VIOLA MONOSTORINÉ GROLMUSZ

RECOVERING STOCK ANALYSTS' LOSS FUNCTIONS FROM BUY/SELL RECOMMENDATIONS

MNB WORKING PAPERS | 4

2023
DECEMBER

RECOVERING STOCK ANALYSTS' LOSS FUNCTIONS FROM BUY/SELL RECOMMENDATIONS

MNB WORKING PAPERS | 4

2023
DECEMBER

MNB Working Papers 2023/4

Recovering Stock Analysts' Loss Functions from Buy/Sell Recommendations *

(Részvényelemzők veszteségfüggvényeinek visszanyerése vétel/eladás ajánlásaikból)

Written by Viola Monostoriné Grolmusz**

Budapest, December 2023

Published by the Magyar Nemzeti Bank

Publisher in charge: Eszter Hergár

Krisztina körút 55., H-1013 Budapest

www.mnb.hu

ISSN 1585-5600 (online)

*I am heavily indebted to Róbert Lieli for his guiding work as my Ph.D. advisor. In addition, I am very grateful for Sergey Lychagin, Zsolt Darvas, Lajos Szabó, Péter Gábriel and Anna Naszódi for their valuable comments and suggestions.

**Magyar Nemzeti Bank (Central Bank of Hungary), e-mail: monostorinev@mnb.hu.

Contents

Abstract

I carry out an empirical analysis to recover stock analysts' loss functions from observations on forecasts, actual realizations and a proxy for the publicly observed part of the analyst's information set. The forecasts I use are analyst stock (buy/hold/sell) recommendations for two Blue Chip stocks. I estimate an asymmetry parameter that captures the analyst's relative cost from overpredicting versus underpredicting the stock performance. I find that the results are sensitive to the categorization of 'hold' recommendations. When substituting 'holds' with the recommendation from the previous period, in most cases the estimated bounds for the asymmetry parameter suggest that analysts are more likely to issue a 'false buy' than a 'false sell' recommendation. This is in line with the frequent statement from the analyst recommendations literature, that optimism relative to the consensus is rewarded in analyst recommendations. By shedding light on the direction of bias in individual analysts' stock recommendations, we can better understand the operation of financial markets and we can build more accurate models by controlling for these biases.

JEL: C53, G17.

Keywords: Loss functions, Binary forecasting, Preference recovery.

Összefoglaló

Empirikus elemzésemben részvényelemzők veszteségfüggvényeit határozom meg egyedi előrejelzéseik, a célváltozó megvalósult értéke és az elemző információs halmazának köztudott részhalmaza ismeretében. Az általam használt előrejelzések két Blue Chip részvényre vonatkozó elemzői részvényajánlások (vétel/tartás/eladás). Megbecsülök egy aszimmetriaparamétert, amely az elemzőnek a részvény teljesítményének túl-, illetve alulbecsléséből származó relatív költségeit ragadja meg. Eredményeim érzékenyek a "tartás"-ajánlások kategorizálására. Ha a "tartás" helyett az előző időszak ajánlását használjuk, legtöbbször az aszimmetriaparaméter becsült határai arra utalnak, hogy az elemzők nagyobb valószínűséggel adnak ki 'hamis vételi' ajánlást, mint 'hamis eladási' ajánlást. Ez összhangban van az elemzői ajánlások szakirodalmában gyakran szereplő állítással, miszerint a konszenzushoz viszonyított optimizmust jutalmazzák az elemzői ajánlásokban. Az elemzői részvényajánlások mögötti torzítások irányának felderítése hozzájárul a pénzügyi piacok működésének jobb megértéséhez, valamint ahhoz, hogy a torzítást figyelembe véve pontosabb modelleket építhessünk.

1. Introduction

In this paper, I estimate bounds for the parameter characterizing analysts' loss functions in making stock recommendations. In a binary variable forecasting environment, it is possible to set-identify the parameter that accounts for the forecaster's relative cost for overestimating versus underestimating the target even if the stock analyst's information set is not fully observed (Lieli and Stinchcombe (2013)). In this empirical application of the Lieli and Stinchcombe result, I use binarized stock recommendations as forecasts: buy recommendations account for positive, while hold or sell recommendations account for negative forecasts. The forecast is compared to the one-month-ahead price performance of the stock relative to the market. In the estimation, I also use a proxy for the publicly observed part of the forecaster's information set. The proxy I use is the smooth price per equity ratio. I have chosen this proxy by following Campbell and Thompson (2008), who show that the smooth P/E ratio could be used to predict excess stock returns once weak restrictions hold for the signs of coefficients[1].

