Performance Evaluation of Chinese Equity Analysts

Yeguang Chi[1]

July 2018

ABSTRACT

I construct and study a comprehensive novel dataset on Chinese equity analysts. I find the analysts possess significant forecasting skill. First, I find that more favorable recommendations predict better stock performance for at least six months after the recommendation. Second, the value of analyst recommendations is larger for smaller stocks and for the initial analyst coverage on a given stock. Third, I find evidence for analysts' learning on the job. Fourth, I observe some persistence in analyst skill but it quickly mean-reverts. Finally, I show a simple feasible trading strategy following analyst recommendations to outperform the market by 11.60% annually.

Keywords: equity analyst, Chinese stock market, information asymmetry, market efficiency, learning by doing

JEL Codes: G11, G12, G14, G15

[1] Shanghai Advanced Institute of Finance (SAIF), Shanghai Jiaotong University. Email: ygchi@saif.sjtu.edu.cn

1. Introduction

Since its inception in 1991, China's stock market has gone through dramatic changes and rapid development. After more than twenty years of development, China's stock market has become the second largest in the world, with a market capitalization of $6 trillion at the end of 2014. Despite the enormous size of the market, many fundamental empirical questions on this market remain unanswered. One important aspect of the development of the market is related to the increasing participation of finance professionals. In this paper, I investigate the role of sell-side analysts in the Chinese stock market.

The role of analysts is to gather, analyze, and disseminate information to the market by making forecasts and issuing recommendations. A key question is whether analyst recommendations add value to investors. On the one hand, the opaque nature of the Chinese stock market makes the role of sell-side analysts particularly important in terms of providing information to investors. On the other hand, both financial media and academic researchers criticize Chinese sell-side analysts for compromising their objectivity by inflating their recommendations to cater to corporate interests.

Evidence from the United States generally finds that analyst recommendations add value to investors. For example, Stickel (1995) and Womack (1996) document that upgraded stocks tend to outperform downgraded stocks. Barber, Lehavy, McNichols and Trueman (2001), Jegadeesh et al. (2003), and Boni and Womack (2003) find that the stocks with the most favorable recommendations outperform the stocks with the least favorable recommendations. These findings indicate that analysts are able to forecast future stock returns, and investors can benefit from analysts' recommendations if they consider the relative levels of recommendations across stocks, or if they pay attention to changes in recommendations. Similar return-forecasting power is also present in other countries. Jegadeesh and Kim (2006) examine the value of analyst recommendation revisions in G7 countries and find that the value of analyst recommendation revisions is higher in countries with larger stock markets, for instance US and Japan. Moshirian, Ng, Wu (2009) find similar return predictability of analyst recommendations in emerging countries, and also document a stronger positive bias in analyst recommendations and revisions.

These previous studies find that the information content of analyst recommendations and their forecasting powers can vary a lot in different market conditions. Given that relatively few papers examine the role of analysts in the China, I attempt to fill in this gap by providing a comprehensive study on the role of analyst recommendations in the Chinese stock market.

Consistent with the findings in the US and other countries, I find that analysts possess significant return-forecasting skill. First, I find that more favorable recommendations predict better stock performance for at least six months after the recommendation. One major role of the sell-side analysts is to alleviate the information asymmetry between the firms and the investors. I expect analyst recommendations to be of more value for stocks that are subject to more severe information asymmetry. I use three proxies for a stock's information asymmetry: stock size, stock IPO, and the length of a stock's no-coverage period. Smaller stocks are commonly thought to be associated with a higher degree of information asymmetry because fewer investors follow small-cap stocks. Consistent with my intuition, I find that analyst recommendations do provide more value to smaller stocks than larger stocks. Moreover, just after a stock IPO, there is likely a high degree of information asymmetry surrounding the stock because few investors are familiar with the stock. Similarly, when a stock has gone a long time without any analyst coverage, fewer investors are up to date about the company, so information asymmetry is high. Consistent with my intuition, initial analyst coverage on a stock following the stock's IPO provides more value. Additionally, initial analyst coverage on a stock after the stock has gone without any coverage for six months or longer provides more value.

Given that Chinese analysts do exhibit ability in predicting future stock returns, I then ask the question whether this ability depends on analyst experience. Mikhail, Walther and Willis (1997) document that analysts' forecasting accuracy improves as the time they follow the firm increases. I use analyst tenure and analyst-stock tenure (how long an analyst has covered a given stock) to measure analyst's general and firm-specific experiences. I find some evidence for analysts' learning on the job. Analysts who have covered a stock for a longer time exhibit better performance in their recommendations of the stock.

Another related issue is whether analysts' innate ability determines the quality of the recommendation. Extant literature discusses the issue of innate ability of analysts. Jacob, Lys and Neale (1999) find commonalities in analysts' performance across all the companies they follow. Clement, Koonce and Lopez (2007) show that both analysts' innate ability and analysts' experience matter in their forecast accuracy. I attempt to examine whether analysts exhibit innate ability in stock recommendations by studying the persistence of the performance of their recommendations. If some analysts have superior innate ability in producing more valuable recommendations, I should expect that the quality of an analyst's previous recommendations to predict the quality of his future recommendations. Specifically, I sort analysts into deciles by their past performance and follow these deciles of analysts for the next period. I find that analysts whose recommendations outperform others in the past tend to produce better recommendations in the subsequent period. Despite the fact the persistence quickly mean-reverts, an investor can still capture a monthly alpha of 1.76% (t=5.58) by always following the top-decile analysts.

Finally, how much trading profit analysts can generate for investors is an important empirical question. I answer this question by constructing two trading strategies following analyst recommendations. First, a buy-and-hold strategy that invests in the stocks with the most favorable recommendation type (strong buy) outperforms the market by 11.60% annually. Second, a long-short strategy that longs upgraded stocks and shorts downgraded stocks outperforms the market by 8.10% annually, with a close-to-zero loading on the market. These results suggest that both analyst recommendations and changes in analyst recommendations may provide substantial investment value for investors.

The rest of this paper is organized as follows. Section 2 discusses my data. Section 3 evaluates the performance of analyst recommendations with respect to recommendation characteristics, stock characteristics, and analyst characteristics. Section 4 investigates the persistence of analyst performance. Section 5 shows results on two trading strategies following analyst recommendations. Section 6 concludes. The appendix includes additional result and my data dictionary.

2. Data

2.1 Analyst Data

I collect data on Chinese equity analysts from WIND® and GTA®, the two major financial data providers in China. Analyst data in WIND starts from January 2004, whereas GTA starts from January 2001. However, GTA data before January 2004 has relatively few observations compared to data after January 2004. Therefore, I choose to focus on the sample period between January 2004 and October 2015.

The two data sources have a substantial amount of non-overlapped coverage. In the appendix, I lay out a detailed data dictionary as a guide to process, merge, and combine these two datasets. WIND dataset on its own contains 299,418 analyst reports, whereas GTA dataset contains 286,838 analyst reports. After my processing, the resulting combined dataset has a total of 386,103 analyst reports, a significant improvement over either one of the datasets.

In this study, I focus on the recommendation types from analyst reports. In particular, I have five recommendation types in my data: 1, 2, 3, 4, and 5. Type 1 is the most favorable recommendation an analyst can issue. Type 5 is the least favorable recommendation an analyst can issue. Types 1, 2, 3, 4, and 5 are commonly referred to as "strong buy", "buy", "hold", "sell", and "strong sell". Each row in my data represents a unique analyst report. My variables include report date, stock ticker, analyst, analyst firm, recommendation type, and recommendation change (upgrade, no change, or downgrade).

2.2 Stock Data

I collect Chinese A-share stock data from WIND®. I cover all the publicly listed stocks in Shanghai Stock Exchange and Shenzhen Stock Exchange, which also constitute the stock universe of Chinese equity analysts. My dataset includes but is not limited to daily data of stock returns, trading status, market capitalization, high, low, open, close, value-weighted average price (vwap), and major index returns (SSE50, CSI300, and CSI500[2]), annual data of book value at the end of each June, industry classifications following the global industry classification standard (GICS), and IPO dates.

[2] SSE50 is a stock index that represents the largest 50 stocks that trade on the Shanghai Stock Exchange. CSI300 is a stock index that represents the largest 300 stocks that trade on the Shanghai Stock Exchange and Shenzhen Stock Exchange. CSI500 is a stock index that represents the most liquid 500 mid-/small-cap stocks that trade on the Shanghai Stock Exchange and Shenzhen Stock Exchange.

I construct the commonly used risk benchmarks using the Chinese stock data. The summary statistics are shown in Table 1. I report results for two sets of sample periods. The whole sample period is from January 1999 to November 2015. The sub-sample period is from May 2005 to October 2015. I use the sub-sample period benchmark factor return series in section 5 when I evaluate the performance of trading strategies.

3. Analyst Performance Evaluation

In this section, I first show some summary statistics of the Chinese equity analyst data. I then discuss my method of calculating the performance of analyst recommendations. Next, I propose several hypotheses on analyst recommendation performance in relation to recommendation characteristics, stock characteristics, and analyst characteristics. I test these hypotheses in a regression framework. Moreover, I offer corroborating evidence on each of the key explanatory variables considered in the regression. Finally, I also discuss some results on how analysts learn on the job.

3.1 Summary Statistics of Analyst Data

The year 2004 has much less coverage than later years. For my analyses later that require sorting analysts or stocks into decile portfolios, I take the sub-sample from 2005 to 2015, in order to ensure sufficient sample size. In addition, my data in 2015 ends in October, so the 2015 sample is underrepresented.

