

Rankings of financial analysts as means to profits

Artur Aiguzhinov (artur.aiguzhinov@inescporto.pt)^{1,2}, Ana Paula Serra (aserra@fep.up.pt)¹, and Carlos Soares (csoares@fe.up.pt)^{2,3}

¹FEP & CEF.UP, University of Porto

²INESC TEC

³FEUP, University of Porto

November 24, 2015

Abstract

Financial analysts are evaluated based on the value they create for those who follow their recommendations. Some institutions use these evaluations to rank the analysts. Predictions of the most accurate analysts are typically modeled in terms of individual analyst's characteristics. The disadvantage of this approach is that these data are hard to collect. In this paper, we follow a different approach in which we characterize analysts' relative performance on the basis of state variables rather than individual characteristics. We apply a label ranking model to predict the rankings and use them to develop active trading strategies. The results show that it is possible to accurately model the relation between selected attributes and rankings of analysts. In addition, the strategy based on the predicted rankings applied to the S&P500 stocks generates higher returns than that based on the analysts' consensus.

1 Introduction

Rankings of financial analysts is not new in finance. Many agencies develop their procedures to evaluate analysts based on their performance either in forecasting or stock recommendations. Some institutions even hold a “Red Carpet” event to recognize the top analysts. On one hand, for market participants, the rankings may signal who is the best analysts. On the other hand, studies have shown that following the best analysts’ recommendations of buy-sell stocks have statistically insignificant benefits.

In this paper we have an objective to show that rankings can serve as inputs for trading; that is, they can be a direct input for strategy. We base our research on the main assumption that analysts at the top ranks are the best analyst and they are worth to be followed. Naturally, instead of relying on personal expertise in selecting the expected returns, it is best to refer to the specialist in the field, namely, the financial analysts. Using the analysts’ price target information we create a vector of expected returns and use it as a starting point for appropriate trading strategy.

Given the objective of the paper, we have to solve two problems. First, we need to predict the rankings of the analyst and, second, translate these rankings into an operational input for the trading strategy.

For the first problem, we take advantage of the algorithm used for prediction of rankings developed in Aiguzhinov, Soares, and Serra (2010) and adapt it for the case of analysts’ rankings. In short, this algorithm is based on the Bayesian probability and the similarities between the rankings. The solution for the second problem relies on the Black-Litterman (BL) model (Black and Litterman, 1992).

We are particularly confident in this choice of tools given that the BL model and ranking algorithm are both based on the Bayesian framework. Given the results from Aiguzhinov, Serra, and Soares (2015a), we base our rankings on analysts' target prices as it has been shown that strategies based on these rankings are the ones that yield the highest cumulative annualized return.

The paper is organized as follows: **Section 2** provides motivation on use of rankings; **Section 3** outlines label ranking algorithm; **Section 4** summarizes the selection of the independent variables that affect the analysts' rankings; **Section 5** describes a methodology of building the rankings; **Section 6** discusses information environments that influence analysts' decisions; **Section 7** outlines the steps of the trading strategy; **Section 8** describes data used in the study; **Section 9** discusses the results, and **Section 10** concludes.

2 Rankings of financial analysts

In financial literature there has been a long debate on whether financial analysts produce valuable advice. Some argue that following the advice of financial analysts, translated as recommendations of buying, holding, or selling a particular stock, does not yield abnormal returns, i.e., returns that are above the required return to compensate for risk. The Efficient Market Hypothesis (Fama, 1970) states that financial markets are efficient and that any public available information regarding a stock would be immediately reflected in prices; hence, it would be impossible to generate abnormal returns based upon past information.

Yet, several authors have since stressed that there are information-gathering costs and information is not immediately reflected on prices (Grossman and Stiglitz,

1980). As such, prices may not reflect all the available information at all time because if this were the case, those who spent resources to collect and analyze information would not have an incentive to do it, because there would not get any compensation for it.

Many trading strategies try to forecast the price movements relying on the historical prices or estimate the intrinsic value of a company. Obviously, this type of research is associated with significant amount of up-front costs to acquire databases, software, etc. On the other hand, financial analysts have these tools and, presumably, skills to identify stocks that worth be invested. Thus, for an investor, it is cheaper to follow the recommendations of financial analysts rather than perform a proper stock market analysis.

Some authors show that financial analysts' recommendations create value to investors (Womack, 1996; Barber, Lehavy, McNichols, and Trueman, 2001)¹. Assuming that some analysts produce valuable advice it makes sense to rank analysts based on the accuracy of their recommendations.

StarMine rankings are based on financial analysts' accuracy either on TP or EPS forecasts. To rank analysts based on EPS forecasts, StarMine developed a proprietary metric called a Single-stock Estimating Score (SES). This score measures "... [a] relative accuracy; that is, analysts are compared against their peers. An analyst's SES can range from 0 to 100, with 50 representing the average analyst. To get a score higher than 50, an analyst must make estimates that are both significantly different from and more accurate than other analysts' estimates"².

¹Womack (1996) finds that post-recommendation excess returns are not mean-reverting, but are significant and in the direction forecast by the analysts. Barber et al. (2001) finds that over the period of 1986-1996 a portfolio of stocks with the most (least) favorable consensus analyst recommendations yields an average abnormal return of 4.13 (-4.91)%.

²<http://excellence.thomsonreuters.com/award/starmine?award=>

As for target price ranking, StarMine’s methodology compares the portfolios based on analysts recommendations. Portfolios are constructed as follows. For each “Buy” recommendation, the portfolio is one unit long the stock and simultaneously one unit short the benchmark. “Strong buy” gets a larger investment of two units long the stock and two units short the benchmark. “Hold” invests one unit in the benchmark (i.e., an excess return of zero). “Sell” recommendations work in the reverse way. StarMine re-balances its calculations at the end of each month to adjust for analysts revisions (adding, dropping or altering a rating), and when a stock enters or exits an industry grouping.

Recent evidence suggests that top ranked financial analyst affect market participants: prices seem to react more to the recommendations issued by the top-ranked analysts (Emery and Li, 2009). As such, StarMine ranking based models can be used to identify such analysts and generate superior estimates (e.g., SmartEstimates³).

The goal of our study is to predict StarMine rankings. With this purpose, we adapt a Machine Learning algorithm to predict rankings given a set of variables that characterize these rankings. We, further, apply the predicted rankings to build active trading strategies to evaluate quality of predictions against the consensus strategy (giving equal weights to analysts’ recommendations).

Analyst+Awards&award_group=Overall+Analyst+Awards

³http://www.starmine.com/index.phtml?page_set=sm_products&sub_page_set=sm_professional&topic=analytical§ion=accurate_estimates

3 Label ranking algorithm

The classical formalization of a label ranking problem is the following (Vembu and Gärtner, 2010). Let $\mathcal{X} = \{\mathcal{V}_1, \dots, \mathcal{V}_m\}$ be an instance space of variables, such that $\mathcal{V}_a = \{v_{a,1}, \dots, v_{a,n_a}\}$ is the domain of nominal variable a . Also, let $\mathcal{L} = \{\lambda_1, \dots, \lambda_k\}$ be a set of labels, and $\mathcal{Y} = \Pi_{\mathcal{L}}$ be the output space of all possible total orders over \mathcal{L} defined on the permutation space Π . The goal of a label ranking algorithm is to learn a mapping $h : \mathcal{X} \rightarrow \mathcal{Y}$, where h is chosen from a given hypothesis space \mathcal{H} , such that a predefined loss function $\ell : \mathcal{H} \times \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ is minimized. The algorithm learns h from a training set $\mathcal{T} = \{x_i, y_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{X} \times \mathcal{Y}$ of n examples, where $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,m}\} \in \mathcal{X}$ and $y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,k}\} \in \mathcal{Y}$. With time-dependent problem in rankings, we replace the i index with t ; that is $y_t = \{y_{t,1}, y_{t,2}, \dots, y_{t,k}\}$ is the ranking of k labels at time t described by $x_t = \{x_{t,1}, x_{t,2}, \dots, x_{t,m}\}$ at time t .

Consider an example of a time-dependent ranking problem presented in **Table 3**. In this example, we have three brokers ($k = 3$), four independent variables ($m = 4$) and a period of 7 quarters. Our goal is to predict the rankings for period t , given the values of independent variables and rankings known up to period $t - 1$; that is, to predict the ranking for time $t = 7$, we use $n = 6$ ($t \in \{1 \dots 6\}$) examples to train the ranking model.