My empirical results show high sensitivity to the categorization of hold recommendations. When I assume that 'hold' means 'sell', the estimated asymmetry parameters are relatively high. This suggests that we can rule out analysts' extreme reluctance to propose a 'sell'; they are more likely to issue 'false sells' than 'false buys'. However, when categorizing 'hold' into the buy category, the reverse is found: in almost all cases the highest possible values for the asymmetry parameter are ruled out. When imputing 'hold' with the previous recommendation, again the highest values are ruled out in more than half of the cases.

While financial professionals do not all agree on the information content of analyst stock recommendations, their widespread use and several pieces of evidence from the literature confirm that they are in fact relevant and useful forecasts for the future performance of stocks. It has been shown that analysts' earnings forecasts are superior to mechanical time series models (Brown and Rozeff (1978), Bradshaw et al. (2012)). Empirical evidence also shows that recommendations have some investment value, as they are successful in predicting short-run stock returns (Womack (1996), Loh and Mian (2006)). In their 1998 paper, Barber et al. document that an investment strategy based on the consensus recommendations of security analysts earns positive returns. For the analyzed period between 1986 and 1996, purchasing stocks most highly recommended and selling short those with the worst recommendations yielded a return of 102 basis points a month (Barber et al (1998)). The statement from Barber et al. is confirmed by more recent findings as well: see Jegadeesh et al. (2004) and Green (2006).

Another straightforward argument on the relevance of analyst recommendations is that brokerage houses produce and sell them for millions of dollars every year[2]. If they were in fact useless, why would so much money be spent on their production and sale?

We can see that analyst stock recommendations are in fact relevant. This is also confirmed if we look at the massive attention analyst recommendations get in the academic literature (for a comprehensive picture, see the review on the financial analyst forecasting literature by Ramnath et al. (2008)).

We can deduct some important inference from this large body of academic literature on what characteristics of analysts' recommendations are rewarded. First, unsurprisingly, evidence suggests that forecast accuracy is important for an analyst's prestige and career prospects. In their 2003 paper, Hong and Kubik relate earnings forecasts made by security analysts to job separations. They find that forecast accuracy is indeed a substantial factor in an analyst's career outcomes, such as how prestigious is her employer brokerage house, or what kind of stocks is she assigned to cover (Hong and Kubik (2003)). Forecasts are not directly evaluated on their accuracy, but for building reputation and influence among the buy side, it is substantial for the analyst to make the right calls (Hong and Kubik (2003)).

Although accuracy is important, evidence suggests that it is not everything: for the best career perspectives, an analyst also has to publish relatively optimistic recommendations. Controlling for accuracy, analysts who issue a large fraction of forecasts that

[1] An earlier version of this paper appeared in the Spring Wind 2016 conference volume (Grolmusz (2016)).

[2] A first year equity analyst earned a yearly base salary of $68,200 plus a bonus of $48,100 on average in 2013, as reported by the Wall Street Oasis 2013 Compensation Report (Rapoza (2013).

are more optimistic than the consensus are much more likely to move up the career hierarchy ladder (Hong and Kubik (2003)). This observation is confirmed by Lim (2001), among others. Lim argues that incorporating positive bias in earnings forecasts is a rational action.

Anecdotal evidence also supports the above statement. Lim argues that it is widely known throughout the financial analyst profession that a negative report on a company might result in the involved company's management limiting or eliminating the pessimistic analyst's information flow (Lim (2001)). Other pieces of anecdotal evidence emphasize that analysts need to go along with the management's optimistic projections, or if they do not, they risk being passed over for more loyal analysts (Hong and Kubik (2003), Lim (2001)). The importance of following the management's guidelines is even higher for young and inexperienced analysts, as their risk of unfavorable job separation is much higher than it is for their older colleagues (Hong et al. (2000)). This is the reason why younger analysts tend to avoid making bold forecasts and are more likely to herd (Hong et al. (2000)).