In Table 2.A, I see that a generally increasing trend in analyst coverage of the Chinese stock market. From 2005 to 2014, the number of Chinese equity analysts has increased from 779 to 1,493; the number of analyst reports has increased from 20,127 to 42,451. At the same time, the number of stocks covered by analysts has increased from 827 in 2005 to 2,322 in 2015. In 2005, Chinese equity analysts covered 827 stocks, or 59.07% of all the publicly listed stocks. In 2015, analysts covered 2,322 stocks, or 80.46% of all the publicly listed stocks.

I observe in the last column of Table 2.A that larger stocks are more likely to be covered by Chinese equity analysts. I define a stock's market-capitalization percentile as the percentage of stocks that are smaller than the stock on the report day. 0.00% corresponds to the smallest stock, whereas 100.00% corresponds to the largest stock. I report the equal-weighted average of stocks' market-capitalization percentiles each year. I observe that the average market-capitalization percentile is above 50.00% in all years. This phenomenon is due to two reasons: (1) stocks not at all covered by analysts are small stocks, and (2) among the stocks covered, analysts publish reports more often on larger stock than they do on smaller stocks. As the stock coverage increases from 2005 to 2015, I observe a consistent decreasing trend in the stocks' average market-capitalization percentile.

In Table 2.B, I categorize analyst recommendation types and recommendation changes. Panel A reports summary statistics by recommendation type. There are five recommendation types: type 1 corresponds to the most favorable recommendation (strong buy); type 2 corresponds to the second most favorable recommendation (buy); type 3 corresponds to the neutral recommendation (hold); type 4 corresponds to the second least favorable recommendation (sell); type 5 corresponds to the least favorable recommendation (strong sell). It is striking to see that types 4 and 5 account for a total of 0.90% of the whole sample. Clearly, the overall distribution is heavily tilted towards favorable recommendations. Type-1 and type-2 recommendations account for 85.68% of the whole sample. The recommendation change is more evenly distributed. Upgrades and downgrades account for 6.70% and 6.16% of the whole sample, respectively.

3.2 Definition of Performance

For a given stock covered by an analyst report, I compute a measure for the stock's performance following the report. First, I construct a size-value cohort to which the stock belongs. Specifically, at the end of each June, I sort all the stocks into 5-by-5 size-value cohorts. I measure size by a stock's market capitalization. I measure value by a stock's book-to-market ratio. I then independently sort the stocks into 5 size bins and 5 value bins. Finally, I assign a stock into one out of the 25 size-value cohort depending on which size bin and value bin it belongs to. Next, for a given horizon ranging from 5 trading days (1 week) to 126 trading days (6 months), I calculate the stock's return as well as its size-value cohort's value-weighted return. I define the stock's performance as the difference between the stock's return and its size-value cohort's value-weighted return. In other words, the stock's performance measure can be thought of as an outperformance relative to its respective cohort of stocks.

In unreported results, I also construct other performance proxies and find similar results for analyses conducted in this paper. For example, instead of value-weighted cohort returns, I also consider equal-weighted cohort returns. In addition, I also consider industry in cohort construction by dividing the stocks into 10-by-2-by-2 industry-size-value cohorts. Moreover, I consider an index-hedged performance measure by hedging out the stock's exposure to two major tradable indices: CSI300 and CSI500, which offer a fairly comprehensive coverage of the Chinese stock market by covering the 300 largest stocks and 500 most liquid mid-/small-cap stocks. Finally, I account for the daily return limits in my sample construction. If a stock cannot be traded on the report day because it has reached the daily return limit of either +10% or –10%, I eliminate it from my sample. My results in this paper do not change qualitatively after this adjustment. For interested readers, I will provide additional analyses upon request.

3.3 Hypotheses and Regression Framework

I now propose several hypotheses with regard to performance evaluation of analyst recommendations. I test my hypotheses in a regression framework. I set the independent variable to be the performance proxy I discussed in sub-section 3.2. I test for several different horizons: 5 trading days ($R_{t+1\sim t+5}^{\alpha\_SV}$), 21 trading days ($R_{t+1\sim t+21}^{\alpha\_SV}$), 63 trading days ($R_{t+1\sim t+63}^{\alpha\_SV}$), and 126 trading days ($R_{t+1\sim t+126}^{\alpha\_SV}$). I do not extend my test periods beyond six months because most analysts update their recommendations more frequently than every six months. For example, the median updating interval in my sample is 85 calendar days, less than three months. Next, I choose my explanatory variables to understand relevant factors for performance evaluation.

First and foremost, I want to test if analyst recommendation types have return forecasting powers. More formally, I test the following hypothesis:

> *The more favorable recommendation type a stock receives, the better the stock will perform in the future.*

As there are five recommendation types, ranging from the most favorable type 1 (strong buy) to least favorable type 5 (strong sell), I construct binary/fixed-effect variables for each recommendation type (*rec1, rec2, rec3, rec4, rec5*). I omit type 3 and include type 1, 2, 4 and 5 variables in my regression. I refer to recommendation type 3 as neutral, recommendation types 1 and 2 as favorable, and recommendation types 4 and 5 as unfavorable.

Second, I want to test if analyst recommendation changes have return forecasting powers. More formally, I test the following hypothesis:

> *The more favorable recommendation change a stock receives, the better the stock will perform in the future.*

I define the recommendation change variable (*rec_chg*) as the difference between the last recommendation type and the new recommendation type for a given analyst and stock. For example, if the last recommendation type is 3 and the new recommendation type is 1, then the recommendation change is 2. For another example, if the last recommendation type is 3 and the new recommendation type is 4, then the recommendation change is –1. If there is no change in recommendation type, then the recommendation change is 0. Finally, if there is no previous coverage by the analyst on the stock, then I also set the recommendation change to be 0. I refer to *rec_chg>0* as upgrades, *rec_chg<0* as downgrades, and *rec_chg=0* as neutral or no change.

Third, I want to test if analysts' return-forecasting powers are larger when information asymmetry is more severe by using stock size as a proxy for information asymmetry. More formally, I test the following hypothesis:

> *The value of analyst recommendations is larger for smaller stocks.*

I compute the natural log of stocks' market capitalization: *ln(mktcap)*, and create interaction terms between *ln(mktcap)* and *rec1, rec2, rec4, rec5,* and *rec_chg* respectively. Using these interaction terms as my explanatory variables, I test to see if a favorable recommendation type/change leads to a higher performance for smaller stocks than for larger stocks, and if an unfavorable recommendation type/change leads to a lower performance for smaller stocks than for larger stocks.

In the next two hypotheses, I follow the same intuition that analysts produce more value when they cover stocks with more information asymmetry. In particular, I use stock IPO and the length of a stock's no-coverage period as proxies for information asymmetry. Just after a stock IPO, there is likely a high degree of information asymmetry because few investors are familiar with the stock. Similarly, when a stock has gone a long time without any analyst coverage, fewer investors are up to date about the company, so information asymmetry is high.

Fourth, I want to test for an "absolutely-initial-coverage effect". More formally, I test the following hypothesis:

> *The very first analyst recommendation on a stock since its IPO produces better performance than later analyst recommendations on the stock.*

I compute the absolutely-initial-coverage indicator (*abs_init*) as follows. It takes value 1 if the analyst recommendation on a stock is the very first since the stock's IPO and 0 otherwise. More precisely, I impose a 21-trading-day threshold after the IPO. That is, I start recording the very first analyst report only after 21 trading days from the stock's IPO date. I do this for two reasons: first, there is usually a dark period between IPO date and the date when analysts are allowed to initiate coverage; second, Chinese IPOs tend to have a long string of +10% (upper limit on daily return) returns due to return limits the exchange sets for the stock on its IPO date. By imposing a 21-trading-day filter, I alleviate this concern of mechanical price appreciation. I then create interaction terms between *abs_init* and *rec1, rec2, rec4,* and *rec5* respectively. Using these interaction terms as my explanatory variables, I test to see if the very first coverage on a stock produces a higher/lower performance when the recommendation is favorable/unfavorable than later coverage on the stock.

Fifth, I want to test for a "relatively-initial-coverage effect". More formally, I test the following hypothesis:

> *The first analyst recommendation on a stock that has been without coverage for at least six months produces better performance than other analyst recommendations on the stock.*

I compute the relatively-initial-coverage indicator (*rel_init*) as follows. It takes value 1 if the analyst recommendation on a stock is the very first since the stock's last coverage that must be at least six months before and 0 otherwise. I then create interaction terms between *rel_init* and *rec1, rec2, rec4, rec5,* and *rec_chg* respectively. Using these interaction terms as my explanatory variables, I test to see if the first coverage after a long silent period on a stock produces higher/lower performance when the recommendation is favorable/unfavorable than other coverage on the stock.

Finally, I want to test if analyst experience with a stock affects his return forecasting powers on the stock. More formally, I test the following hypothesis:

> *The longer an analyst has covered a stock, the better performance his recommendations on the stock will have.*

I compute the number of calendar days an analyst has covered a stock up to the report day: *analyst_stock_tenure*, and create interaction terms between *analyst_stock_tenure* and *rec_chg>0* (i.e. indicator for upgrade), and *rec_chg<0* (i.e. indicator for downgrade) respectively. Using these interaction terms as my explanatory variables, I test if longer analyst experience with a stock contributes to better performance, as measured by his upgrade and downgrade decisions.