3.1 Naive Bayes algorithm for label ranking

The naive Bayes for label ranking (NBLR) will output the ranking with the higher $P_{LR}(y|x)$ value (Aiguzhinov et al., 2010):

$$\begin{aligned}\hat{y} &= \arg \max_{y \in \Pi_{\mathcal{L}}} P_{LR}(y|x) = \\ &= \arg \max_{y \in \Pi_{\mathcal{L}}} P_{LR}(y) \prod_{i=1}^m P_{LR}(x_i|y)\end{aligned}\tag{1}$$

where $P_{LR}(y)$ and $P_{LR}(x_i|y)$ are the prior and conditional label ranking probabilities of a nominal variable x of attribute a , (v_a), respectively, and they are given as follows:

$$P_{LR}(y) = \frac{\sum_{i=1}^n \rho(y, y_i)}{n}\tag{2}$$

$$P_{LR}(x_i|y) = \frac{\sum_{i: x_a = v_a} \rho(y, y_i)}{|\{i : x_a = v_a\}|}\tag{3}$$

where $\rho(y, y_i)$ is the similarity between rankings obtained from the Spearman ranking correlation:

$$\rho(y, y_i) = 1 - \frac{6 \sum_{j=1}^k (y - y_{i,j})^2}{k^3 - k}\tag{4}$$

Similarity and probability are different concepts; however, a connection has been established between probabilities and the general Euclidean distance measure (Vogt, Godden, and Bajorath, 2007). It states that maximizing the likelihood is equivalent to minimizing the distance (i.e., maximizing the similarity) in a Euclidean space. The predicted ranking for an example x_i is the one that will receive

the maximum posterior label ranking probability $P_{LR}(y|x_i)$.

3.1.1 Continuous independent variables

In its most basic form, the naive Bayes algorithm cannot deal with continuous attributes. The same happens with its adaptation for label ranking (Aiguzhinov et al., 2010). However, there are versions of the naive Bayes algorithm for classification that support continuous variables (Bouckaert, 2005). The authors modify the conditional label ranking probability by utilizing the Gaussian distribution of the independent variables. We apply the same approach in defining the conditional probability of label rankings:

$$P_{LR}(x|y) = \frac{1}{\sqrt{2\pi}\sigma(x|y)} e^{-\frac{(x-\mu(x|y))^2}{2\sigma^2(x|y)}} \quad (5)$$

where $\mu(x|y)$ and $\sigma^2(x|y)$ weighted mean and weighted variance, defined as follows:

$$\mu(x|y) = \frac{\sum_{i=1}^n \rho(y, y_i) x}{\sum_{i=1}^n \rho(y, y_i)} \quad (6)$$

$$\sigma^2(x|y) = \frac{\sum_{i=1}^n \rho(y, y_i) [x - \mu(x|y)]^2}{\sum_{i=1}^n \rho(y, y_i)} \quad (7)$$

3.2 Time series of rankings

The time dependent label ranking (TDLR) problem takes the inter-temporal dependence between the rankings into account. That is, rankings that are similar to the most recent ones are more likely to appear. To capture this, we propose the

weighted TDLR prior probability:

$$P_{TDLR}(y_t) = \frac{\sum_{t=1}^n w_t \rho(y, y_t)}{\sum_{t=1}^n w_t} \quad (8)$$

where $w = \{w_1, \dots, w_n\} \rightarrow \mathbf{w}$ is the vector of weights calculated from the exponential function $\mathbf{w} = b^{\frac{1-\{t\}_1^n}{t}}$. Parameter $b \in \{1 \dots \infty\}$ sets the degree of the “memory” for the past rankings, i.e., the larger b , the more weight is given to the last known ranking (i.e., at $t - 1$) and the weight diminishes to the rankings known at $t = 1$.

As for the conditional label ranking probability, the equation for the weighted mean (Equation (5)) becomes:

$$\mu(x_t|y_t) = \frac{\sum_{t=1}^n w_t \rho(y, y_t) x_t}{\sum_{t=1}^n \rho(y, y_t)} \quad (9)$$

$$\sigma^2(x_t|y_t) = \frac{\sum_{i=1}^n w_i \rho(y, y_t) [x_t - \mu(x_t|y)]^2}{\sum_{i=1}^n \rho(y, y_t)} \quad (10)$$

4 State variables

Several studies try to analyze factors that affect the performance of analysts (Clement, 1999; Brown and Mohammad, 2003; Jegadeesh, Kim, Krische, and Lee, 2004). However, most of these papers look at the individual characteristics of analysts such as their job experience, their affiliation, education background, industry specializations, etc. These variables are very important to characterize the relative performance of the analysts in general but they miss the “state-of-the-world” component, i.e., variables that all analysts are affected. We believe that rankings of analysts capture this component in full.

Ranking means that there are differences in opinion among the analysts concerning the future performance of a company. This implies that there is a variability (dispersion) in analysts' forecasts for a given stock in a given quarter (Diether, Malloy, and Scherbina, 2002). Thus, we can analyze the analysts' forecasts dispersion in terms of its origin and factors that affect it; hence, assuming the same variables influence the rankings. It follows that the variation in rankings is due to the different ability of analysts to interpret informational from different states of the environment (e.g., whether the market is bull or bear). We, thus, select and analyze variables that describe these states of environment.

To capture the full spectrum of the analyst's decision making process, we select variables based on different levels of information availability: analyst-specific, firm-specific and general economy. In each level, we want a variable to be responsible for information asymmetry and uncertainty. Thus, we believe that these two domains are responsible for the differences in analysts' opinions.

4.1 Analysts-specific variables

On analysts' level, we want to capture the asymmetry and uncertainty among the analysts (Barron, Kim, Lim, and Stevens, 1998; Barron, Stanford, and Yu, 2009; Zhang, 2006; Sheng and Thevenot, 2012). Particularly, Barron et al. (2009) point out that the reason for dispersion is either uncertainty or information asymmetry. They find that prior to earnings announcement the uncertainty component prevail, whereas around the time of earnings announcement, information asymmetry is responsible for changes in analysts' opinions.

To capture the states of the dispersion, we use the same set of variables defined

in (Barron et al., 2009, page 333):

$$SE = (ACT - \overline{FE})^2$$

$$disp = \sum_{j=1}^k \frac{(FE_j - \overline{FE})^2}{(k-1)} \quad (11)$$

$$uncert = \sum_{j=1}^k \left(1 - \frac{1}{k}\right) \times disp + SE \quad (12)$$

$$assym = 1 - \frac{SE - \frac{disp}{k}}{uncert} \quad (13)$$

where SE is the squared error in mean forecast; \overline{FE} is the average per analysts EPS forecast error (see 5.1); and k is the number of analysts in a given quarter for a given stock.

Equation (11) calculates the dispersion among the analysts which is a variance of EPS forecasts of all analysts for a given stock. Equation (12) defines the Uncertainty component of the dispersion per Barron et al. (2009). Equation (13) is the proxy for information asymmetry which a function of dispersion, squared mean error, and a number of EPS forecasts.

4.2 Firm-based variables

To be consistent with the two paradigms that characterize the state of the analysts, we split the firm-based variables based on their influence on analysts' opinions.

4.2.1 Uncertainty

The following are the set of the variables and their definitions that we think are responsible for the uncertainty component.

Business risk. Business risk is associated with the uncertainty in operating results, especially, in operating earnings (Hill and Stone, 1980). An increase in business risk entails an increase in *ex-ante* volatility of the reported earnings (Parkash, Dhaliwal, and Salatka, 1995). We select book-to-market ratio as a proxy for the business risk measurement.

$$btm = \frac{EQUITY}{MKT.CAP} = \frac{Tot.assts - Tot.liab}{Stocks \times Price} \quad (14)$$

where Stocks is the number of stocks outstanding and Price is the close stock price on last day of a quarter.

Financial risk. Financial risk is responsible for the uncertainty of the future earnings. More debt implies more variability in earnings as managers would try to maximize the value of a stock using the debt; thus, having high risk of default in the future or taking high risk investment projects. The debt-to-equity ratio is used to capture the financial risk (Parkash et al., 1995). We use short-term debt from balance sheet (Notes payable) as a measure for debt.

$$dte = \frac{DEBT}{EQUITY} = \frac{ShortTermDebt}{Tot.assts - Tot.liab} \quad (15)$$

Size. The firm size can be used as a proxy for amount of information available for a firm. Thus, larger firm has more news coverage which reduces uncertainty.