Different theories on the driving forces behind creating analyst recommendations suggest different implications for the direction of bias in the observed recommendations. The above arguments support low risk aversion in analysts for making buy-side recommendations: as analysts are rewarded for issuing relatively optimistic recommendations, they tend to incorporate a positive bias into their recommendations. However, sound arguments for the reverse can also be found. Consider that if an analyst issues a buy recommendation, then in the case of underperformance of the stock, her client will lose money for sure. However, if the analyst recommendation is 'sell', then the client can still lose in the sense of opportunity cost, but it might not be as painful for her (due to loss aversion), and the client might not even observe the performance of the stock as it is not anymore in her portfolio. This argument suggests that a risk-averse analyst should only issue a 'buy', if the probability of the stock outperforming the market is very high. Thus, analysts should be motivated to avoid making overly optimistic recommendations.

The contribution of this paper to the literature is twofold. First, I derive confidence intervals for the bounds of the loss function asymmetry parameter introduced by Lieli and Stinchcombe (2013), and second, I develop an empirical application of their theoretical result in a binary forecasting setting. More concretely, I inspect stock analysts' relative costs for overprediting versus underpredicting the stock's performance, by using a flexible and general method that has not been used up to now. By doing this, I am able to draw conclusions on the relative empirical relevance of the above two channels.

The remainder of the paper is organized as follows. In section 2, I outline the theoretical background for preference recovery in a binary forecasting environment, relying on the results from Lieli and Stinchcombe (2013). In section 3, I introduce the methodology and the data used in the empirical application. Section 4 presents and interprets the results, while the last section concludes.

2. Preference Recovery in a Binary Forecasting Environment

The theoretical background for the empirical investigation used in this paper comes from the 2013 paper of Lieli and Stinchcombe. In a binary variable forecasting environment, Lieli and Stinchcombe's paper provides a set identifiction result for the parameter characterizing the forecaster's loss function. In this section, I summarize this theoretical result.

2.1 EXPECTED LOSS MINIMIZATION PROBLEM

Let Y_t be the time series of binary values, and $\widehat{Y}_t$ be the time series of Y_t's forecasts made in the previous period ($\widehat{Y}_t = \widehat{Y_{t|t-1}}$); $t = 1, 2, \ldots, T$; $T < \infty$. In a binary variable forecasting setting, $Y_t, \widehat{Y}_t \in \{0, 1\}$, and a loss function can be represented in the following way:

	$Y_t = 1$	$Y_t = 0$
$\widehat{Y}_t = 1$	0	$\ell(1, 0)$
$\widehat{Y}_t = 0$	$\ell(0, 1)$	0

Where $\ell(\widehat{Y}_t, Y_t)$ is the loss from forecasting $\widehat{Y}_t$ when the realization will be Y_t. We assume that $\ell(1,0) \geq 0$ and $\ell(0,1) \geq 0$. For expected loss minimizing forecasters, assuming that the loss is zero when the forecaster *hits* the target ($Y_t = \widehat{Y}_t$) is true without loss of generality[3].

We assume that forecasters produce their forecasts by minimizing expected loss. Let I_t denote the information set of the forecaster. Then the forecaster solves the following problem:

$$min_{\widehat{Y}_t \in \{0,1\}} \ell(\widehat{Y}_t, 0) P(Y_t = 0 \mid I_t) + \ell(\widehat{Y}_t, 1) P(Y_t = 1 \mid I_t)$$

The solution to this problem is to predict one if $P(Y_t = 1 \mid I_t) > c$, where $c = \frac{1}{1 + \frac{\ell(0,1)}{\ell(1,0)}}$, $c \in [0, 1]$. Let us denote c as the asymmetry parameter. The asymmetry parameter depends on the forecaster's relative loss from overpredicting versus underpredicting the target. It is the parameter I would like to estimate.

The key identification problem is that the econometrician does not observe the whole information set on which the forecast is based, but only the public part of it. Following Lieli and Stinchcombe (2013), let us partition the information set I_t into two subsets: the part that the econometrician also observes, Z_t, and the private information of the forecaster, Z_t'. The forecast is based on the whole information set that is only partly observed by the econometrician, that is, the forecaster predicts one if $p_{Z_t, Z_t'} \equiv P(Y_t = 1 \mid Z_t, Z_t') > c$. Therefore, the econometrician cannot identify the asymmetry parameter exactly, she can only estimate a set in which the parameter lies (Lieli and Stinchcombe (2013)).

