0.0931***	0.0907***	0.0914***	0.0923***	-0.1163***	-0.1081***	-0.1086***	-0.1091***
	(0.0028)	(0.0029)	(0.0030)	(0.0030)	(0.0048)	(0.0050)	(0.0053)	(0.0053)
Observations	52,048	52,048	52,048	52,048	32,548	32,548	32,548	32,548
R-squared	0.0348	0.0284	0.0241	0.0243	0.0329	0.0274	0.0255	0.0256

TABLE 5

Performance of Recommendations

This table analyzes the performance of buy and sell recommendations issued by market/industry/total benchmarkers. Our sample period is between 9/2002 and 12/2014. Each recommendation is paired with a propensity score matched (control) unit according to the procedure described in Table A1. The table reports performance measures for the sample of recommendations and the corresponding sample of control units. In Panel A, the performance variable for each recommendation (control unit) is a dummy equal to 1 if the recommendation (control unit) achieved its stated objective. For a buy recommendation, the stated objective from an industry (market) [total] benchmarker is $R_{industry}$+*target* (R_{market}+*target*)[*target*], so achieving the objective means R-$R_{industry}$ - *target*>*0* (R-R_{market}-*target*>*0*)[*R-target*>*0*]. For a sell recommendation, the stated objective from an industry (market) [total] benchmarker is $R_{industry}$ - *target* (R_{market} - *target*)[*target*], so achieving the objective means R-$R_{industry}$ + *target*<*0* (R-R_{market}+*target*<*0*)[*R-target*<*0*]. For a control unit, the stated objective is the same as in its corresponding recommendation. In Panel B, the performance variable is the difference between the cumulative stock return and the stated objective. In Panel C, the performance variable is the raw return. Returns associated with a recommendation (the stock return R, the industry return $R_{industry}$ and the market return R_{market}) are computed during the stated life span of a recommendation—the period in which the recommendation advice is kept alive. This is the period between the recommendation issuance and the earliest of (i) 12 months following the recommendation issuance and (ii) the date when the recommendation advice is changed (e.g., though a cancelation or an upgrade/downgrade by the same analyst). Returns associated with a control unit are computed for the period starting with the control unit issuance date (as defined in Table A1) and with the same number of days as the stated life span of its corresponding recommendation. P-values for test of difference of proportions is reported under the column *Diff (p-value)*.

Panel A: Proportion of Recommendations Achieving the Stated Objective

	Buys				Sells			
		% achieving the objective				% achieving the objective		
	# obs	Buy	Control	Diff p-value	# obs	Sell	Control	Diff p-value
Industry	13,079	51.0%	45.6%	0.0000	4,351	59.8%	49.7%	0.0000
Market	11,876	50.9%	44.7%	0.0000	2,149	54.6%	42.3%	0.0000
Total	19,409	43.8%	38.4%	0.0000	2,627	43.9%	25.3%	0.0000

Panel B: Return in Excess of the Stated Objective

	Buys						Sells					
	Recommendation		Control		Diff (p-value)		Recommendation		Control		Diff (p-value)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Industry	0.0456	0.0053	0.0119	-0.0210	0.0000	0.0000	-0.0422	-0.0491	0.0449	0.0013	0.0000	0.0000
Market	0.0551	0.0060	0.0044	-0.0272	0.0000	0.0000	0.0010	-0.0277	0.0668	0.0295	0.0000	0.0000
Total	-0.0052	-0.0518	-0.0321	-0.0781	0.0000	0.0000	0.0925	0.0480	0.1943	0.1416	0.0000	0.0000