An investor is likely to find private information about larger firm more valuable than the same information about smaller firm (Bhushan, 1989).

Size is measured as the market value (MV) of the firm as follows:

$$size = \log(\text{Price} \times \text{Stocks}) \quad (16)$$

Consistent with the literature, we use log of market value.

Return variability. Return variability influence the uncertainty regarding future earnings (Diether et al., 2002; Henley, Carruth, and Dickerson, 2003). An increase in variability of the abnormal returns is positively correlated with the uncertainty about the earnings; thus, affecting the dispersion among the analysts. To calculate the return variability, we use method provided in Sousa and Serra (2008), where stock return volatility is decomposed into market and stock specific components as follows:

$$\begin{aligned} \sigma_{mkt}^2 &= \sum_{d \in t} (R_{mkt,d} - \mu_{mkt})^2 \\ \sigma^2 &= \sum_{d \in t} (R_d - R_{mkt,d})^2 \\ s.ret = Var(R_t) &= \sigma_{mkt}^2 + \sigma^2 \end{aligned} \quad (17)$$

where $R_{mkt,t}$ is the market return over sample period; μ_{mkt} is the mean of daily market returns; R_t is an individual stock return; d is the number of trading days in period t .

4.2.2 Information Asymmetry variables

Accruals. Accruals, as a part of earnings, is one of the variables that cause the information asymmetry between managers of a firm and investors. Studies have shown that presence of asymmetry is a necessary condition for the earnings management (Trueman and Titman, 1988; Richardson, 2000). To be more specific, it is the discretionary part of the accruals that causes the information inefficiency in the earnings management (Richardson, 2000; Ahmed, Nainar, and Zhou, 2005). We calculated total accruals-to-total assets ratio defined in Creamer and Stolfo (2009):

$$accr = \frac{\Delta C.As - \Delta Cash - (\Delta C.Lb. - \Delta C.Lb.D) - \Delta T - D\&A_t}{(T.As. - T.As._{t-4})/2} \quad (18)$$

where $\Delta X = X_t - X_{t-1}$; C.As – current assets; C.Lb – current liabilities; C.Lb.D – debt in current liabilities; T – deferred taxes; D&A – depreciation and amortization; and T.A – total assets.

Sector-based variables. The industry specific variables that cause the dispersion in the analysts' forecasts are connected with the uncertainty concept. One of the variables that is suggested to capture is the variability in the industry Producer Price Index (PPI) (Henley et al., 2003).

$$sec.ret = \sigma(\log PPI_{sec}) \quad (19)$$

where $\sigma(\log \text{PPI}_{sec})$ is the standard deviation of the log of SIC sectors' produce price index.

4.3 Macroeconomics variables

In the last set of the state variables, we want to capture the macroeconomic conditions which affect the analysts' dispersion. For example, different states of the economy are based on different levels of “GNP–inflation” combinations (Lev and Thiagarajan, 1993; Hope and Kang, 2005). When economy is booming, i.e. “high GNP-low inflation” state, Lev and Thiagarajan (1993) observe the significant increase in firms' Capital Expenditures coefficient. This implies that firms start enjoy capital investment due to the low cost of capital. This state of the economy produces less uncertainty. In the “medium GNP-high inflation” state of the economy, there is an increase in R&D expenditures, which, from the above mentioned analysis, may spur high level of information asymmetry based on the increase R&D activities. Finally, in the “low GNP-high inflation” state, Lev and Thiagarajan (1993) observe the Doubtful Receivables coefficient is the largest implying that at this recession state many firms go bankrupt or default on the loans – a signal of high uncertainty in the economy. All these states produce the dispersion of the analysts' forecasts.

We select the following set of the macroeconomic variables:

- gnp = Gross National Product;
- $infl$ = Inflation rate;
- $t.bill$ = Interest rate (90-days T-bill rate);

- $vix.ret$ = Market variability (CBOE VIX index)

5 Target rankings of financial analysts

Analysts are ranked on the basis of Proportional Mean Absolute Forecast Error (PMAFE) that measures the accuracy of a forecast (Clement, 1999; Brown, 2001; Ertimur, Sunder, and Sunder, 2007). We use both target price and EPS accuracy to build the rankings.

5.1 Target Price ranking

We define the forecast daily error FE_j as the absolute value of the difference between analyst's target price TP_j and the daily stock price P for each stock:

$$FE_j^{TP} = |P - TP_j| \quad (20)$$

The PMAFE is given as:

$$PMAFE_j^{TP} = \frac{FE_j^{TP}}{\overline{FE^{TP}}} \quad (21)$$

where $\overline{FE^{TP}}$ is the average forecasting error across analysts. The target price is fixed over the quarter unless it gets revised.

The rank is average analyst's $PMAFE_j^{TP}$ over a particular quarter:

$$\overline{PMAFE_j^{TP}} = \frac{1}{d} \sum_{i=1}^d PMAFE_{j,i}^{TP} \quad (22)$$

$$rank_j = \text{rank}_{j=1}^k \left\{ \overline{PMAFE_j^{TP}} \right\} \quad (23)$$

5.2 EPS ranking

To compute the EPS rankings, we apply the same procedure as above:

$$FE_j^{EPS} = |ACT - PRED_j| \quad (24)$$

$$PMAFE_j^{EPS} = \frac{FE_j^{EPS}}{\overline{FE^{EPS}}} \quad (25)$$

$$rank_j = \text{rank}_{j=1}^k \left\{ PMAFE_j^{EPS} \right\} \quad (26)$$

where ACT and $PRED_j$ are the actual quarterly EPS and analyst j 's EPS forecast for stock.

6 Analysts' information environment

To proceed with the ranking prediction, we need to establish which information we will be using to initially rank analysts.

6.1 Past information sets

Different analysts' ranks are obtained if we select different time horizons. If we use only the most recent information, we will capture the recent performance of the analysts. This, of course, is more sensitive to unique episodes (e.g., a quarter which has been surprisingly good or bad). If, alternatively, we opt to incorporate the entire analyst performance, the ranking is less affected by such events, yet

it may not reflect the current analyst ability. We use two information sets: the first uses only the information about the analyst's performance in period $t - 1$; the second, uses all the available information for that particular analyst. We call the former the *recent* rankings and the latter the *all-time* rankings.

In addition to these rankings, we also create a hypothetical scenario that assumes we anticipate perfectly the future analyst accuracy performance that would only be available at the end of t . We call this the *true* rankings.

Formalizing information sets considered are:

- the *true* rankings

$$\text{rank}_{j,t} = \text{rank}_{j,t} \quad (27)$$

- the *recent* rankings

$$\text{rank}_{j,t} = \text{rank}_{j,t-1} \quad (28)$$

- the *all-time* rankings

$$\text{rank}_{j,t} = \frac{1}{T} \sum_{t=1}^T \text{rank}_{j,t} \quad (29)$$

where $\text{rank}_{j,t}$ is analyst j rank at time t . These rankings will serve as baselines to assess the quality of the predicted rankings.

6.2 Dynamic states

For ranking predictions, the past information sets are no longer valid as we model the variables that affect analysts' performance with *true* rankings. For this reason, it is necessary to take into account the dynamics of independent variables from one time period to another. We propose the following methods:

- `static`: no dynamics in the state of the variables, i.e., independent variables used as they are: $x_{\Delta t} = x_t$;
- `diff`: first-difference of the variables, i.e., $x_{\Delta t} = x_t - x_{t-1}$;
- `random`: in time series decomposition of the independent variables, it is an unobserved component: $x_{\Delta t} = T(t) + S(t) + \epsilon(t)$, where $T(t)$ - trend, $S(t)$ - seasonal part and $\epsilon(t)$ - random part of time series decomposition.
- `roll.sd`: rolling 8 quarters standard deviation of the independent variables (Zivot and Wang, 2003):

$$\begin{aligned}\mu_t(8) &= \frac{1}{8} \sum_{i=0}^7 x_{t-i} \\ \sigma_t^2(8) &= \frac{1}{7} \sum_{i=0}^7 (x_{t-i} - \mu(8))^2\end{aligned}\tag{30}$$

Each of these methods produces a different set of attributes which corresponds to different predicted rankings. Overall, in each time period we would have seven different rankings: one is the perfect foresight ranking (*true*), two are based on the different sizes of analysts' past information (*recent* and *all-time*), and the rest are from the predicted model with different dynamic states (`static`, `diff`, `random`, and `roll.sd`).