As a result, I are running the following OLS regression for four different holding periods, i.e. *k=5, 21, 63, or 126 trading days*:

$$
\begin{aligned}
R_{t+1\sim t+k}^{\alpha\_SV} = {} & constant + b_1 * rec1 + b_2 * rec2 + b_3 * rec4 + b_4 * rec5 + c * rec\_chg \\
& + d_1 * ln(mktcap) * rec1 + d_2 * ln(mktcap) * rec2 + d_3 * ln(mktcap) \\
& * rec4 + d_4 * ln(mktcap) * rec5 + d_5 * ln(mktcap) * rec\_chg \\
& + f_1 * abs\_init * rec1 + f_2 * abs\_init * rec2 + f_3 * abs\_init * rec4 \\
& + f_4 * abs\_init * rec5 + g_1 * rel\_init * rec1 + g_2 * rel\_init \\
& * rec2 \qquad + g_3 * rel\_init * rec4 + g_4 * rel\_init * rec5 + g_5 * rel\_init \\
& * rec\_chg \\
& + analyst\_stock\_tenure * (rec\_chg > 0) \\
& + analyst\_stock\_tenure * (rec\_chg < 0) + e_{t+1\sim t+k} \qquad (1)
\end{aligned}
$$

3.4 Regression Analysis

I report regression results in Table 3. Next I will analyze regression results with respect to each of the hypotheses in sub-section 3.3.

3.4.1 Recommendation Type

Across all four horizons, I observe significant outperformance of favorable recommendations over neutral (*rec3* omitted) recommendations. For example, in the 5 days following the analyst recommendation, *rec1* outperforms *rec3* by 1.04% (t=27.84), and *rec2* outperforms *rec3* by 0.50% (t=14.12). This pattern lasts well into 6 months after the analyst recommendation. In the 126 days following the analyst recommendation, *rec1* outperforms *rec3* by 4.03% (t=22.52), and *rec2* outperforms *rec3* by 2.53% (t=14.92). On the other hand, I also observe underperformance of unfavorable recommendations over neutral recommendations. For example, in the 21 days following the analyst recommendation, *rec4* underperforms *rec3* by 0.77% (t=–2.55), and *rec5* underperforms *rec3* by 1.07% (t=–2.40). It is clear that analysts' recommendation types carry return-forecasting powers.

I plot the average cumulative performance in Figure 1. On average, just as the multiple regression results suggest, there is a clear monotonic pattern among recommendation types. Stocks with more favorable recommendation types outperform stocks with less favorable recommendations.

3.4.2 Recommendation Change

Across all four horizons, I observe significantly positive loading on *rec_chg*, which suggests analyst upgrades lead to better performance and analyst downgrades lead to worse performance. For example, in the 5 days following the analyst recommendation, a 1-level upgrade/downgrade leads to 0.34% (t=11.48) better/worse performance. This pattern lasts well into 6 months after the analyst recommendation. In the 126 days following the analyst recommendation, a 1-level upgrade/downgrade leads to 1.01% (t=7.27) better/worse performance. It is clear that analysts' upgrade and downgrade decisions carry return forecasting powers.

In Figure 2, I plot the average cumulative performance of stocks following an upgrade or a downgrade relative to stocks that have had no change in analyst recommendations. On average, just as the multiple regression results suggest, stocks upgraded/downgraded by analysts outperform/underperform stocks that have had no change in analyst recommendations.

3.4.3 Stock Size

Across all four horizons, I observe significantly negative loadings on *ln(mktcap)*'s interaction term with *rec1* and *rec2*, which suggests smaller stocks experience a larger performance improvement following favorable analyst recommendations. I also observe significantly negative loadings on the *ln(mktcap)*'s interaction term with *rec_chg,* which suggests smaller stocks outperform/underperform more than larger stocks following an analyst upgrade/downgrade. It is clear that analyst recommendations provide more value to smaller stocks than to larger stocks. It is intuitive because fewer analysts cover small stocks, which leads to more room for value-added from good research. On the contrary, many analysts cover a big stock, which leads to less room for value-added from good research.

I divide my sample into two halves by the sample-median of stock size: the small-stock half and the large-stock half. In Figure 3a, I plot the average cumulative performance of the small stocks following an upgrade relative to the small stocks that have had no change in analyst recommendations, and the average cumulative performance of the large stocks following an upgrade relative to the large stocks that have had no change in analyst recommendations. In Figure 3b, I plot the counterpart to Figure 3a but with downgrades. On average, just as the multiple regression results suggest, an upgrade in smaller stocks lead to a larger outperformance than larger stocks, whereas a downgrade in smaller stocks lead to a larger underperformance than larger stocks.

3.4.4 Initial Coverage

Across most horizons, I observe significantly positive loadings on *abs_init*'s interaction term with *rec1* and *rec2*, as well as on *rel_init*'s interaction term with *rec1* and *rec2*. These results suggest that initial coverage on a stock (either the absolutely first report on the stock, or the first report on the stock after at least six months without any coverage) exhibit more return-forecasting power. The market reacts to such initial coverage strongly in the short term. For example, in the 5 days following the recommendation, an absolutely initial coverage of type-1 recommendation boosts the stock performance by 5.18% (t=13.58), whereas a relatively initial coverage of type-1 recommendation boosts the performance by 3.41% (t=19.58). This effect lasts into 6-month period. For example, in the 126 days following the recommendation, an absolutely initial coverage of type-1 recommendation boosts the stock performance by 6.92% (t=3.81), whereas a relatively initial coverage of type-1 recommendation boosts the performance by 4.06% (t=4.72). Similar to my intuition for the size effect, I believe that the initial coverage helps capture the degree of information asymmetry surrounding a stock. When a stock has not been covered at all (absolutely initial coverage), or has not been covered for a long time (relatively initial coverage), analyst recommendations add much more value.

3.4.5 Analyst Experience with a Stock

Lastly, I investigate analyst learning on the job by focusing on the explanatory variable *analyst_stock_tenure* that measures the calendar days an analyst has been covering a stock up to the report day. I do not observe significant results across all horizons. From the interaction term with *(rec_chg>0)*, I see that upgrades from more experienced analysts get a stronger market reaction in the 5 days following the analyst recommendation. Specifically, one standard deviation increase in *analyst_stock_tenure* leads to a performance improvement of 0.16% (t=3.52). But this effect is transitory. In the 3 months following the analyst recommendation, it reverts to –0.01% (t=–0.08). From the interaction term with *(rec_chg<0)*, I see that downgrades from more experienced analysts lead to significant underperformance in the 3-month (–0.43%, t=–3.06) and 6-month (–0.55%, t=–2.45) horizons.

In unreported results, I include interaction terms between *analyst_stock_tenure* and *rec1, rec2, rec4, and rec5*. I also replace *analyst_stock_tenure* with *analyst_tenure* that measures the length of analyst tenure independent of the stocks he covers. I observe weak and mixed evidence for analysts' learning on the job. In a separate exercise, I test if analysts' tenure (independent of stock) positively correlates with their performance. For each analyst, I focus on the first report he publishes on a stock, so as to isolate the *analyst_stock_tenure* effect. I do not find evidence that corroborates analyst learning on the job. That is, analysts with a longer tenure do not outperform analysts with a shorter tenure on their first-time coverage of a stock.

4. Analyst Performance Persistence

Now that I have discovered evidence for analyst skill, I want to know if analyst performance is persistent. Specifically, I test to see if analysts that have outperformed in the last period still outperform in the next period. To measure an analyst's performance, I focus on his favorable (rec=1 or 2) and unfavorable (rec=4 or 5) recommendations and ignore his neutral recommendations (rec=3). For the favorable recommendations, I follow them for one period and record the outperformance measured against its size-value cohort's value-weighted returns as alpha. For the unfavorable recommendations, I follow them for one period and record the outperformance measured against its size-value cohort's value-weighted returns, and then take the negative value of the outperformance as alpha. I then calculate the average alpha of the analyst's favorable and unfavorable recommendations to be his alpha.

4.1 Persistence at Monthly Frequency

In month m–1, I follow the methodology above and employ a holding period of one month to compute each analyst's alpha by averaging his recommendations' alphas for the month. I sort these analysts into decile portfolios by their alphas in month m–1. I then follow these decile portfolios of analysts and compute their alphas for month m and m+1. I report results in the top panel of Table 4. Columns under m–1 report the average alpha and t-statistic in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio. Columns under m and m+1 report the average alpha and t-statistic in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio for month m and m+1.

In month m–1, the decile portfolios' alphas exhibit a wide range. The worst performing analyst decile produces a significantly negative monthly alpha of –12.04% (t=–39.02). The best performing analyst decile produces a significantly positive monthly alpha of 19.02% (t=27.17). The long-short analyst portfolio produces an alpha as high as 31.06% (t=33.35).

In the next month m, I still observe a largely monotonic trend in the alpha of the analyst decile portfolios. Outperforming analysts during the past month still tend to outperform one month later. However, the difference in performance variation shrinks significantly. The long-short analyst portfolio now only shows a monthly alpha of 3.13% (t=8.23). In other words, analyst performance tends to mean-revert. This mean-reversion pattern extends into month m+1, when the monotonic pattern in the alpha of the analyst decile portfolios largely disappears. The long-short analyst portfolio now only produces a monthly alpha of 0.63% (t=1.94).

For my definition of alpha, I need to hold the stock for one month in order to compute alpha. Therefore, from an investor's perspective, there is a 1-month lag in computing analyst alpha. The investor does not observe analyst alpha in month m–1 until the end of month m. As a result, the investor can only feasibly trade to extract the alpha in month m+1, which is far weaker than the alpha in month m. For example, in month m, the best performing analyst decile produces an alpha of 3.30% (t=9.81); in month m+1, the best performing analyst decile produces a much smaller alpha of 1.76% (t=5.58).

4.2 Persistence at Weekly Frequency

In week w–1, I follow the methodology above and employ a holding period of one week to compute each analyst's alpha by averaging his recommendations' alphas for the week. I sort these analysts into decile portfolios by their alphas in week w–1. I then follow these decile portfolios of analysts and compute their alphas for week w and w+1. I report results in the bottom panel of Table 4. Columns under w–1 report the average alpha and t-statistic in each decile analyst portfolio, as well as those of a long-short (10–

1) portfolio. Columns under w and w+1 report the average alpha and t-statistic in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio for week w and w+1.