Let us define the unconditional and sample probabilities of Y_t and $\widehat{Y}_t$ in the following way:

[3] This fact is due to the following standardization: $\ell^c(\widehat{Y}_t, Y_t) = \ell(\widehat{Y}_t, Y_t) - \ell(Y_t, Y_t)$, where ℓ^c is the canonical form of the loss function (Lieli and Stinchcombe (2013)).

$$p = P(Y_t = 1)$$
$$q = P(\widehat{Y}_t = 1)$$
$$\widehat{p}_T = \frac{1}{T}\sum_{t=1}^{T} Y_t$$
$$\widehat{q}_T = \frac{1}{T}\sum_{t=1}^{T} \widehat{Y}_t$$

Then, we can define p_{Z_t} and q_{Z_t} as probabilities conditional on Z_t; the part of the forecaster's information set that the econometrician also observes:

$$p_{Z_t} = P(Y_t = 1 \mid Z_t)$$
$$q_{Z_t} = P(\widehat{Y}_t = 1 \mid Z_t) = P(p_{Z_t,Z_t'} > c \mid Z_t),$$

where q_{Z_t} is the proportion of times $\widehat{Y}_t = 1$ is observed conditional on Z_t. It is true by the law of iterated expectations [4], that $E[p_{Z_t,Z_t'} \mid Z_t] = p_{Z_t}$.

Using this relationship, Lieli and Stinchcombe (2013) derive the following bounds for the asymmetry parameter:

$$\frac{p_{Z_t} - q_{Z_t}}{1 - q_{Z_t}} \leq c \leq \frac{p_{Z_t}}{q_{Z_t}}.$$

Let us denote the lower bound as L_t, and the upper bound as U_t: $L_t = \frac{p_{Z_t} - q_{Z_t}}{1 - q_{Z_t}}, U_t = \frac{p_{Z_t}}{q_{Z_t}}$. It is easy to show that $L_t \leq U_t$. It can happen that $U_t \geq 1$ or $L_t \leq 0$, in these cases the bound is not informative. p_{Z_t} and q_{Z_t} could be estimated from the data using logit regressions, and using these estimates, we can give lower and upper bounds L_t and U_t for c.

Lieli and Stinchcombe highlight that their result is very general, as there are no assumptions about the number of omitted variables Z_t', nor about their distributions. This makes loss function parameter identification possible in a general framework.

2.2 CONFIDENCE INTERVALS

To check the statistical significance of the estimates, we need to derive confidence intervals. I do this by setting up a central limit theorem for the averages $\widehat{p}_T$ and $\widehat{q}_T$, and derive the variances for the estimated upper and lower bounds that are approximated as linear combinations of $\widehat{p}_T$ and $\widehat{q}_T$.

Definition:

Let Γ_h be the following:

$$\Gamma_h = E\left[\begin{pmatrix} Y_t - p \\ \widehat{Y}_t - q \end{pmatrix} \begin{pmatrix} Y_{t-h} - p & \widehat{Y}_{t-h} - q \end{pmatrix}\right], h = 0, \pm 1, \pm 2, \ldots$$

[4] $p_{Z_t} = E[Y_t \mid Z_t] \overset{\text{LIE}}{=} E[E(Y_t \mid Z_t, Z_t') \mid Z_t] = E[p_{Z_t,Z_t'} \mid Z_t]$

L and U are the lower and upper bounds for the asymmetry parameter c:

$$L = \frac{p-q}{1-q} \leq c \leq \frac{p}{q} = U;$$

$$\widehat{L}_T = \frac{\widehat{p}_T - \widehat{q}_T}{1-\widehat{q}_T},\ \widehat{U}_T = \frac{\widehat{p}_T}{\widehat{q}_T}$$

Assumptions:

- Y_t and $\widehat{Y}_t$ are weakly stationary,
- Y_t and $\widehat{Y}_t$ have absolutely summable covariances: $\sum_{h=0}^{\infty} \Gamma_h < \infty$.