We selected variables that describe the information environment consistent with Aiguzhinov, Serra, and Soares (2015b). We use variable that have more than 10% contribution to the rankings. **Table 2** demonstrates the total discriminative power of state variables for different states for EPS ranking. The variables that contribute the most to the rankings are: $\{uncert; disp; s.ret; gnp; infl; t.bill\}$.

7 Trading Strategy

We follow the Black-Litterman procedure developed in Aiguzhinov et al. (2015a):

1. For each stock, at the beginning of quarter t , we use predicted rankings of all analysts that we expect to be at the end of the quarter t ;
2. Based on these predicted rankings and analysts' price targets, we define Q_t and Ω_t (see (7.1) and (7.2));
3. Using market information available at the last day of quarter $t - 1$, we obtain the market inputs;
4. Apply BL model to get optimized portfolio weights and buy/sell stocks accordingly;

The model requires from an investor two inputs: the vector of expected returns and the confidence of these returns. The vector of returns is where we rely on the knowledge of the analysts. We use two types of expected returns: 1) ones that are based on the consensus among analysts about future stock performance; 2) ones that are based on the rankings of analysts.

7.1 Defining Q

For the consensus strategy, we use median of expected returns for a particular stock:

$$Q_{cons} = \text{median} \{r_j\} \quad (31)$$

where $r_j = TP_j/P - 1$ is last known analyst's j expected return computed using the analyst price target TP_j and stock price P ⁴.

For the strategies that weight the analysts' estimates of expected return the weight of each analyst j is based on his/her rank such that the top analyst has the weight of 1 and then the weights diminish as the rank increases.

$$w_j = 1 - \frac{\text{rank}_j - \min \{\text{rank}\}}{\max \{\text{rank}\}} \quad (32)$$

where rank_j is the predicted analyst j rank (Section 6)

The expected rank-weighted return is thus:

$$Q_{rank} = \frac{\sum_{j=1}^k (w_j \times r_j)}{\sum_{j=1}^k w_j} \quad (33)$$

7.2 Defining the confidence of expected returns Ω

The confidence of Q is given by the coefficient of variation (CV) of forecasting errors:

$$CV = \frac{\sigma(FE)}{\overline{FE}} \quad (34)$$

where σ and \overline{FE} are the standard deviation and the mean of the forecast errors across analysts for TP. A low value of CV reflects consensual estimates of future prices.

⁴Consistent with the literature, we use stock price 3 days *ex-ante* the TP announcement. This is done to avoid any information leakage around new TP announcement day (Bonini, Zanetti, Bianchini, and Salvi, 2010)

8 Data and experimental setup

8.1 Database and sample

We focus our sample on the S&P500 stocks. The period of the experiments runs from the first quarter of 2001 until the last quarter of 2009. We get the analysts price target and EPS forecast data from ThomsonReuters I/B/E/S dataset; the list of S&P constituents and stock daily prices data are from DataStream as well as the market capitalization data.

Over the sample period, the total number of Equity Research Firms (ERF)⁵ in TP dataset is 477, covering 502 stocks. Given the fact that financial analysts commonly issue TP with the one year horizon⁶, we assume that analysts keep their TP forecasts valid for one calendar year unless it is revised. After one year we assume that TP recommendation expires.

Consistent with other studies on analysts' expected returns that work with price targets (Bradshaw, 2002; Brav and Lehavy, 2003; Da and Schaumburg, 2011), we truncate the sample of $TP/P - 1$ at the 5th percentile (values below -0.14) and at the 95th percentile (values above 0.99). This is done due to occurrence of the extreme values. Most of these extreme values are driven by misalignment errors found on I/B/E/S data⁷. To implement ranking, we require that a stock had at least three equity research firms per quarter and that a equity research firm has to be active in covering a particular stock for at least 3 years (12 quarters). After all the data requirements, our final sample number of equity research firms

⁵We use words "analyst" and "Equity Research Firm" interchangeably.

⁶According to Wharton Research Data Services (WRDS), 92.33% of all price targets reported in I/B/E/S have a 12-month horizon (Glushkov, 2009).

⁷We found some differences between the DataStream and I/B/E/S the databases. In some cases the stock-splits and the dividends were not properly adjusted.

issued target prices is 152 covering 419 S&P500 stocks. Overall, the number of observations ($\text{Stock} \times \text{ERF} \times \text{Quarter}$) is reduced from 134336 (initial) to 90743 (filtered).

In the case of EPS forecasts, the initial file of quarterly EPS forecast consists of 437 ERFs covering 516 stocks. Considering the ranking data requirement, our final sample of EPS forecasts consist of 157 ERFs covering 402 S&P500 stocks. The total number of observation is 80185.

Table 4 presents descriptive statistics of the price targets (panel A) and EPS forecasts (panel B). We observe that, on average, the analysts issue 5.52 and 5.6 of price targets and EPS forecasts per quarter respectively.

8.2 Ranking contingency results

We check for analysts' ranking consistency as follows. In one particular quarter (t), we place analysts at one of the bins which corresponds to a tercile: *top*, *medium*, *bottom*. We, then, check analysts position at the immediate next quarter ($t + 1$) and after one year ($t + 4$).

Beforehand, we convert the rankings into scores as follows:

$$\text{score}_j = \frac{\text{rank}_j}{\max \text{rank}} \quad (35)$$

To get the cross-sectional values of scores across different stocks, we take the average of score_j

$$\overline{\text{score}}_j = \frac{1}{M} \sum_{i=1}^M \text{score}_{j,i} \quad (36)$$

where M is number of stocks followed by a particular analyst j .

Table 1 shows a contingency analysis of the ranks. Panel A shows the dynamics of each tercile for rankings based on target price accuracy for the *recent* and the *all-time* rankings. We observe that analysts exhibit strong ranking consistency as, on average, they stay at the same tercile after one quarter. For the *recent* case, of the *top* (*bottom*) most accurate (inaccurate) analysts in the previous quarter 67.79% (69.69%) remain in that same tercile after one quarter. After one year the corresponding figures are lower respectively 46.04% and 41.08% for the *top* and *bottom* terciles. In case of the *all-time*, the analyst consistency is even more profound with 92.35% (92.13%) of analysts that stayed on *top* (*bottom*) in previous quarter remained in the same tercile in the next quarter. Even after one year, the consistency does not change much with 81.76% of analysts stayed on *top* and 79.6% remained at the *bottom*.

In the case of EPS (panel B), for the *recent* ranking 67.87% and 54.86% (51.88% and 35.29%) of the analysts remained in the *top* and *bottom* terciles, respectively, after one quarter (year). For the case of *all-time* ranking, 67.19% and 53.03% (53.1% and 33.71%) of the analysts stayed on *top* and *bottom* respectively after one quarter (year).

These results are consistent with the recent findings of Hilary and Hsu (2013) on analyst forecast consistency.

8.3 Views: descriptive statistics

Table 5 presents the descriptive statistics of the analysts' expected returns. The expected returns are computed comparing TP estimates with actual prices. To form the smart strategies we compute rank-weighted estimates where weights are

given by the TP rankings.

Bradshaw (2002) reports analyst average expected returns for the period of 2000–2009 and 206 ERFs of 24%. Da and Schaumburg (2011) report an average expected return of 40% for the period of 1996–2004. Zhou (2013) finds an average expected return of 96% for the sample period of 2000–2009. These figures suggest that analysts are overly optimistic.

Panel A of **Table 5** show the statistics for the consensus expectations as defined in **Equation (31)**. As mentioned above in **7.1**, the consensus views have equal weights among the analysts, regardless of their ranks; thus, for the cases of *true*, *recent*, and *all-time*, the median is the same regardless of knowing or not the present or past rankings (Q_{cons} in **Equation (31)**). As such, the mean, median, and standard deviation are the same and independent of analysts' information environment. However, since views also include the confidence (**Equation (34)**), which is based on analysts past performance, the results of the trading strategy based on consensus expectations will be different for the *recent* and the *all-time* information sets.

Panel B of the table shows the TP accuracy weighted average expected returns. For each information environment (*true*, *recent*, *all-time*, *static*, *diff*, *random*, and *roll.sd*) the average expected return is respectively 0.15%, 0.16%, 0.12%, 0.15%, 0.15%, 0.16%, and 0.15%.