I observe similar mean-reversion patterns at weekly frequency. In week w–1, the long-short analyst portfolio produces an alpha as high as 19.38% (t=57.15). In the next week w, the long-short analyst portfolio only shows a weekly alpha of 0.79% (t=5.17). In the next week w+1, the long-short analyst portfolio now only produces a weekly alpha of 0.41% (t=2.39).

At both monthly and weekly frequencies, despite the fast mean-reversion in analyst performance, I do still observe statistically and economically significant alphas exhibited by the best analyst cohort. For example, at the monthly frequency, an investor can feasibly follow the recommendations of the best analysts in the top decile with a one-month lag and capture a monthly alpha of 1.76% (t=5.58). At the weekly frequency, an investor can feasibly follow the recommendations of the best analysts in the top decile with a one-week lag and capture a weekly alpha of 1.05% (t=7.85). There are at least two caveats that prevent an investor to fully capture these alphas: (1) shorting a stock is costly and sometimes unavailable, and (2) trading costs would damper the returns, especially for the weekly frequency. In the appendix, I extend the evaluation period from m–1 to m–3~m–1, and from w–1 to w–4~w–1, in order to get a longer track record of analysts in computing their alphas. I find similar results, as shown in Table A.1. In the next section, I will formally investigate two trading strategies following analyst recommendations.

5. Evaluation of Trading Strategies

In this section, I form two trading strategies based on analyst recommendations. I show that a simple feasible buy-and-hold strategy can significantly outperform the market and other common risk benchmarks. I also show that a long-short strategy can achieve both a significant alpha and market neutrality.

5.1 Long-Only Trading Strategy

I form a simple and feasible buy-and-hold trading portfolio following analyst recommendations. I focus on the stocks with the most favorable recommendation type (recommendation=1). On the report day, because I do not know if the report is released before or after the market close, I cannot necessarily establish my long positions on that day. But by the next trading day's open, I would have had all the information of reports that came out the day before.

Hence, I form my trading strategy as follows: from January 2005[3] until October 2015, at the open of each trading day, I aggregate all the reports that came out during the last day, and invest one unit of capital equally among the stocks with recommendation=1. If a report came out during the weekend or a holiday (i.e. not a trading day), I process the report and take action at the open of the next trading day. I then hold the stocks in my portfolio for three months, or 63 trading days, before I sell them. No matter how many stocks qualify for my trading strategy on a given trading day, I always invest one total unit of capital split equally among them. If there are no qualified stocks during the last report day, then I do nothing on the trading day.

In forming the trading strategy, I also take into consideration of the daily return limits placed on Chinese stocks. That is, the stocks have a daily return ceiling at +10% and a daily return floor at –10%. If at the open, a qualified stock has already jumped up 10%, and stays that way throughout the day, then I eliminate this stock from my trading portfolio simply because I cannot buy it. If the stock does not stay at +10% throughout the day, then I include that stock in my trading portfolio but replace its open price with its value-weighted average price (vwap) for that day.

I then compute the daily portfolio return series by dividing the portfolio value at the day's close by the portfolio value at the previous day's close. Note that for the positions of unit value established at the day's open, I count it as part of the portfolio value at the previous day's close, in order to facilitate the return calculation. Next, I compound the portfolio's daily returns series to calculate the portfolio's monthly return series. It takes 63 trading days for the portfolio to fully form, by which time there is one unit capital entering and one unit of capital exiting every day. Therefore my fully formed portfolio's monthly return series starts in May 2005.

Next, I employ three benchmark models to evaluate the trading portfolio's performance: CAPM, the Fama and French (1993) 3-factor model (FF3F), and Carhart's

[3] 2004 had many fewer reports than 2005 and onwards. So I start my trading strategy from 2005.

(1997) 4-factor model (FF3F+MOM). To measure performance, these models use variants of the time-series regression

$$R_{pt} - R_{ft} = \alpha_p + b_p\left(R_{mt} - R_{ft}\right) + s_p SMB_t + h_p HML_t + m_p MOM_t + e_{pt} \qquad (2)$$

In this regression, R_{pt} is the return on the trading portfolio for month *t*, R_{ft} is the risk-free rate for month *t*. For the lack of 1-month US Treasury bill rate counterpart in the Chinese Treasury bond market, I use the 3-month deposit rate as a proxy for risk-free rate. R_{mt} is the market return (the return on a value-weighted portfolio of all Chinese domestic stocks), SMB_t and HML_t are the size and value-growth returns as in Fama and French (1993), MOM_t is the Fama-French version of Carhart's (1997) momentum return, α_p is the average return left unexplained by the benchmark model, and e_{pt} is the residual. All factor returns are based on the Chinese stock market data. Details on factor construction are included in Table 1. The regression without MOM_t is the FF3F model. The regression with only $R_{mt} - R_{ft}$ as the only explanatory variable is what I call the CAPM model.

I report the regression results on the top panel of Table 5. Under the CAPM, my trading portfolio produces a significantly positive annual alpha of 11.60% (t=3.09), despite having an indistinguishable-from-one (b=0.99, t=–0.43) loading on the market. Under the 3-factor model, the annual alpha is still significantly positive at 11.29% (t=4.52). I observe a significantly positive loading on SMB_t and significantly negative loading on HML_t. My trading strategy tends to load more on small-cap growth stocks. Under the 4-factor model, the annual alpha is even more significantly positive at 11.69% (t=4.99). I observe a significantly positive loading on MOM_t. My long-only trading strategy tends to chase past winners.

Consistent with results in Section 3, here I offer a different perspective on evaluating the value of Chinese equity analysts' forecasting power reflected by recommendation type. Picking the best-recommended stocks and following a simple buy-and-hold strategy produces a statistically and economically significant alpha for investors. Next, I focus on the analyst upgrades and downgrades and form a long-short trading portfolio.

5.2 Long-Short Trading Strategy

Shorting in the Chinese stock market is costly. Most brokers charge over 8% annualized fees for shorting stocks. Furthermore, the stock exchanges only allow a subset of all publicly listed stocks to be shorted. As a result, this long-short strategy is more of a thought experiment and less of a feasible strategy. Nonetheless, it should still provide valuable insights into the performance evaluation of analysts' upgrade and downgrade decisions.

Similar to the long-only trading portfolio, I form two portfolios here: the long-leg portfolio and the short-leg portfolio. For the long-leg portfolio, I focus on the stocks that had an upgrade, instead of a recommendation of value 1. Then I form the long-leg portfolio exactly as I do the long-only portfolio, as shown in sub-section 5.1. Similarly, for the short-leg portfolio, I focus on the stocks that had a downgrade. Then I form the short-leg portfolio exactly as I do the long-only portfolio, as shown in sub-section 5.1.

To circumvent the return limit problem, I follow the same procedure for the long-leg portfolio as I did the long-only portfolio. I revise the procedure for the short-leg portfolio as follows. If at the open, a downgraded stock has already dropped 10%, and stays that way throughout the day, then I eliminate this stock from my trading portfolio simply because I cannot short it. If the stock does not stay at –10% throughout the day, then I include that stock in my trading portfolio but replace its open price with its value-weighted average price (vwap) for that day.

Next, I compute the daily return series of both the long-leg and the short-leg portfolios. I then define the daily return series of the long-short portfolio by the difference between the long-leg portfolio's daily return and the short-leg portfolio's daily return (i.e. $R_{pt} = R_{long-leg,t} - R_{short-leg,t}$). Finally, I compound the long-short portfolio's daily return series to form its monthly return series.

I report its performance evaluation results on the bottom panel of Table 5. Under the CAPM, my trading portfolio produces a significantly positive annual alpha of 8.10% (t=4.22). Moreover, my long-short portfolio has an indistinguishable-from-zero (b=–0.02, t=–1.24) loading on the market. It is effectively market-neutral. Under the 3-factor model, the annual alpha is more significantly positive at 9.06% (t=4.62). Under the 4-factor model, the annual alpha is even more significantly positive at 9.11% (t=4.74). I observe a significantly positive loading on MOM_t, but not very significant loadings on SMB_t or HML_t. My long-only trading strategy tends to chase past winners.

Consistent with results in Section 3, here I offer a different perspective on evaluating the value of Chinese equity analysts' forecasting power reflected by recommendation change. Loading upgraded stocks and dumping downgraded stocks yield a statistically and economically significant alpha for investors.

5.3 Loading on MOM_t

Despite the fact that the momentum factor is not a priced factor in the Chinese stock market, Chinese equity analysts load positively on it. I see it clearly from the 4-facor model results for both the long-only portfolio and the long-short portfolio. The long-only portfolio's loading on MOM_t is 0.24 (t=4.30). The long-short portfolio's loading on MOM_t is 0.10 (t=2.35). To offer some additional evidence, I investigate the past stock performance rankings in each recommendation type and recommendation change.

I report the summary statistics in Table 6. Panel A summarizes by recommendation type: the stocks' past return percentile (0.00%/100.00% corresponds to the lowest/highest stock returns). Panel B summarizes by recommendation change: the stocks' past return percentile (0.00%/100.00% corresponds to the lowest/highest stock returns). The columns record the return percentiles of the past 5 days, 10 days, 20 days, 40 days, 60 days, and 120 days, respectively. For all horizons considered, I observe a monotonic pattern in stocks' past return ranking. Analysts tend to issue more favorable recommendations to stocks that have performed better in the past. Analysts also tend to upgrade/downgrade stocks that have performed better/worse in the past.