Theorem 1. *Distribution of $\widehat{U_T}$ and $\widehat{L}_T$*

1. $\sqrt{T}(\widehat{U_T} - U) \xrightarrow{d} N(0, \lambda_U' V \lambda_U)$,

   where $V = \sum_{h=-\infty}^{\infty} \Gamma_h$, *and* $\lambda_U = \begin{pmatrix} \frac{1}{q} \\ -\frac{p}{q^2} \end{pmatrix}$

2. $\sqrt{T}(\widehat{L}_T - L) \xrightarrow{d} N(0, \lambda_L' V \lambda_L)$,

   where $\lambda_L = \begin{pmatrix} 1 \\ \frac{p-1}{(1-q)^2} \end{pmatrix}$

Theorem 2. *Distribution of $\widehat{U_T}$ and $\widehat{L}_T$*

1. $\sqrt{T}(\widehat{U_T} - U) \xrightarrow{d} N(0, \lambda_U' V \lambda_U)$,

   where $V = \sum_{h=-\infty}^{\infty} \Gamma_h$, *and* $\lambda_U = \begin{pmatrix} \frac{1}{q} \\ -\frac{p}{q^2} \end{pmatrix}$

2. $\sqrt{T}(\widehat{L}_T - L) \xrightarrow{d} N(0, \lambda_L' V \lambda_L)$,

   where $\lambda_L = \begin{pmatrix} 1 \\ \frac{p-1}{(1-q)^2} \end{pmatrix}$

The proof is based on the central limit theorem. Appendix A contains the sketch of the proof.

3. Empirical Strategy and Data

In this section, I show the empirical strategy based on the theory outlined in section 2 that I use for the set-identification of the asymmetry parameter from analyst stock recommendations.

3.1 EMPIRICAL STRATEGY

To estimate the bounds given in section 2, we need to estimate p_{Z_t} and q_{Z_t}. If Z_t is an empty set, $\widehat{p}_T$ and $\widehat{q}_T$ are used to give the unconditional bounds for the asymmetry parameter.

If Z_t is non-empty, then p_{Z_t} and q_{Z_t} could be estimated using fitted values from the following logit regressions, using observations collected over time:

$$\widehat{p_{Z_t}} = logit(Z_t' \widehat{\beta_p}) = \frac{1}{1+e^{-Z_t' \widehat{\beta}_p}}$$
$$\widehat{q_{Z_t}} = logit(Z_t' \widehat{\beta_q}) = \frac{1}{1+e^{-Z_t' \widehat{\beta}_q}},$$

One can use the time series $\widehat{p_{Z_t}}$ and $\widehat{q_{Z_t}}$ (t=1, 2, ..., T) to derive L_t and U_t for every t. We could use different definitions for the overall bounds for c. We can either take max L_t and min U_t to be lower and upper bounds, respectively, or we could choose the minimum range min $(U_t - L_t)$ and denote its bounds as the overall highest and lowest bound. I use the latter method in the empirical exercise.

3.2 DATA

As forecast data, $\widehat{Y}_t$, I use monthly analyst stock recommendations for shares of Goldman Sachs and 3M Company. I have chosen these Blue Chip stocks because they are highly liquid and I have access to many individual analyst recommendations on them[5]. Analyst stock recommendations are usually published using similar rating scales, categorized into three to five levels. I standardize the different scales and binarize the recommendations in the following way: take $\widehat{Y}_t = 1$ if the recommendation is strong buy, buy, or equivalent, and take $\widehat{Y}_t = 0$ for sell, and strong sell recommendations. I impute missing observations with the previous recommendation. The categorization of hold recommendations is not straightforward, I use three different ways for treating these observations: imputing by zero (equivalent to sell), imputing by one (equivalent to buy), and imputing with the recommendation from the previous period. Imputing with the previous recommendation can be argued for if we treat 'holds' similarly to missing observations; I assume that an analyst issues a hold recommendation if she does not have any new information or expectation on the future behavior of stock price.

The time series I compare the forecasts to is Y_t, called the actual or realized series. I define Y_t to be one if the price growth[6] of Goldman or 3M Co. is positive and higher than the growth of the Dow Jones Industrial Average in one month from making the forecast:

$Y_t = 1$ if $\frac{P_{D,t+1}}{P_{D,t}} < \frac{P_{G,t+1}}{P_{G,t}}$ and $\frac{P_{G,t+1}}{P_{G,t}} > 1$

$Y_t = 0$ if $\frac{P_{D,t+1}}{P_{D,t}} \geq \frac{P_{G,t+1}}{P_{G,t}}$ or $\frac{P_{G,t+1}}{P_{G,t}} < 1$

[5] I use a Bloomberg terminal and Reuters Eikon for data collection.