9 Empirical Results

9.1 Ranking predictions

We report the ranking prediction results in [Table 6](#). Panel A (B) presents annual and total ranking accuracy measured as the average Spearman rank correlation between the *true* and all other label ranking methods obtained from the analysts' target price (EPS) forecasts. Specifically, the table shows if the label ranking model can better predict rankings than a “no-model” setup. In panel A, we observe that predicted rankings based on target prices are constantly outperformed the *all-time* baseline but fail the *recent*. Thus, our model can predict more accurately rankings of those analysts' whose relative performance in setting target prices is based on the whole history.

The average accuracy of the predicted rankings based on the EPS forecasts (panel B) demonstrates quite different results. We report that rankings predicted with the `diff` state of the variables are more accurate than those of the baselines. Thus, our label ranking model can predict rankings of analysts' who issue EPS forecast with higher accuracy than those obtained not just from the *all-time* analysts' past performance information (as in case of the target prices) but also from the *recent* set.

[Figure 1](#) depicts the reported average Spearman correlations. We apply the local polynomial regression fitting (Cleveland, Grosse, and Shyu, 1992) to smooth the series for a better presentation. The top panel of [Figure 1](#) plots the average accuracy of rankings based on the price targets. We observe the accuracy of rankings based on the *recent* information are constantly above all others; only by the end of the sample period it begins to drop. The *all-time* case, on the contrary, drops

rapidly at the beginning of the sample period and continues to decrease in value until the end of the period. The accuracy of predicted rankings falls in between of the *recent* and the *all-time* cases. The cases of `diff` and `random` show a constant increase in accuracy starting from 2005-2006 and, by the end of the sample period, the `random` almost reaches the level of the accuracy of *recent* case while the `diff` is still in the upward trend.

The bottom panel of **Figure 1** plots the case of the EPS based rankings. Looking at baselines rankings, i.e., rankings that obtained without label ranking model, we observe that for the first half of the sample period (from 2000 until 2004) the *recent* rankings demonstrate a constant increase in the average ranking accuracy. This means that during this period, rankings of the analysts, on average, did not change from one quarter to another. Starting from 2004, the average accuracy of the *recent* rankings begins to decrease until the end of the sample period. A somewhat interesting pattern shows the average accuracy of the *all-time* ranking: similar to the *recent*, it also has been increasing until 2005 but running below the *recent*. However, after 2005, the *all-time* ranking accuracy outran the *recent*. Moreover, in times of the financial crisis of 2007-2009, the *all-time* demonstrated an increasing trend in ranking accuracy. As far as the predicted rankings are concerned, we observe that the `diff` average accuracy has been constantly above the other methods and only during the period of crisis it dropped significantly. In fact, except for the `random`, all predicted methods reach the maximum average accuracy at around 2007, a pre-crisis time, and all, except for the *all-time*, show a downward trend afterwards.

Overall, the results of experiment of the predicting the rankings show that it is possible to model the independent variables under different dynamic states

with the rankings. Moreover, the label ranking model can predict the rankings that outperform, in terms of average accuracy, the ones obtained from the past analysts’ relative performance. Concretely, we report that the predicted rankings based on the EPS forecasts outperform the baselines under the `diff` dynamic state of independent variables.

9.2 Trading strategies

We perform a back-test of trading strategy consistent with Aiguzhinov et al. (2015a); namely, we build the “smart estimates” from rankings based on the price targets. The results are presented in [Table 7](#). Panel A reports the performance of *Market* (passive strategy). This strategy showed annualized cumulative return of -3.03% and annualized Sharpe ratio of -0.18 . The average number of stocks used per quarter is 499.98 and the turnover ratio of strategy is 0.05 which demonstrates the ins/outs of the S&P 500 constituents list.

Panel B of the table demonstrates the results of trading with consensus among analysts about price targets. The annualized cumulative return of this strategy under the *recent* (*all-time*) information set is 0.12% (0.31%) and the Sharpe ratio is 0.01 (0.02). This strategy outperforms the *Market* in both of the information sets with the *all-time* outrunning the *recent* in terms of the annualized cumulative returns.

Panel C of the table demonstrates the results of trading based on analysts’ rankings. We observe that consistent with our assumption, the *true* resulted in the maximum possible annual cumulative return and the Sharpe ratio (4.32% and 0.29 respectively). Given the hypothetical assumption of the *true* set, it is not feasible

to implement. The next best feasible strategy is `diff` which is based on the predicted rankings from our label ranking model. This strategy yields an annualized cumulative return of 0.83% and the Sharpe ratio of 0.05 which are higher than those of the *CONS* and *Market* strategies. Moreover, we report that all strategies based on the predicted rankings with the dynamic states yield higher annualized cumulative returns than those based on the *recent* and the *all-time* rankings. The analysis of the sub-periods performance of the `diff` strategy is depicted in **Table 8**. The table shows the value of Sharpe ratio for the 5-year periods. We observe that the `diff` strategy was dominant in most of the periods.

To test the significance of the annualized cumulative returns of strategies based on the predicted rankings we perform a null-hypothesis pairwise test when *null* is the difference in returns is zero. **Table 9** presents the test results. We report that the returns of the strategies based on the predicted rankings with the `diff` and the `random` dynamic states are statistically significant at 1% when compared with the returns of all other strategies. The test accepts the *null* for the returns obtained for the strategies based on the predicted rankings of the `static` and the `roll.sd` states.

Figure 2 plots the graphical representation of the cumulative returns for all trading strategies. We observe that for the case of strategy based on the analysts' rankings, the *true* strategy is always on top of all others. This implies that in the settings where analysts' expected returns and rankings are based on price targets, an investor can gain a maximum results from trading strategy.

10 Conclusion

Some institutions, such as StarMine, rank financial analysts based on their accuracy and investment value performance. These rankings are published and are relevant: stocks favored by top-ranked analysts will probably receive more attention from investors. Therefore, there is a growing interest in understanding the relative performance of analysts.

In this paper we show that it is possible to model analysts' rankings and variables that affect them. With recent findings from Machine Learning body of research in label ranking, we adapted the algorithm to predict the rankings of financial analysts based on price targets and EPS forecasts. We report that, in case of price targets, our predicted rankings are more accurate than those obtained from using information of the whole history of analysts performance. For the rankings based on EPS forecasts, our model is able to predict rankings that are better than those of the *all-time* and the *recent* baselines. Moreover, the supremacy of our model above the baselines occurs when the variables that characterize analysts' information environment exhibit a stationary behavior expressed as the first-difference.

We also performed a back-test of active trading using the predicted rankings as inputs for the Black-Litterman model. The results showed that the strategies based on the analysts' rankings outperform, in terms of the annualized cumulative return, a strategy based on the analysts' consensus. Of the ranking based trading strategies, the maximum annualized cumulative return yields a strategy that is based on the predicted rankings with the first-difference of state variables.

The results of our work open many opportunities for future research. For ex-

ample, in this paper we use the classical interpretation of the Black-Litterman model where risk is measured as a standard deviation. Recent work suggests utilizing more complex measures such as value-at-risk and high-moments approaches.

References

- Ahmed, A., S. Nainar, and J. Zhou (2005). Do analysts' earnings forecasts fully reflect the information in accruals? *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration* 22(4), 329–342.
- Aiguzhinov, A., A. Serra, and C. Soares (2015a). Are rankings of financial analysts useful for investors? <http://ssrn.com/abstract=2680094>. Last checked on 26/10/2015.
- Aiguzhinov, A., A. Serra, and C. Soares (2015b). Understanding rankings of financial analysts. <http://ssrn.com/abstract=2680301>. Last checked on 26/10/2015.
- Aiguzhinov, A., C. Soares, and A. Serra (2010). A similarity-based adaptation of naive Bayes for label ranking: Application to the metalearning problem of algorithm recommendation. In B. Pfahringer, G. Holmes, and A. Hoffmann (Eds.), *Discovery Science*, Volume 6332 of *Lecture Notes in Computer Science*, pp. 16–26. Springer Berlin Heidelberg.
- Barber, B., R. Lehavy, M. McNichols, and B. Trueman (2001). Can investors profit from the prophets? Security analyst recommendations and stock returns. *The Journal of Finance* 56(2), 531–563.
- Barron, O., O. Kim, S. Lim, and D. Stevens (1998). Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73(4), 421–433.