TABLE 6

Industry- and Market Neutral Portfolio Alphas

This table presents statistics (mean return, Sharpe ratio and Fama-French 4-factor alpha) for returns on industry- and market-neutral portfolios formed based on recommendations and the type(s) of benchmarker(s) issuing the recommendation. Sampling follows the requirements in Boni and Womack (2006) of only using stocks trading in NYSE, Nasdaq or Amex, with share codes of 10, 11, 12, 18, 30 and 31, and priced above $5. The sample period is September 2002 through December 2014. Panel A designs portfolios using all available stocks that survive the data requirements in Boni and Womack (2006). Panel B repeats the process based on a breakdown of the original sample into subsamples of small vs. large stocks. For this, each month we compute the distribution of firm size, proxied by market value of equity. We then define small (large) stock as one whose firm size in the month before the recommendation announcement date is below (above) the median value of lagged firm size. Panel C uses a breakdown of the original sample into subsamples of low- vs. high-coverage stocks. For stock j at month t, we define analyst coverage as the number of analysts that have issued at least one recommendations for j in the 12 months preceding t. We then compute the monthly distribution of analyst coverage. Finally, we define low-coverage (high-coverage) stock as one whose coverage in the month before the recommendation announcement date is below (above) the median value of lagged analyst coverage. Each row analyzes returns of a portfolio based on recommendations from one specific set of brokers—according to the type of benchmark adopted by these brokers. To construct each portfolio, at the end of each month we compute for each stock the number of recommendation upgrades minus the number of recommendation downgrades from that set of brokers. If the measure is positive (negative), the stock is considered net upgraded (net downgraded). The portfolio consists of taking long (short) positions in net upgraded (net downgraded) stocks at the end of the month and keeping the positions open over the following month. Portfolio returns are computed first within each industry then equal-weighted across the industries. The portfolio's monthly excess returns are regressed on the four Fama-French factors (excess market return, SMB, HML, and UMD).

Panel A: All stocks

	Mean	Sharpe	FF 4-factor alpha	
Benchmarker	return	ratio	Coeff	t-stat
1. Industry/Market/Total	0.300	0.169	0.148	1.57
2. Market/Total	0.310	0.124	0.120	0.93
3. Industry	0.528	0.296	0.400	3.38
4. Market/Total - Industry (2 minus 3)	-0.219	-0.165	-0.400	-2.40

Table 6. (*Continued*)

Panel B: Breakdown by firm size

Sample: Small stocks

Benchmarker	Mean return	Sharpe ratio	FF 4-factor alpha Coeff	FF 4-factor alpha t-stat
1. Industry/Market/Total	0.579	0.325	0.428	3.52
2. Market/Total	0.445	0.137	0.278	1.39
3. Industry	0.738	0.282	0.620	3.27
4. Market/Total - Industry (2 minus 3)	-0.294	-0.127	-0.460	-1.70

Sample: Large stocks

Benchmarker	Mean return	Sharpe ratio	FF 4-factor alpha Coeff	FF 4-factor alpha t-stat
1. Industry/Market/Total	0.186	0.055	0.054	0.49
2. Market/Total	0.277	0.074	0.090	0.50
3. Industry	0.336	0.0883	0.295	1.48
4. Market/Total - Industry (2 minus 3)	-0.059	-0.062	-0.329	-1.40

Panel C: Breakdown by coverage

Sample: Low-coverage stocks

Benchmarker	Mean return	Sharpe ratio	FF 4-factor alpha Coeff	FF 4-factor alpha t-stat
1. Industry/Market/Total	0.508	0.275	0.341	2.85
2. Market/Total	0.316	0.088	0.162	0.85
3. Industry	0.827	0.304	0.723	3.63
4. Market/Total - Industry (2 minus 3)	-0.511	-0.185	-0.680	-2.40

Sample: High-coverage stocks

Benchmarker	Mean return	Sharpe ratio	FF 4-factor alpha Coeff	FF 4-factor alpha t-stat
1. Industry/Market/Total	0.145	0.024	0.002	0.02
2. Market/Total	0.123	0.004	0.000	-0.02
3. Industry	0.271	0.057	0.243	1.05
4. Market/Total - Industry (2 minus 3)	-0.148	-0.076	-0.360	-1.30

FIGURE 1

End-of-Month Distribution of Outstanding Recommendations

This figure presents, for each month between September 2002 and December 2014, the fractions of buys and strong buys (denoted as buys) and of sells and strong sells (denoted as sells) among the outstanding recommendations issued by market, total, and industry benchmarkers. Only brokerage houses which issued at least 200 recommendations to U.S. firms during our sample period (9/2002 – 12/2014) are included in the analysis.