[6] Price is taken to be the end-of-month closing price of Goldman and 3M Co. stocks. Analyst recommendations are also published at the end of each month.

where G: Goldman or 3M co., D: Dow Jones index

I compare the two stocks to the Dow Jones index, as Goldman Sachs and 3M Co. stocks are classic Blue Chip stocks. The length of the time series varies from analyst to analyst: it starts in 2003 the earliest (but in most cases, only after 2009), end ends in November 2016.

I present unconditional results along with conditional bounds, for which I include explanatory variables in the logit regressions. The included variable is a proxy for the public part of the analyst's information set used to make the recommendation. I follow Campbell and Thompson (2008), and use the smooth P/E ratio as a proxy for the analyst's information set. The data I use was accessed using Bloomberg and Reuters Eikon.

4. Empirical Results

In this section, I show and interpret the results from the empirical analysis. The estimation gives an upper and a lower bound for the asymmetry parameter of each analyst. The unconditional bounds are the estimates based on the sample averages $\widehat{p}_T$ and $\widehat{q}_T$. In the conditional case, upper and lower bounds are estimated based on the logit regression for every period t. Then, the largest lower bound L_t and smallest upper bound U_t are presented as the conditional bounds for the sample period.

4.1 INTERPRETATION

How could we interpret the results; e.g. what does a $[0, 0.25]$ result mean? Ruling out the highest values for the asymmetry parameter means that the representative analyst is not extremely risk-averse in proposing a buy strategy. In this case, let us assume that the asymmetry parameter takes its highest estimated value, 0.25. Then, by writing up the definition for c:

$$0 \leq \frac{1}{1 + \frac{\ell(0,1)}{\ell(1,0)}} \leq 0.25$$

$$\downarrow$$

$$3 \leq \frac{\ell(0,1)}{\ell(1,0)}$$

$$\downarrow$$

$$3 \times \ell(1,0) \leq \ell(0,1),$$

This means that a 'false sell' is at least three times as costly as 'false buy'. This would make the analyst reluctant to propose a sell strategy. If the upper bound is below 0.5, the analyst has asymmetric loss: she is more inclined to overpredict the target than to underpredict it. On the other hand, when the lower bound is above 0.5, the analyst is more likely to issue more pessimistic recommendations than overly optimistic ones.

It is important to analyze the relationship between the variation in the time series and their consequences on c in more detail. Let me show the consequences on c, when there is absolutely no variation in the recommendation series. If the analyst recommends to sell the stock and the recommendation stays the same ($\widehat{Y}_t = 0$) throughout the entire time series, then $p \in]0, 1[$ and $q = 0$. We assume that there is some variation in the binarized actual series.

$$\widehat{L}_T = p$$

$$\widehat{U}_T = \frac{p}{0} \rightarrow \infty$$

Similarly, if the analyst recommends to 'buy' the stock and the recommendation stays the same ($\widehat{Y}_t = 1$) throughout the entire time series, then $p \in]0, 1[$ and $q = 0$. We assume that there is some variation in the binarized actual series.

$$\widehat{L}_T = \frac{p-1}{0} \rightarrow -\infty$$

$$\widehat{U}_T = p$$

4.2 RESULTS

Table 1 shows the results for Goldman Sachs stocks, analyzed by fifteen brokerage houses in the sample. When we categorize hold as zero (hold is the same as a sell), we see that in eight cases, the lowest *c*'s are ruled out. This suggests that for these eight analysts, a 'false buy' is likely costlier than a 'false sell'. We cannot conclude that these analysts have undoubtedly asymmetric loss functions, as the lower bonds are below 0.5. These results are in line with the argument for high risk aversion in making buy side recommendations: it is less costly for the analyst to suggest a sell (or hold), as he expects the client not to observe the stock's price performance after taking it out from the portfolio. If there are many hold recommendations in the time series, observing high asymmetry parameters might be due to the categorization of 'holds' as 'sells'.

The unconditional bounds for Oppenheimer's analyst are uninformative. This is because there are exactly as many ones in the binary actual series than in the binary recommendation series. Therefore, $\widehat{p}_T = \widehat{q}_T$, and hence $\widehat{L}_T = 0$ and $\widehat{U}_T = 1$. For the rest of the sample (six analysts out of the fifteen), the highest *c*'s are ruled out: a 'false sell' is likely to be costlier than a 'false buy'.