- Barron, O., M. Stanford, and Y. Yu (2009). Further evidence on the relation between analysts' forecast dispersion and stock returns. *Contemporary Accounting Research* 26(2), 329–357.
- Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics* 11(2-3), 255–274.
- Black, F. and R. Litterman (1992). Global portfolio optimization. *Financial Analysts Journal* 48(5), 28–43.
- Bonini, S., L. Zanetti, R. Bianchini, and A. Salvi (2010). Target price accuracy in equity research. *Journal of Business Finance & Accounting* 37(9-10), 1177–1217.
- Bouckaert, R. R. (2005). Naive Bayes classifiers that perform well with continuous variables. In *AI 2004: Advances in Artificial Intelligence*, Volume 3339 of *Lecture Notes in Computer Science*, pp. 85–116.
- Bradshaw, M. T. (2002). The use of target prices to justify sell-side analysts' stock recommendations. *Accounting Horizons* 16(1), 27–41.
- Brav, A. and R. Lehavy (2003). An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. *The Journal of Finance* 58(5), 1933–1968.
- Brown, L. (2001). How important is past analyst earnings forecast accuracy? *Financial Analysts Journal* 57(6), 44–49.
- Brown, L. and E. Mohammad (2003). The predictive value of analyst characteristics. *Journal of Accounting, Auditing and Finance* 18(4), 625–647.
- Clement, M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27(3), 285–303.
- Cleveland, W. S., E. Grosse, and W. M. Shyu (1992). Local regression models. *Statistical models in S*, 309–376.

- Creamer, G. and S. Stolfo (2009). A link mining algorithm for earnings forecast and trading. *Data Mining and Knowledge Discovery* 18(3), 419–445.
- Da, Z. and E. Schaumburg (2011). Relative valuation and analyst target price forecasts. *Journal of Financial Markets* 14(1), 161–192.
- Diether, K., C. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance* 57(5), 2113–2141.
- Emery, D. and X. Li (2009). Are the Wall Street analyst rankings popularity contests? *Journal of Financial and Quantitative Analysis* 44(2), 411.
- Ertimur, Y., J. Sunder, and S. Sunder (2007). Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts. *Journal of Accounting Research* 45(3), 567–606.
- Fama, E. (1970). Efficient capital markets: A review of empirical work. *The Journal of Finance* 25, 383–417.
- Glushkov, D. (2009). Overview of IBES on WRDS: Research and data issues. <http://goo.gl/M0T1tw>. Last checked on: 27/01/2014.
- Grossman, S. and J. Stiglitz (1980). On the impossibility of informationally efficient prices. *The American Economic Review* 70(3), 393–408.
- Henley, A., A. Carruth, and A. Dickerson (2003). Industry-wide versus firm-specific uncertainty and investment: British company panel data evidence. *Economics Letters* 78(1), 87–92.
- Hilary, G. and C. Hsu (2013). Analyst forecast consistency. *The Journal of Finance* 68(1), 271–297.
- Hill, N. and B. Stone (1980). Accounting betas, systematic operating risk, and financial leverage: A risk-composition approach to the determinants of systematic risk. *Journal of Financial and Quantitative Analysis* 15(03), 595–637.
- Hope, O. and T. Kang (2005). The association between macroeconomic uncertainty and analysts' forecast accuracy. *Journal of International Accounting Research* 4(1), 23–38.

- Jegadeesh, N., J. Kim, S. Krische, and C. Lee (2004). Analyzing the analysts: When do recommendations add value? *The Journal of Finance* 59(3), 1083–1124.
- Lev, B. and S. Thiagarajan (1993). Fundamental information analysis. *Journal of Accounting research* 31(2), 190–215.
- Parkash, M., D. Dhaliwal, and W. Salatka (1995). How certain firm-specific characteristics affect the accuracy and dispersion of analysts' forecasts:: A latent variables approach. *Journal of Business Research* 34(3), 161–169.
- Richardson, V. J. (2000). Information asymmetry and earnings management: Some evidence. *Review of Quantitative Finance and Accounting* 15(4), 325–347.
- Sheng, X. and M. Thevenot (2012). A new measure of earnings forecast uncertainty. *Journal of Accounting and Economics* 53(1), 21–33.
- Sousa, S. and A. Serra (2008). What drives idiosyncratic volatility over time? *Portuguese Economic Journal* 7(3), 155–181.
- Trueman, B. and S. Titman (1988). An explanation for accounting income smoothing. *Journal of Accounting Research* 26, 127–139.
- Vembu, S. and T. Gärtner (2010). Label ranking algorithms: A survey. In J. Fürnkranz and E. Hüllermeier (Eds.), *Preference Learning*, pp. 45–64. Springer.
- Vogt, M., J. Godden, and J. Bajorath (2007). Bayesian interpretation of a distance function for navigating high-dimensional descriptor spaces. *Journal of Chemical Information and Modeling* 47(1), 39–46.
- Womack, K. (1996, March). Do brokerage analysts' recommendations have investment value? *The Journal of Finance* 51(1), 137–168.
- Zhang, X. (2006). Information uncertainty and stock returns. *The Journal of Finance* 61(1), 105–137.

- Zhou, J. (2013). Analysts' target prices and stock recommendations. *Accounting and Finance Research* 2(1), p1.
- Zivot, E. and J. Wang (2003). *Modeling financial time series with S-PLUS*, Volume 191. Springer Berlin Heidelberg.

Table 1

Analysts' ranking consistency

This contingency table shows changes in analysts' *top*, *middle*, *bottom* ranking bins. Panel A (Panel B) depicts the dynamics of the analysts' ranks based on the accuracy in target prices (EPS forecasts). Rankings of the *recent* is the case of ranking information know at $t - 1$ and the *all-time* is the case of using all ranking information for up to $t - 1$.

	top_t			$middle_t$			$bottom_t$		
	<i>top</i>	<i>mid</i>	<i>bottom</i>	<i>top</i>	<i>mid</i>	<i>bottom</i>	<i>top</i>	<i>mid</i>	<i>bottom</i>
Panel A: TP									
					$t + 1$				
<i>recent</i>	67.8	22.1	10.1	29.9	48.3	21.8	13.3	17.0	69.7
<i>all-time</i>	92.3	7.2	0.5	9.0	83.3	7.8	0.5	7.4	92.1
					$t + 4$				
<i>recent</i>	46.0	28.1	25.9	39.2	29.4	31.4	32.3	26.7	41.1
<i>all-time</i>	81.8	15.2	3.0	19.2	64.3	16.5	2.9	17.5	79.6
Panel B: EPS									
					$t + 1$				
<i>recent</i>	48.3	26.2	26.1	48.1	26.3	25.9	45.8	26.3	28.7
<i>all-time</i>	89.7	9.0	1.3	11.5	78.0	10.5	1.2	10.1	88.8
					$t + 4$				
<i>recent</i>	47.0	28.2	26.1	45.0	27.2	28.5	43.5	27.9	31.4
<i>all-time</i>	79.3	16.6	4.1	20.9	60.2	18.9	4.2	18.9	76.9

Table 2

Discriminative power contribution

The table demonstrates the contributions (in %) of each of the variables to changes in analysts' rankings. State `static` is the state with no dynamics in values of the variables, `diff` is the state with first-difference in values, `random` is the state that captures the random part of values time-series decomposition, and `roll.sd` is the state of values sliding 8 quarters standard deviation.

Variable	<code>static</code>	<code>diff</code>	<code>random</code>	<code>roll.sd</code>
<code>uncert</code>	12.84	17.53	9.69	0.03
<code>assym</code>	3.88	4.79	3.17	0.01
<code>disp</code>	10.52	12.27	5.83	0.09
<code>btm</code>	2.34	2.47	5.72	4.29
<code>size</code>	0.39	0.32	1.30	1.54
<code>dte</code>	4.46	2.72	4.81	3.47
<code>accr</code>	0.90	0.77	1.21	1.42
<code>s.ret</code>	7.48	8.11	12.83	19.38
<code>sec.ret</code>	0.13	1.49	3.65	2.21
<code>gnp</code>	12.81	15.56	18.44	27.31
<code>infl</code>	18.16	24.90	13.43	20.83
<code>vix.ret</code>	0.08	0.22	1.43	0.19
<code>t.bill</code>	26.01	8.86	18.47	19.22
Total	100.00	100.00	100.00	100.00

Table 3
Example of label ranking problem

Period	\mathcal{V}_1	\mathcal{V}_2	\mathcal{V}_3	\mathcal{V}_4	Ranks		
					Alex	Brown	Credit
1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	1	2	3
2	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	2	3	1
3	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	1	2	3
4	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	3	2	1
5	$x_{5,1}$	$x_{5,2}$	$x_{5,3}$	$x_{5,4}$	3	2	1
6	$x_{6,1}$	$x_{6,2}$	$x_{6,3}$	$x_{6,4}$	2	1	3
7	$x_{7,1}$	$x_{7,2}$	$x_{7,3}$	$x_{7,4}$	1	2	3

Table 4

Sample Statistics

This table shows the average number of target prices (panel A) and EPS forecasts (panel B) per stock per quarter.