Conceptually, these results are not surprising. On the one hand, it is easy for an analyst to issue a favorable recommendation when the stock has performed well in the past. On the other hand, it is tough for an analyst to act in a contrarian manner to go against the flow. Practically, the positive correlation between analyst recommendations and past stock returns in the short horizon encroaches investment value. This is because in the short horizon (e.g. 20d/1m), the reversal factor is both statistically and economically significant in the Chinese stock market. Table 1's last column echoes this point. For the sample period between 2005 and 2015, average monthly return for the reversal factor is 1.26% (t=3.62). That is, past-month winners significantly underperform past-month losers. Investors should be aware of such short-term-momentum-chasing behavior of analysts when evaluating analyst recommendations.

6. Conclusion

In this paper, I construct a comprehensive novel dataset on Chinese equity analysts from two best-known data providers of Chinese financial data: WIND and GTA. My final dataset offers a 30% sample-size increase to the WIND dataset and 45% sample-size increase to the GTA dataset.

Using this novel dataset, I evaluate the performance of Chinese equity analysts. I find significant return-forecasting powers from analyst recommendations. More favorable recommendations predict better stock performance for at least six months. Analyst upgrades and downgrades also forecast stock performance. Stocks with upgrades outperform stocks with no change in their recommendations. Stocks with downgrades underperform stocks with no change in their recommendations. Consistent with my intuition on the positive correlation between recommendation value and information asymmetry, I find that the value of analyst recommendations is stronger for smaller stocks and for initial coverage on stocks.

Next, I investigate analysts' performance persistence. I find some evidence for performance persistence on both monthly and weekly frequencies. Despite the fact the persistence quickly mean-reverts, an investor can still feasibly follow the best analysts to add alpha.

Finally, I formally investigate two trading strategies following analyst recommendations. The first strategy focuses on the most favorable recommendations and forms a feasible buy-and-hold portfolio. The trading portfolio beats the market by 11.60% (t=3.09) annually. The second strategy focuses on the upgrades/downgrades and forms a long-short portfolio. The trading portfolio is market-neutral and produces an annual alpha of 8.10% (t=4.22).

REFERENCE

Barber, B., Lehavy, R., Mcnichols, M., and Trueman, B., 2001, Can investors profit from the prophets? Security analyst recommendations and stock returns, *Journal of Finance*, 56(2), 531–563.

Clement, M.B., Koonce, L., and Lopez T.J., 2007, The roles of task-specific forecasting experience and innate ability in understanding analyst forecasting performance, *Journal of Accounting and Economics*, 44(3), 378–398.

Deng, Y. and Xu Y., 2011, Do institutional investors have superior stock selection ability in china? *China Journal of Accounting Research*, 4(3), 107–119.

Fama, Eugene F., 1965, The behavior of stock market prices, *Journal of Business* 38, 34–105.

Fama, Eugene F., 1970, Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, 25(2), 383–417.

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

Grossman, Sanford, and Joseph E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *American Economic Review* Vol. 70, No. 3, 393–408.

Jacob, J., Lys, T. Z., & Neale, M. A., 1999, Expertise in forecasting performance of security analysts, *Journal of Accounting and Economics,* 28(1), 51–82.

Jegadeesh, N., and Kim W. (2006). Value of analyst recommendations: International evidence, *Journal of Financial Markets*, 9(3), 274-309.

Jiang, G J., Lu L, and Zhu D., 2014, The Information Content of Analyst Recommendation Revisions– Evidence from the Chinese Stock Market, *Pacific-Basin Finance Journal,* 29(C)*,* 1–17.

Mikhail, M. B., Walther, B. R., & Willis, R. H., 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research*, *35*, 131-157.

Rubin, A., Segal, B., and Segal, D., The interpretation of unanticipated news arrival and analysts' skill, *Journal of Financial and Quantitative Analysis*, forthcoming

Stickel, S. E., 1995, The anatomy of the performance of buy and sell recommendations, *Financial Analyst Journal*, 51(5), 25-39.

Womack, Kent L., 1996, Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance*, 51(1), 137-67.

Xu, N., Jiang X., Chan K. C., and Yi Z., 2013, Analyst coverage, optimism, and stock price crash risk: evidence from China, *Pacific-Basin Finance Journal* 25, 217-239.

Yu, F., 2008, Analyst coverage and earnings management, *Journal of Financial Economics*, 88(2), 245-271.

APPENDIX

Appendix A

Table A.1

Analyst Performance Persistence with Extended Estimation Period

The table reports results on the analyst performance persistence. The top panel reports results at monthly frequency. In months m-3 through m–1, I compute an analyst's performance by averaging his recommendations' performance. Specifically, I focus on the favorable (rec= 1 or 2) and unfavorable (rec=4 or 5) recommendations and ignore the neutral (rec=3) recommendations. For the favorable recommendations (rec=1 or 2), I follow them for one month and record the outperformance measured against its size-value cohort's value-weighted returns as alpha. For the unfavorable recommendations (rec=4 or 5), I follow them for one month and record the outperformance measured against its size-value cohort's value-weighted returns, and then take the negative value as alpha. I then take the average alpha of all favorable and unfavorable recommendations to be the alpha of the analyst for month m–1. I then sort analysts into decile portfolios by their alphas in month m–1. Columns under m–1 report the average alpha and t-stat in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio. I then follow the same decile portfolios of analysts and compute their alphas for month m and m+1. Columns under m and m+1 report the average alpha and t-stat in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio for month m and m+1.

The bottom reports results in a similar fashion, but at weekly frequency, and with an estimation period from week w-4 to week w-1.

Monthly	m-3~m-1		m		m+1	
	alpha	t(alpha)	alpha	t(alpha)	alpha	t(alpha)
1 (lowest)	-9.72	-(39.82)	0.65	(2.81)	1.04	(4.80)
2	-4.86	-(33.40)	0.27	(1.63)	0.85	(4.96)
3	-2.94	-(26.26)	0.45	(2.84)	0.80	(4.71)
4	-1.51	-(16.13)	0.69	(3.75)	0.91	(5.67)
5	-0.24	-(2.70)	0.70	(4.41)	0.85	(5.29)
6	1.04	(10.93)	0.89	(5.33)	0.95	(6.02)
7	2.43	(21.41)	1.02	(5.79)	0.91	(5.06)
8	4.15	(27.94)	1.51	(7.36)	1.12	(5.59)
9	6.74	(31.23)	1.86	(7.65)	1.48	(6.70)
10 (highest)	15.28	(28.94)	2.56	(7.57)	1.54	(4.65)
10-1	25.00	(33.90)	1.92	(5.00)	0.50	(1.40)
Weekly	w-4~w-1		w		w+1	
	alpha	t(alpha)	alpha	t(alpha)	alpha	t(alpha)
1 (lowest)	-6.16	-(72.98)	0.82	(8.86)	0.77	(8.89)
2	-3.15	-(64.54)	0.64	(6.93)	0.84	(9.56)
3	-1.91	-(57.16)	0.68	(7.01)	0.80	(7.66)
4	-1.00	-(40.01)	0.61	(6.45)	0.64	(8.86)
5	-0.18	-(8.12)	0.66	(8.47)	0.61	(7.09)
6	0.64	(24.54)	0.82	(9.88)	0.71	(9.10)
7	1.54	(44.04)	0.61	(8.90)	0.81	(9.43)
8	2.67	(54.74)	0.92	(10.86)	0.78	(9.63)
9	4.37	(58.93)	0.95	(10.79)	0.91	(10.76)
10 (highest)	9.96	(53.71)	1.19	(10.51)	0.98	(9.70)
10-1	16.12	(64.80)	0.36	(2.73)	0.21	(1.90)

Appendix B: Data Dictionary

1. Data Sources

I collect data on Chinese equity analysts from WIND® and GTA®, the two major financial data providers in China. Analyst data in WIND starts from January 2004, whereas GTA starts from January 2001. However, GTA data before January 2004 has relatively few observations compared to data after January 2004. Therefore, I decide to focus on the sample period between January 2004 and October 2015. After deleting observations that have missing values for critical variables (i.e. stock ticker, report date, analyst firm name, analyst, and recommendation type), I summarize the data as follows. As I will show in more details in section 3 of the data dictionary, a significant amount of non-overlapped coverage exists between these two datasets.

	WIND	GTA
# of reports	298,548	266,245
# of analysts	3,563	4,284
# of analyst firms	72	137

2. Data Description

The following table shows the variables I construct for this study.

Variable Name	Variable Label
obs	observation id
rptdt	report date
stkcd	stock ticker
brokername	name of the broker/analyst firm
author1	first author of the report
author2	second author of the report, if available
author3	third author of the report, if available
stdrank	standard rank of the recommendation, aka recommendation type: 1 to 5
rankchg	recommendation change
dsfrom	identifier for the merging step from which the data point is generated

3. Combining Datasets

I go through nine steps to clean, merge, and combine the two datasets to arrive at a complete final dataset I use for my analysis.

Step 1: extract observations with brokername unique to WIND and GTA

WIND and GTA's coverage on analyst firms are not the same. In this step, I focus on extracting the analyst firms that belong uniquely to each dataset. WIND covers 3 analyst firms (*brokername*) that GTA does not cover. GTA covers 65 analyst firms that WIND does not cover. I take the variable values as given because there is no overlap between

WIND and GTA data for this sub-sample. I group those observations unique to WIND into a sub-dataset WINDONLY, which consists of 115 observations. I group those observations unique to GTA into a sub-dataset GTAONLY, which consists of 38,725 observations. Next, I attempt to match overlapped observations from WIND and GTA.

Step 2: merge by five identifiers

If a pair of observations from WIND and GTA has the same value in the following five fields: *stkcd, rptdt, brokername, author1,* and *stdrank*, I consider it a perfect match. I group these matched observations into a sub-dataset FINISH1, which consists of 143,064 observations.

Step 3: merge by four identifiers

If a pair of observations from WIND and GTA has the same value in the following four fields: *stkcd, rptdt, brokername, author1*, but has different values in the field *stdrank*, I consider it a likely match. I then manually check the reports from both sources and choose the correct value for *stdrank*. I group these matched observations into a sub-dataset FINISH2, which consists of 27,962 observations.