The conditional bound intervals are narrower in all cases for hold=0 (column 2). This suggests that the smooth P/E ratio bears some forecasting power for stock price performance. In three cases (Wells Fargo, Macquarie and Oppenheimer) the estimated upper bound is lower than the estimated lower bound. In these cases, the estimated bounds are not informative.

When categorizing 'holds' as 1 (buy), the results change significantly. In all but one case, the highest asymmetry parameters are ruled out, suggesting that a 'false sell' is costlier than a 'false buy'[7]. This is in line with the argument for low risk aversion in making buy side recommendations: analysts might be biased towards optimistic recommendations. Analysts who are relatively more optimistic in their stock recommendations than the consensus can expect better career prospects, as it was shown by Hong and Kubik (2003).

The upper bounds are around 0.5 in most cases, suggesting certain asymmetry for *c*[8]. The conditional logit regressions produce results where the intervals for *c* become even narrower. E.g., we can conclude that Vining Sparks analysts are at least 5.25 times more likely to produce a 'false buy' than a 'false sell', when making recommendations for Goldman Sachs stocks.

In the last specification, we treat 'holds' similar to missing values and impute them with the previous recommendation. Depending on the exact time series, i.e. the typical recommendation and number of 'holds', this produces similar bounds as the hold=0 or the hold=1 categorization: in ten cases, the bounds are the same as in columns 1-2 (hold=0), and in five cases, they are equivalent to treating 'hold' as 1.

[7] The lower asymmetry parameters are ruled out in the estimated bounds for Societe Generale. This time series does not contain any hold recommendations, only 'sells'.

[8] I have not yet calculated the confidence intervals for the unconditional bounds. However, taking into consideration that in most cases, the upper bound or the lower bound is uninformative (i.e. $\widehat{L}_T = 0$ or $\widehat{U}_T = 1$) , it appears that the confidence intervals will be wide. This might change the interpretation of the results.

Table 1
Bounds for the asymmetry parameter, Goldman Sachs stocks

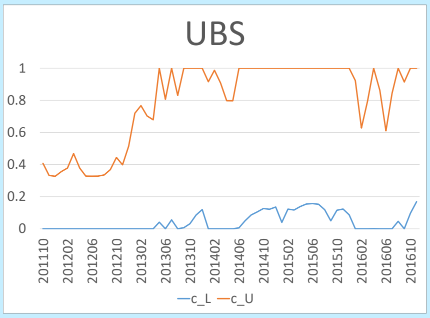
	hold=0		hold=1		hold=prev	
analyst's firm	unconditional	conditional	unconditional	conditional	unconditional	conditional
Wells Fargo	[0.14, 1]	[0.67, 0.5]	[0, 0.44]	[0, 0.29]	[0, 0.44]	[0, 0.29]
Nomura	[0.22, 1]	[0.34, 0.76]	[0, 0.47]	[0, 0.35]	[0.22, 1]	[0.34, 0.76]
Morgan Stanley	[0.27, 1]	[0.65, 0.68]	[0, 0.44]	[0, 0.3]	[0.27, 1]	[0.65, 0.68]
JMP	[0.38, 1]	[0.43, 1]	[0, 0.4]	[0, 0.35]	[0, 0.63]	[0, 0.44]
Barcleys	[0.41, 1]	[0.49, 1]	[0, 0.41]	[0, 0.32]	[0.41, 1]	[0.49, 1]
Macquarie	[0.43, 1]	[0.62, 0.23]	[0, 0.44]	[0, 0.23]	[0.43, 1]	[0.62, 0.23]
Societe Generale	[0.44, 1]	[0.53, 1]	[0.44, 1]	[0.53, 1]	[0.44, 1]	[0.53, 1]
RBC	[0.46, 1]	[0.52, 1]	[0, 0.53]	[0.28, 0.43]	[0.46, 1]	[0.52, 1]
Oppenheimer	[0,1]	[0.68, 0.36]	[0, 0.45]	[0, 0.31]	[0, 0.45]	[0, 0.31]
Atlantic	[0, 0.28]	[0, 0.02]	[0, 0.28]	[0, 0.02]	[0, 0.28]	[0, 0.02]
Credit Suisse	[0, 0.37]	[0, 0.18]	[0, 0.37]	[0, 0.18]	[0, 0.37]	[0, 0.18]
Vining Sparks	[0, 0.48]	[0.15, 0.23]	[0, 0.46]	[0, 0.16]	[0, 0.46]	[0, 0.16]
Rafferty	[0, 0.62]	[0, 0.34]	[0, 0.4]	[0, 0.25]	[0, 0.4]	[0, 0.25]
UBS	[0, 0.71]	[0.17, 0.33]	[0, 0.24]	[0, 0.2]	[0, 0.24]	[0, 0.2]
Evercore	[0, 0.93]	[0.1, 0.8]	[0, 0.41]	[0, 0.24]	[0, 0.94]	[0.1, 0.8]