	Min	Mean	Median	Max	Std.dev
Panel A: TP					
1999	3	4.144	4	10	1.549
2000	3	4.512	4	14	1.806
2001	3	4.873	4	16	2.188
2002	3	5.436	5	19	2.733
2003	3	5.763	5	21	3.073
2004	3	5.931	5	21	3.177
2005	3	6.042	5	21	3.207
2006	3	5.991	5	20	3.094
2007	3	5.755	5	20	2.953
2008	3	5.325	5	18	2.602
2009	3	4.535	4	18	1.958
Total	3	5.522	5	21	2.858
Panel B: EPS					
1999	3	5.165	4	17	2.758
2000	3	5.017	4	17	2.561
2001	3	5.418	5	18	2.758
2002	3	5.532	5	21	2.965
2003	3	5.643	5	24	3.034
2004	3	5.835	5	22	3.293
2005	3	5.933	5	24	3.328
2006	3	6.042	5	26	3.383
2007	3	5.841	5	22	3.103
2008	3	5.318	4	22	2.653
2009	3	4.998	4	20	2.425
Total	3	5.601	5	26	3.024

Table 5

Descriptive statistics of views

This table shows the descriptive statistics of views (expected returns) based on the consensus (median) among the analysts (panel A) and target price rankings (panel B). State *static* is the state with no dynamics in values of the variables, *diff* is the state with first-difference in values, *random* is the state that captures the random part of values time-series decomposition, and *roll.sd* is the state of values sliding 8 quarters standard deviation.

	Mean (in %)	Median (in %)	Std.dev
Panel A: Consensus			
<i>true</i>	18.610	16.889	0.120
<i>recent</i>	18.610	16.889	0.120
<i>all-time</i>	18.610	16.889	0.120
<i>static</i>	18.610	16.889	0.120
<i>diff</i>	18.610	16.889	0.120
<i>random</i>	18.610	16.889	0.120
<i>roll.sd</i>	18.610	16.889	0.120
Panel B: TP			
<i>true</i>	14.876	13.380	0.096
<i>recent</i>	15.742	14.314	0.098
<i>all-time</i>	12.459	10.591	0.089
<i>static</i>	15.346	13.859	0.095
<i>diff</i>	15.367	13.797	0.097
<i>random</i>	15.735	14.245	0.096
<i>roll.sd</i>	15.174	13.658	0.095

Table 6

Average ranking accuracy

The table presents the average Spearman correlation between *true* and predicted rankings that are based on accuracy of price target (panel A) and on EPS forecasts (panel B) compared to baselines: *true* shows the case of the known future information; *recent* is the case of ranking information know at $t - 1$, and the *all-time* is the case of using all ranking information for up to $t - 1$. State *static* is the state with no dynamics in values of the variables, *diff* is the state with first-difference in values, *random* is the state that captures the random part of values time-series decomposition, and *roll.sd* is the state of values sliding 8 quarters standard deviation.

Year	<i>true</i>	<i>recent</i>	<i>all-time</i>	<i>static</i>	<i>diff</i>	<i>random</i>	<i>roll.sd</i>
Panel A: TP							
2000	1.0000	0.5988	0.4918	0.5695	0.5706	0.5312	0.5728
2001	1.0000	0.6600	0.4515	0.5355	0.6166	0.6058	0.6030
2002	1.0000	0.6092	0.3770	0.3960	0.4436	0.5310	0.4374
2003	1.0000	0.6168	0.3379	0.4812	0.4771	0.5372	0.5515
2004	1.0000	0.5500	0.2542	0.4031	0.2559	0.4610	0.2517
2005	1.0000	0.5862	0.2965	0.5114	0.4622	0.4100	0.5466
2006	1.0000	0.5908	0.3087	0.5323	0.4583	0.5804	0.4279
2007	1.0000	0.5787	0.2440	0.4823	0.2443	0.4703	0.3483
2008	1.0000	0.6573	0.3012	0.4720	0.5380	0.5247	0.6346
2009	1.0000	0.5720	0.2093	0.3902	0.4434	0.5507	0.3231
Total	1.0000	0.6007	0.3175	0.4746	0.4415	0.5167	0.4642
Panel B: EPS							
2000	1.0000	0.0344	0.0381	0.0420	0.0501	0.0461	0.0486
2001	1.0000	-0.0028	-0.0147	-0.0178	-0.0026	-0.0271	-0.0081
2002	1.0000	0.0386	0.0132	-0.0001	-0.0049	0.0089	0.0113
2003	1.0000	0.0489	0.0279	0.0380	0.0632	0.0263	0.0431
2004	1.0000	0.0255	0.0397	0.0021	0.0092	0.0184	-0.0414
2005	1.0000	0.0465	0.0369	0.0450	0.0297	0.0246	0.0380
2006	1.0000	0.0187	0.0416	0.0162	0.0367	0.0062	0.0332
2007	1.0000	0.0358	0.0246	0.0468	0.0535	0.0392	0.0469
2008	1.0000	0.0066	0.0222	0.0318	0.0110	0.0194	0.0110
2009	1.0000	0.0106	0.0240	0.0261	0.0198	0.0087	0.0228
Total	1.0000	0.0261	0.0264	0.0243	0.0268	0.0173	0.0205

Table 7

Trading strategy performance

The table presents the annualized cumulative statistics of the strategy performance based on PT rankings. *true* is actual ranking of the analysts. *recent* is the rankings from the last period. *all-time* is the average rank of an analyst for up to the last period. Trading period is from 2000Q1 until 2009Q4. Panel A presents the results from the passive strategy. Panel B summarizes the results of the strategy with rankings based on consensus in price targets. Panel C shows the case of the strategy with rankings based on price targets. State *static* is the state with no dynamics in values of the variables; *diff* is the state with first-difference in values; *random* is the state that captures the random part of values time-series decomposition; and *roll.sd* is the state of values sliding 8 quarters standard deviation.

Strategy	Annualized cum. return (in %)	Annualized Std. dev (in %)	Sharpe ratio	Average num. stock	Average turnover rate
Panel A					
<i>Market</i>	-3.032	16.654	-0.182	499	0.053
Panel B: CONS					
<i>recent</i>	0.116	15.948	0.007	283	0.256
<i>all-time</i>	0.314	15.773	0.020	283	0.228
Panel C: TP					
<i>true</i>	4.325	14.697	0.294	283	0.345
<i>recent</i>	0.282	15.662	0.018	284	0.264
<i>all-time</i>	0.689	15.565	0.044	284	0.256
<i>static</i>	0.547	15.759	0.035	251	0.266
<i>diff</i>	0.830	15.742	0.053	251	0.276
<i>random</i>	0.690	15.715	0.044	251	0.262
<i>roll.sd</i>	0.738	15.726	0.047	251	0.270

Table 8

Trading strategy performance: Sharpe ratio

This table presents the Sharpe ratio of each of the trading strategies: the passive (*Market*) and the active (consensus and smart estimates) calculated for different holding periods. Panel A represents the perfect foresight information set; panels B and C show, respectively, the recent and the all history analysts' performance. State *static* is the state with no dynamics in values of the variables, *diff* is the state with first-difference in values, *random* is the state that captures the random part of values time-series decomposition, and *roll.sd* is the state of values sliding 8 quarters standard deviation.