Most of these cases come from a few analyst firms. WIND and GTA have different ways of recording *stdrank* from these few analyst firms. I address this difference in a systematic fashion and choose a more reasonable *stdrank* out of the two. I do not elaborate my manual process here but will provide detailed documentation on this step upon request.

Step 4: relax rptdt match to within three days

If a pair of observations from WIND and GTA has the same value in the following four fields: *stkcd, brokername, author1,* and *stdrank*, but has different values in the field *rptdt*, I consider it a likely match. I then relax the match on *rptdt* to within three days of each other. My reasoning for choosing three days as my threshold is that WIND and GTA may record *rptdt* with different lags. For example, if a report comes out on Friday after market close, WIND may capture this report on Friday, whereas GTA may capture this report the next Monday. A lag of up to three days seems a reasonable cutoff. I group these matched observations into a sub-dataset FINISH3, which consists of 2,342 observations. Also for the next steps, I keep the relaxation on the *rptdt* match to within three days.

Step 5: fix author names

If a pair of observations from WIND and GTA has matched values in the following four fields: *stkcd, rptdt, brokername,* and *stdrank*, but has different values in the field *author1*, I consider it a likely match. I then manually check the report to construct *author1* by ourselves and compare with WIND and GTA records. I discover that GTA records analyst names much more precisely than WIND. I decide to keep GTA's version of *author1* instead of WIND's version of *author1*. I group these matched observations into a sub-dataset FINISH4, which consists of 1,889 observations.

Step 6: fix brokername

If a pair of observations from WIND and GTA has matched values in the following four fields: *stkcd, rptdt, author1,* and *stdrank*, but has different values in the field *brokername*, I consider it a likely match. I then manually check these cases. I omit the procedure here but will provide detailed documentation upon request. I group these matched observations into a sub-dataset FINISH5, which consists of 140 observations.

Step 7: fix recommendation types

With *rptdt* matching relaxed to within 3 days, I re-implement step 3. I are able to produce 427 more observations into a sub-dataset FINISH6.

Step 8: fix recommendation types and analyst names

If a pair of observations from WIND and GTA has matched values in the following three fields: *stkcd, rptdt,* and *stdrank*, but has different values in the fields *author1* and *brokername*, I consider it a likely match. I then manually check these cases. I omit the procedure here but will provide detailed documentation upon request. I group these matched observations into a sub-dataset FINISH7, which consists of 63 observations.

Step 9: process remaining unmatched observations

After my exhaustive matching steps 2 to 8, the remaining observations are non-overlapping between WIND and GTA. In other words, they are unique to WIND or GTA. I group those observations unique to WIND into a sub-dataset WINDONLY2, which consists of 122,313 observations. I group those observations unique to GTA into a sub-dataset GTAONLY2, which consists of 49,063 observations.

The following table summarizes the nine data steps.

Step	Sub-dataset	Count
1	WINDONLY	115
	GTAONLY	38,725
2	FINISH1	143,064
3	FINISH2	27,962
4	FINISH3	2,342
5	FINISH4	1,889
6	FINISH5	140
7	FINISH6	427
8	FINISH7	63
9	WINDONLY2	122,313
	GTAONLY2	49,063
SUM		386,103

4. Final Dataset

The following table shows the summary statistics of the final dataset.

	WIND	GTA	FINAL
# of reports	298,548	266,245	386,103
# of analysts	3,563	4,284	4,492
# of firms	72	134	137

Figures

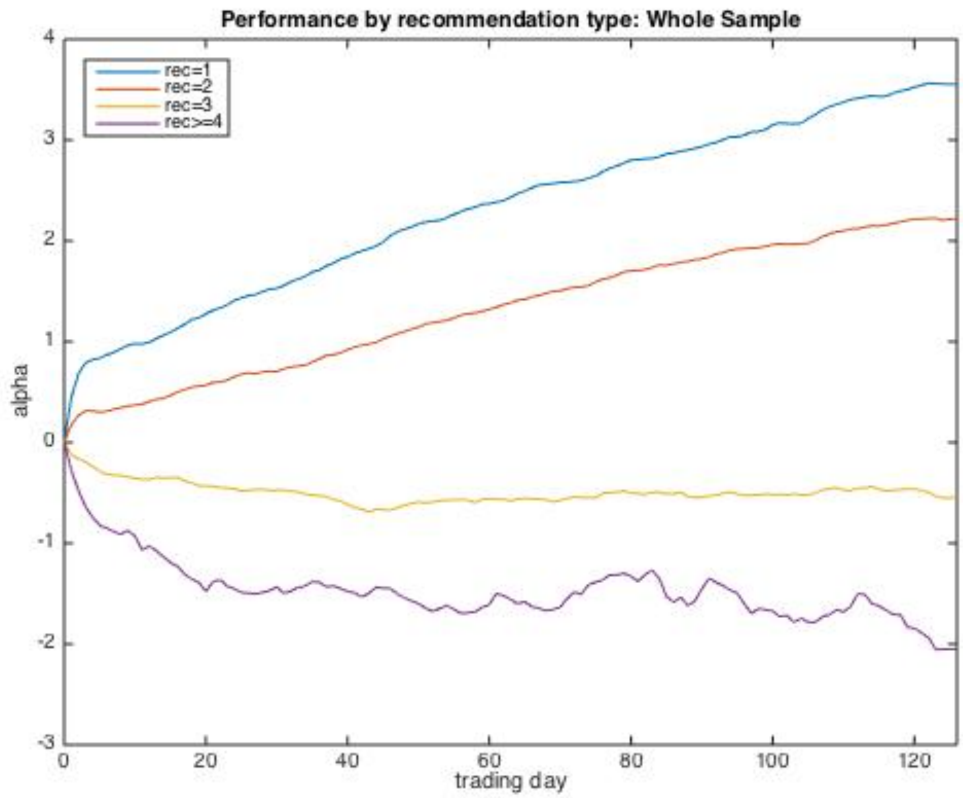

Figure 1: Average cumulative performance for stocks with rec=1, 2, 3, and >=4 for up to 126 days after analyst recommendations.

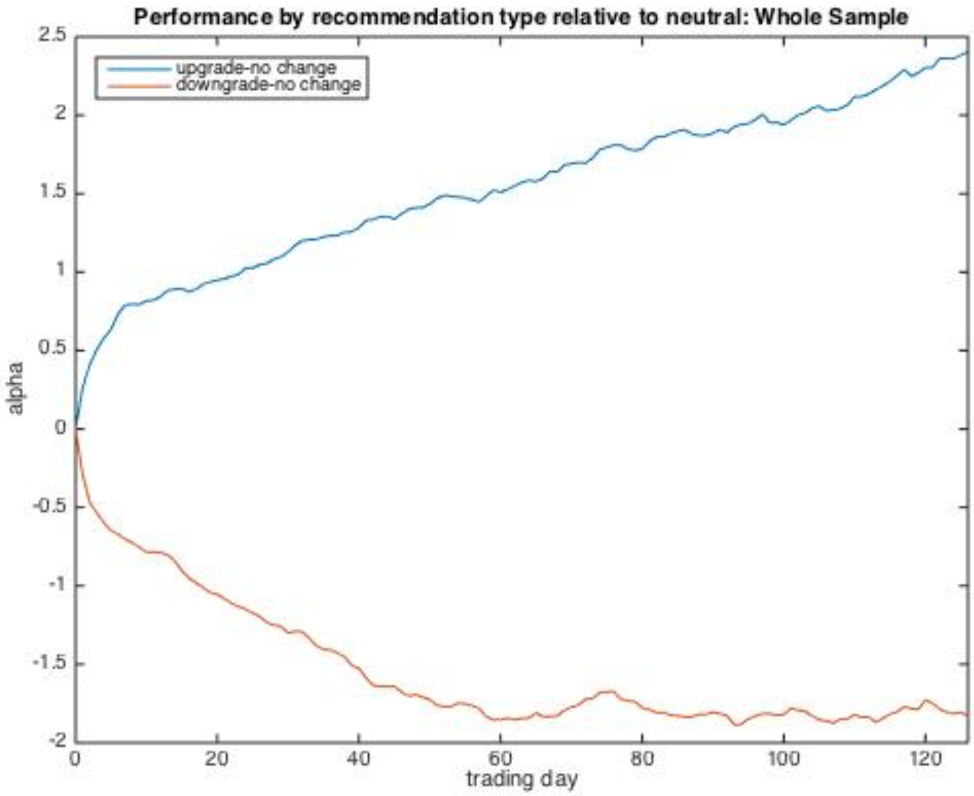

Figure 2: Average cumulative performance for stocks with rec_chg>0 (upgrade) relative to stocks with rec_chg=0 (no change), and stocks with rec_chg<0 (downgrade) relative to stocks with rec_chg=0 (no change), for up to 126 days after analyst recommendations.

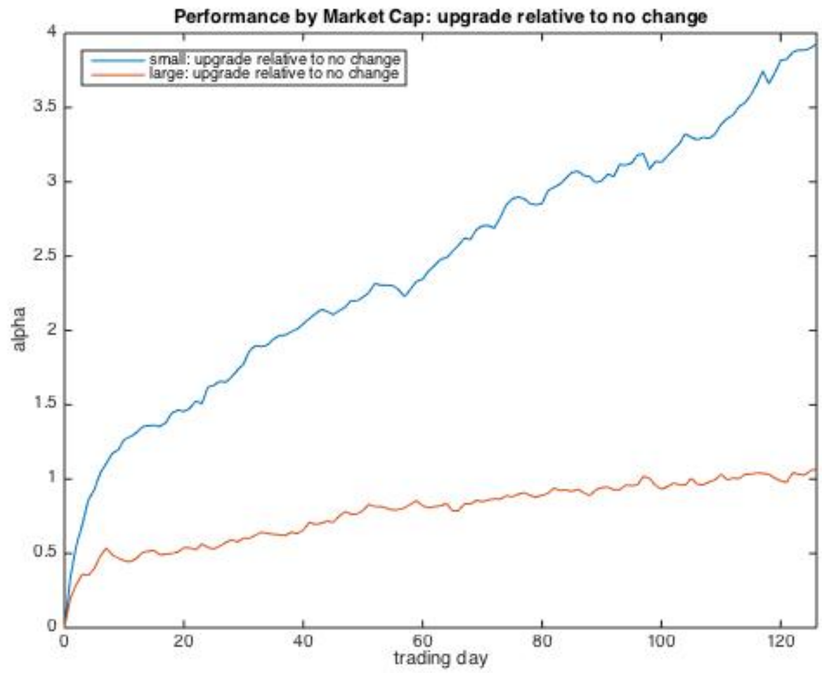

Figure 3a: Average cumulative performance for small/large stocks with rec_chg>0 (upgrade) relative to small/large stocks with rec_chg=0 (no change), for up to 126 days after analyst recommendations.

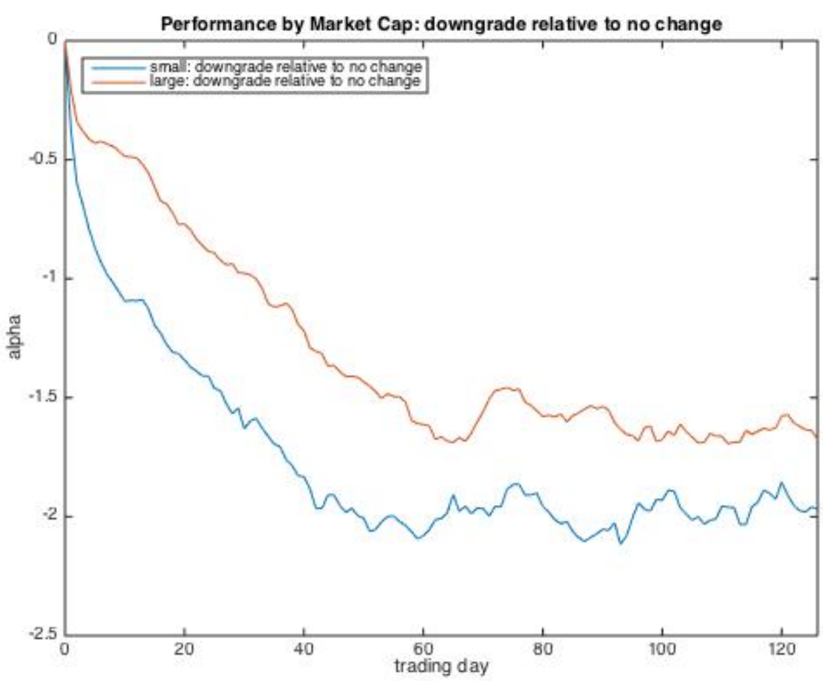

Figure 3b: Average cumulative performance for small/large stocks with rec_chg<0 (downgrade) relative to small/large stocks with rec_chg=0 (no change), for up to 126 days after analyst recommendations.

Tables

Table 1
Monthly Benchmark Factor Returns in the Chinese Stock Market

For the market risk premium $R_m - R_f$, R_m is taken as the value-weighted one-month return on stocks publicly listed on the Shenzhen A and Shanghai A stock exchanges, which represent all eligible stocks for Chinese stock mutual funds. Weights are monthly market-cap values. R_f is the risk free return, proxied by the 3-month Chinese household savings deposit rate. Since this rate is reported as an annual rate, I divide it by 12 to get a monthly R_f. Finally, the excess market return factor was constructed as the market return R_m less the risk free rate R_f.

For the computation of *SMB* and *HML*, each stock is categorized as "big" or "small" based on whether it is above or below the median market-cap. Stocks are also classified as "high", "medium" or "low" BE/ME ratio based on June BE/ME ratio for each stock. Stocks with BE/ME ratios in the top 30th percentile of all BE/ME ratios for publicly listed Chinese A stocks were classified as "high", while stocks with BE/ME ratios in the bottom 30th percentile were classified as "low". Stocks with BE/ME ratios in the middle 40 percentiles (30th to 70th percentile) were classified as "medium". Six portfolios were formed annually, i.e. Small/High, Small/Medium, Small/Low, Big/High, Big/Medium, and Big/Low. The value-weighted monthly returns for each portfolio were computed using monthly market-cap data, and the monthly factors are determined as follows: *SMB* is just the equal-weighted average of returns on the "Small" portfolios minus the equal-weighted average of returns on the "Big" portfolios. *HML* is similarly the equal-weighted average of returns on the "High" portfolios minus the equal-weighted average of returns on the "Low" portfolios.

The momentum factor (*MOM*) and reversal factor (REV) were constructed by forming six portfolios monthly, using monthly market-cap to construct small and big portfolios much like in the computation of *SMB* and *HML*. However, for the momentum and reversal factors, the size portfolios are formed monthly instead of annually. For the momentum factor, the total return from 12 months prior to 2 months prior is computed for each stock. Monthly momentum portfolios are formed based on this prior return measure, with the bottom 30th percentile of stocks (i.e. those stocks with the lowest return from 12 months ago to 2 months ago) being classified as "low", and the top 30th percentile of prior return stocks being classified as "high". For the reversal factor, the total return of the last months less the last trading day is computed for each stock. Monthly reversal portfolios are formed based on this prior return measure, with the bottom 30th percentile of stocks (i.e. those stocks with the lowest return from 12 months ago to 2 months ago) being classified as "low", and the top 30th percentile of prior return stocks being classified as "high". The middle 40 percent, from 30th percentile to 70th percentile, are classified as "medium" reversal stocks. Then I form six portfolios by intersecting the momentum and reversal portfolios with the size portfolios. The monthly momentum factor itself is MOM = 1/2 *(return on Big/High + return on Small/High) – 1/2 *(return on Big/Low + return on Small/Low). The monthly reversal factor itself is REV = 1/2 *(return on Big/Low + return on Small/Low) – 1/2 *(return on Big/High + return on Small/High).

Finally, the whole sample period is July 1998 to November 2015. The sub-sample period is May 2005 to October 2015, which I use for the regression on evaluating the performance of the trading strategy.

Average Monthly Return

Sample Period	Rm-Rf	SMB	HML	MOM	REV
01/1999 ~ 11/2015	1.00%	0.85%	0.54%	-0.15%	0.97%
	(1.67)	(2.70)	(2.08)	-(0.58)	(3.92)
05/2005 ~ 10/2015	1.65%	1.22%	0.44%	-0.50%	1.26%
	(2.02)	(2.72)	(1.17)	-(1.37)	(3.62)

Table 2.A
Summary Statistics of the Chinese Equity Analysts

The table summarizes by year: the number of analyst reports, analyst firms, analysts, and stocks covered by analysts; the percentage of stocks covered out of the total number of stocks traded; and the covered stocks' average market capitalization percentile (0.00%/100.00% corresponds to the smallest/largest stock). My data coverage of 2015 is till the end of October.

year	# of reports	# of analyst firms	# of analysts	# of stocks covered	% of stock coverage	stocks' average mktcap percentile
2004	3,897	50	476	636	45.92	77.28%
2005	20,127	70	779	827	59.07	79.28%
2006	35,064	56	756	962	65.62	78.98%
2007	21,094	61	998	1,062	66.71	77.60%
2008	27,440	72	1,371	1,090	65.31	77.83%
2009	31,788	82	1,467	1,324	74.89	74.81%
2010	34,678	87	1,531	1,651	77.99	70.87%
2011	43,984	89	1,317	1,934	80.62	67.29%
2012	46,822	86	1,448	1,960	76.74	66.99%
2013	43,651	80	1,536	1,764	69.01	69.72%
2014	42,451	77	1,493	1,966	73.33	69.60%
2015	35,107	75	1,292	2,322	80.46	67.64%

Table 2.B
Summary Statistics of the Analyst Recommendations

Panel A summarizes by recommendation type: the number of analyst reports, and the percentage out of the whole sample. There are five recommendation types: type 1 corresponds to the most favorable recommendation (strong buy); type 5 corresponds to the least favorable recommendation (strong sell). Panel B summarizes by recommendation change: the number of analyst reports, and the percentage out of the whole sample. A recommendation change can be upgrade, no change, or downgrade.