Periods	<i>Market</i>	<i>true</i>	<i>recent</i>	<i>all-time</i>	<i>static</i>	<i>diff</i>	<i>random</i>	<i>roll.sd</i>
2000Q1/2004Q4	-0.201	0.395	0.168	0.205	0.218	0.224	0.223	0.208
2001Q1/2005Q4	-0.093	0.425	0.214	0.249	0.232	0.242	0.239	0.222
2002Q1/2006Q4	0.196	0.757	0.490	0.464	0.525	0.538	0.529	0.511
2003Q1/2007Q4	0.925	1.915	1.248	1.245	1.248	1.305	1.250	1.303
2004Q1/2008Q4	-0.435	0.070	-0.305	-0.289	-0.308	-0.292	-0.308	-0.283
2005Q1/2009Q4	-0.158	0.189	-0.125	-0.105	-0.136	-0.106	-0.124	-0.106
All period	-0.182	0.294	0.018	0.044	0.035	0.053	0.044	0.047

Table 9

Significance of cumulative returns

The table demonstrates a pairwise test in difference of the cumulative returns of all strategies vs. those based on the predicted rankings. Case of *true* shows the known future information; *recent* is the case of ranking information know at $t - 1$, and the *all-time* is the case of using all ranking information for up to $t - 1$. State *static* is the state with no dynamics in values of the variables, *diff* is the state with first-difference in values, *random* is the state that captures the random part of values time-series decomposition, and *roll.sd* is the state of values sliding 8 quarters standard deviation.

	static		diff		random		roll.sd	
	t value	Pr(> t)	t value	Pr(> t)	t value	Pr(> t)	t value	Pr(> t)
<i>true</i>	8.329	0.000	8.383	0.000	8.237	0.000	8.698	0.000
<i>recent</i>	-28.526	0.000	-23.989	0.000	-29.860	0.000	-24.316	0.000
<i>all-time</i>	-5.521	0.000	-13.647	0.000	-8.831	0.000	-8.929	0.000
<i>static</i>	-	-	-6.934	0.000	-8.509	0.000	-0.910	0.368
<i>diff</i>	6.934	0.000	-	-	4.521	0.000	10.374	0.000
<i>random</i>	8.509	0.000	-4.521	0.000	-	-	2.731	0.009
<i>roll.sd</i>	0.910	0.368	-10.374	0.000	-2.731	0.009	-	-
<i>Market</i>	-15.357	0.000	-15.168	0.000	-15.361	0.000	-15.166	0.000

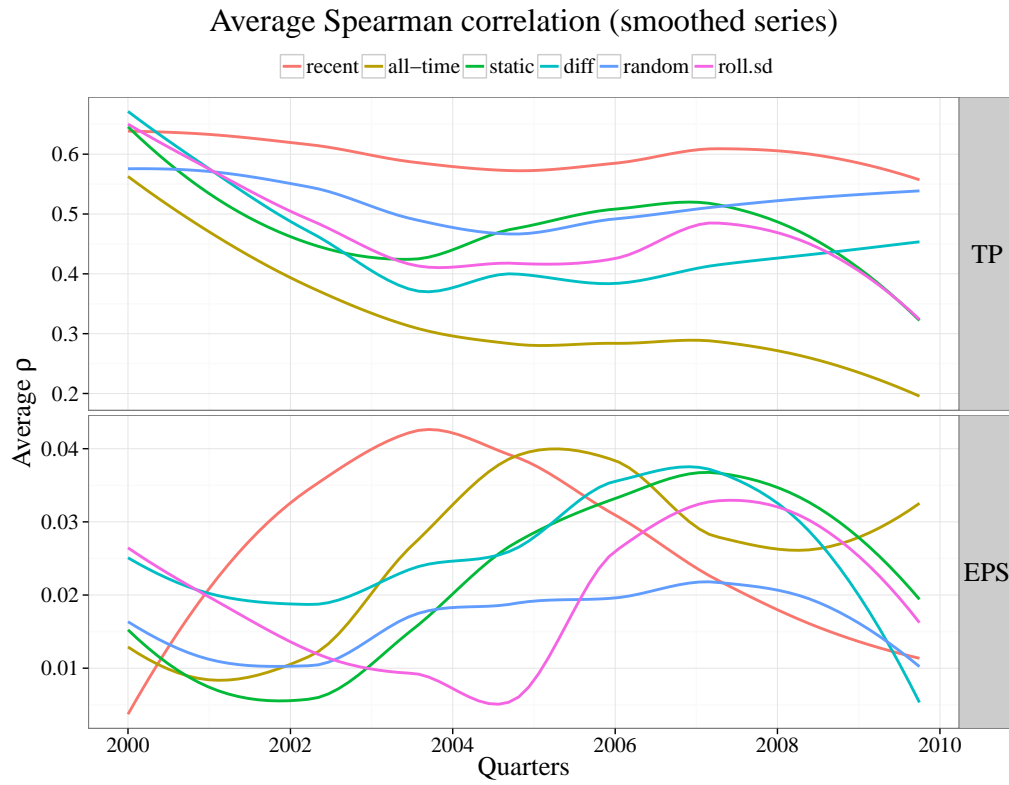
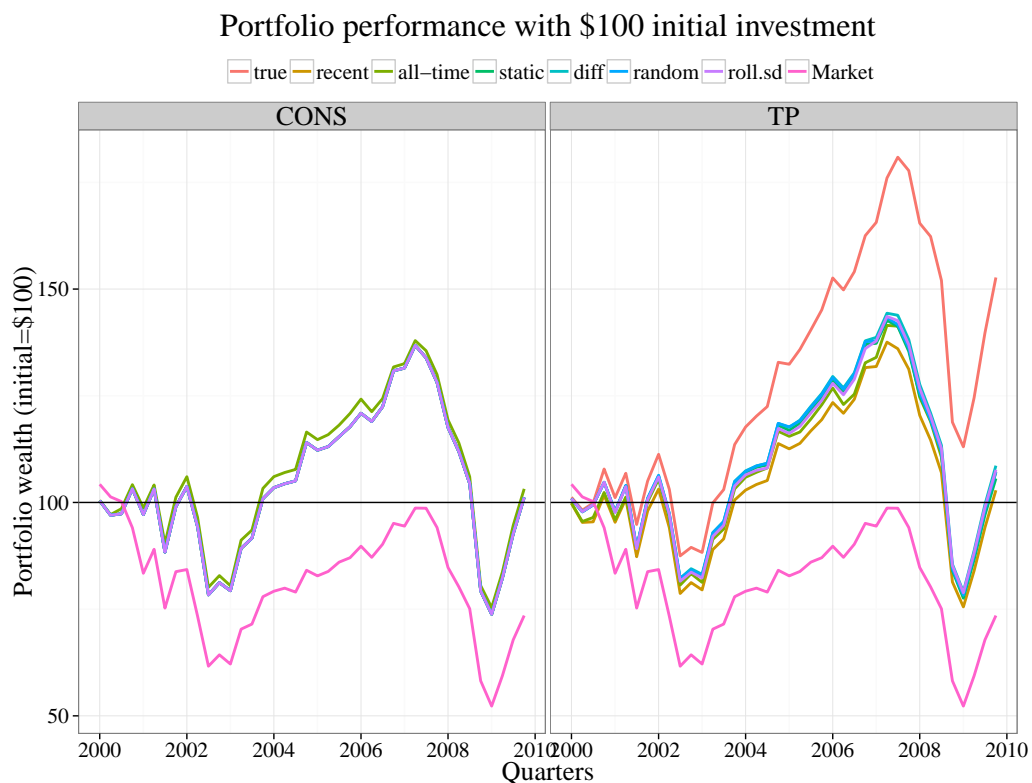


Figure 1

Average Spearman correlation (ρ)

The figure plots the smoothed series of the average Spearman correlation between *true* and predicted rankings resulted from static and dynamic states of variables: *static* is the state with no dynamics in values of the variables, *diff* is the state with first-difference in values, *random* is the state that captures the random part of values time-series decomposition, and *roll.sd* is the state of values sliding 8 quarters standard deviation. Smoothing is done by applying a local polynomial regression fitting (loess smoothing) (Cleveland et al., 1992). The top panel shows results of rankings based on the EPS forecast accuracy, the bottom are based on the price target accuracy.



Performance of BL model

In this figure we show the quarterly performance of the cumulative portfolio wealth for all strategies. State `static` is the state with no dynamics in values of the variables, `diff` is the state with first-difference in values, `random` is the state that captures the random part of values time-series decomposition, and `roll.sd` is the state of values sliding 8 quarters standard deviation.