Panel A: by recommendation type

	# of reports	% of sample
1(best)	152,776	39.57
2	178,044	46.11
3	51,809	13.42
4	2,589	0.67
5(worst)	885	0.23

Panel B: by recommendation change

	# of reports	% of sample
upgrade	25,869	6.70
no change	336,438	87.14
downgrade	23,796	6.16

Table 3
Summary Results of Regression (1)

The table reports the coefficients, t-stats and R-square for regression (1) estimated on horizons of 5 trading days or 1 week, 21 trading days or 1 month, 63 trading days or 3 months, and 126 trading days or 6 months. Explanatory variables rec1/rec2/rec4/rec5 takes value 1 if recommendation type is 1/2/4/5 and 0 otherwise; rec3 is omitted; rec_chg is the difference between the new recommendation type and the old recommendation type (positive for upgrades, negative for downgrades, zero if no change); ln(mktcap) is the natural logarithm of the market capitalization at the report date; abs_init takes value 1 if the report is the first on a stock after one month since its IPO, and takes value 0 for all reports on the stock thereafter; rel_init takes value 1 if the report is the first on a stock after the stock has had no coverage for the past six months, and takes value 0 otherwise; rec_chg>0 is an indicator for upgrades; rec_chg<0 is an indicator for downgrades; analyst_stock_tenure measures the number of days a given analyst has covered a given stock. I further normalize the variables ln(mktcap) and analyst_stock_tenure so that they are standard normal. All t-stats greater than 1.96 are marked in yellow. All t-stats less than –1.96 are marked in red. The sample period is January 2004 through October 2015. The sample size is 386,103.

	t+1~t+5		t+1~t+21		t+1~t+63		t+1~t+126	
	coef	(t-stat)	coef	(t-stat)	coef	(t-stat)	coef	(t-stat)
constant	-0.22	-(6.91)	-0.33	-(5.92)	-0.41	-(4.26)	-0.33	-(2.18)
rec1	1.04	(27.84)	1.65	(25.13)	2.91	(25.65)	4.03	(22.52)
rec2	0.50	(14.12)	0.90	(14.45)	1.78	(16.56)	2.53	(14.92)
rec4	-0.48	-(2.78)	-0.77	-(2.55)	-1.12	-(2.17)	-2.35	-(2.88)
rec5	-0.86	-(3.40)	-1.07	-(2.40)	-0.92	-(1.20)	-1.74	-(1.45)
rec_chg	0.34	(11.48)	0.55	(10.73)	0.88	(9.92)	1.01	(7.27)
ln(mktcap)*rec1	-0.44	-(20.90)	-0.79	-(21.60)	-1.31	-(20.53)	-2.01	-(19.81)
ln(mktcap)*rec2	-0.17	-(8.71)	-0.39	-(11.59)	-0.66	-(11.39)	-0.92	-(10.10)
ln(mktcap)*rec4	-0.37	-(2.33)	-0.27	-(0.97)	-0.84	-(1.74)	-1.87	-(2.45)
ln(mktcap)*rec5	-0.31	-(1.28)	0.07	(0.16)	-0.24	-(0.33)	1.29	(1.10)
ln(mktcap)*rec_chg	-0.16	-(4.91)	-0.23	-(4.13)	-0.29	-(3.01)	-0.48	-(3.19)
abs_init*rec1	5.18	(13.58)	6.21	(9.28)	6.52	(5.66)	6.92	(3.81)
abs_init*rec2	0.50	(1.75)	1.40	(2.79)	2.26	(2.62)	2.72	(2.00)
abs_init*rec4	-1.18	-(1.26)	-5.61	-(3.41)	-3.64	-(1.29)	-3.86	-(0.87)
abs_init*rec5	-1.06	-(0.70)	-1.80	-(0.68)	-1.83	-(0.40)	-0.89	-(0.12)
rel_init*rec1	3.41	(19.58)	4.91	(16.05)	4.11	(7.74)	4.06	(4.72)
rel_init*rec2	1.12	(8.92)	1.26	(5.73)	0.95	(2.50)	0.71	(1.17)
rel_init*rec4	-1.25	-(1.44)	-2.05	-(1.34)	-2.83	-(1.08)	-1.40	-(0.34)
rel_init*rec5	-2.19	-(1.50)	1.09	(0.42)	-0.01	(0.00)	6.36	(0.92)
rel_init*rec_chg	-0.40	-(1.74)	-0.69	-(1.73)	-0.94	-(1.35)	-1.79	-(1.64)
analyst_stock_tenure*(rec_chg>0)	0.16	(3.52)	0.10	(1.24)	-0.01	-(0.08)	0.19	(0.88)
analyst_stock_tenure*(rec_chg<0)	-0.04	-(0.96)	-0.06	-(0.80)	-0.43	-(3.06)	-0.55	-(2.45)
R-square	0.71%		0.58%		0.46%		0.36%	

Sample Size: 386,103

Table 4
Analyst Performance Persistence

The table reports results on the analyst performance persistence. The top panel reports results at monthly frequency. In month m–1, I compute an analyst's alpha by averaging his recommendations' alphas. In particular, I focus on the favorable (rec=1 or 2) and unfavorable (rec=4 or 5) recommendations and ignore the neutral recommendations (rec=3). For the favorable recommendations (rec=1 or 2), I follow them for one month and record the outperformance measured against its size-value cohort's value-weighted returns as alpha. For the unfavorable recommendations (rec=4 or 5), I follow them for one month and record the outperformance measured against its size-value cohort's value-weighted returns, and then take the negative value as alpha. I then take the average alpha of all favorable and unfavorable recommendations to be the alpha of the analyst for month m–1. I then sort analysts into decile portfolios by their alphas in month m–1. Columns under m–1 report the average alpha and t-stat in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio. I then follow the same decile portfolios of analysts and compute their alphas for month m and m+1. Columns under m and m+1 report the average alpha and t-stat in each decile analyst portfolio, as well as those of a long-short (10–1) portfolio for month m and m+1. Alphas are reported in percentage.
The bottom reports results in a similar fashion, but at weekly frequency.

Monthly	**m-1**		**m**		**m+1**	
	alpha	t(alpha)	alpha	t(alpha)	alpha	t(alpha)
1 (lowest)	-12.04	-(39.02)	0.17	(0.69)	1.13	(4.86)
2	-6.66	-(31.54)	0.11	(0.57)	0.91	(4.51)
3	-4.23	-(25.04)	0.42	(2.63)	0.78	(3.92)
4	-2.36	-(15.81)	0.58	(3.24)	0.83	(4.31)
5	-0.64	-(4.55)	0.72	(4.45)	1.06	(6.11)
6	1.09	(7.78)	1.05	(5.58)	1.14	(6.81)
7	2.97	(19.04)	1.35	(7.49)	1.01	(5.61)
8	5.25	(26.87)	1.56	(7.89)	1.03	(5.10)
9	8.65	(30.94)	1.82	(7.75)	1.19	(5.29)
10 (highest)	19.02	(27.17)	3.30	(9.81)	1.76	(5.58)
10-1	31.06	(33.35)	3.13	(8.23)	0.63	(1.94)
Weekly	**w-1**		**w**		**w+1**	
	alpha	t(alpha)	alpha	t(alpha)	alpha	t(alpha)
1 (lowest)	-7.30	-(66.90)	0.59	(5.42)	0.61	(5.42)
2	-4.02	-(60.42)	0.67	(5.67)	0.76	(6.70)
3	-2.52	-(48.99)	0.68	(3.57)	0.52	(4.63)
4	-1.39	-(31.94)	0.44	(4.12)	0.80	(6.87)
5	-0.38	-(9.14)	0.76	(6.53)	0.56	(4.76)
6	0.64	(14.35)	0.66	(5.50)	0.59	(5.31)
7	1.79	(33.45)	0.72	(7.42)	0.83	(7.55)
8	3.24	(45.68)	0.69	(5.84)	0.79	(6.88)
9	5.42	(50.58)	1.02	(8.56)	0.82	(5.69)
10 (highest)	12.08	(43.76)	1.36	(11.54)	1.05	(7.85)
10-1	19.38	(57.15)	0.79	(5.17)	0.41	(2.39)

Table 5
Performance Evaluation of Trading Strategies

The table shows the annualized intercepts (12*α) and *t*-statistics for the intercept (*t*-stat) for the CAPM, three-factor, and four-factor versions of regression (2) estimated on the long-only (rec=1) and long-short (upgrade-downgrade) trading strategies' returns. The table also shows the slopes for factors. For the market slope in the top panel, *t*-stat tests whether *b* is different from 1.0, instead of 0. The period is May 2005 through October 2015.

	12*α	b	s	h	m	R-sq
Long-Only (rec=1)						
CAPM	11.60%	0.99				0.87
	(3.09)	(-0.43)				
3-Factor	11.29%	1.04	0.15	-0.57		0.95
	(4.52)	(1.79)	(3.76)	(-11.41)		
4-Factor	11.69%	1.04	0.18	-0.44	0.24	0.95
	(4.99)	(1.80)	(4.60)	(-7.71)	(4.30)	
Long-Short (upgrade-downgrade)						
CAPM	8.10%	-0.02				0.01
	(4.22)	(-1.24)				
3-Factor	9.06%	-0.02	-0.06	-0.04		0.05
	(4.62)	(-0.88)	(-2.02)	(-0.92)		
4-Factor	9.11%	-0.02	-0.05	0.03	0.10	0.09
	(4.74)	(-0.97)	(-1.73)	(0.57)	(2.35)	

Table 6
Analyst Recommendation and Past Stock Returns

Panel A summarizes by recommendation type: the stocks' past return percentile (0.00%/100.00% corresponds to the lowest/highest stock returns). The columns record the return percentiles of the past 5 days, 10 days, 20 days, 40 days, 60 days, and 120 days, respectively.
Panel B summarizes by recommendation change: the stocks' past return percentile (0.00%/100.00% corresponds to the lowest/highest stock returns). The columns record the return percentiles of the past 5 days, 10 days, 20 days, 40 days, 60 days, and 120 days, respectively.

Panel A: past stock performance ranking by recommendation type

	past5d	past10d	past20d	past40d	past60d	past120d
1 (highest)	55.15%	55.07%	55.36%	55.66%	56.20%	58.50%
2	53.81%	53.45%	53.56%	53.34%	53.67%	55.00%
3	50.73%	50.34%	50.01%	48.91%	48.60%	48.25%
4	48.89%	48.56%	49.06%	47.63%	47.01%	47.07%
5 (lowest)	47.50%	47.59%	46.80%	45.04%	44.43%	43.87%

Panel B: past stock performance ranking by recommendation change

	past5d	past10d	past20d	past40d	past60d	past120d
upgrade	56.12%	56.15%	56.70%	55.96%	55.97%	56.89%
no change	53.84%	53.60%	53.71%	53.68%	54.03%	55.59%
downgrade	51.97%	51.27%	51.21%	50.11%	50.37%	51.16%