Notes: The forecast value is the binarized analyst recommendation for Goldman stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Missing values in the recommendation series are imputed by the previous value. Hold recommendations are categorized as 0 (column 1 and 2), 1 (column 3 and 4), and imputed by the previous value (columns 5 and 6). The actual value is one if price growth for Goldman stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of Goldman Sachs in t.

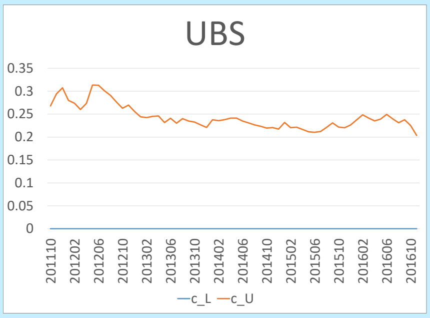
Figures 1, 2, 3 and 4 illustrate the sensitivity of the results on the categorization of 'holds'. In Figure 1, in UBS's case we see that the hold=previous specification gives the same bounds as the hold=1 (buy) specification. However, for Morgan Stanley (Figure 2), the bounds for hold=previous are the same as the bounds for hold=0 (sell). It can also happen that all three specifications produce different bounds (see Figure 3 for JMP), or in the absence of 'holds', all three pairs of estimates are the same (as for Credit Suisse, Figure 4).

The estimates on the other Blue Chip stock, 3M Company are quite similar to the results on Goldman Sachs. When hold is categorized as zero, c is relatively high in six cases (meaning that analysts are not too reluctant to propose a sell strategy). The lower bounds in the unconditional hold=sell case are on average lower than for Goldman Sachs estimates, all six are under 0.5. Therefore, these estimates do not rule out symmetric loss. The four remaining analyst have relatively low asymmetry parameters. For Jefferies and Credit Suisse, we can rule out symmetric loss as the upper bound is below 0.5. The conditional bound intervals become narrower than the unconditional intervals.

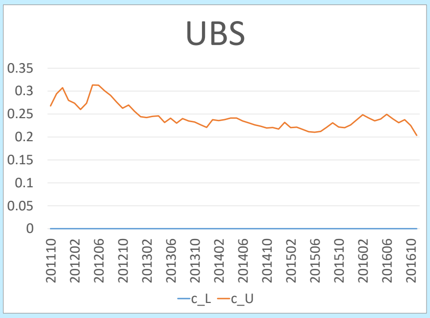
Figure 1
Bounds for c based on conditional probability estimates, UBS analyst recommendation for Goldman stocks



hold=0



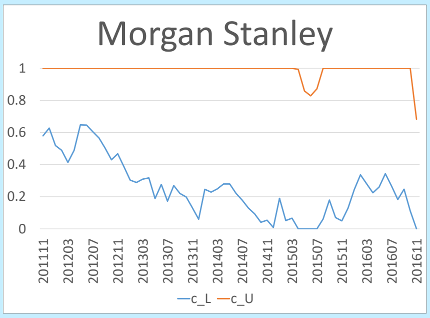
hold=1



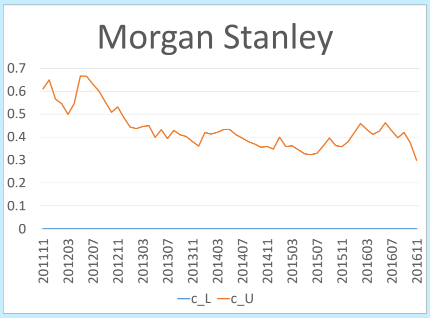
hold=previous

Notes: Bounds for c estimated using conditional logit regression. The forecast value is the binarized analyst recommendation for Goldman stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Hold recommendations are categorized as 0 (a), 1 (b), and imputed by the previous value (c). The actual value is one if price growth for Goldman stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of Goldman Sachs in t.

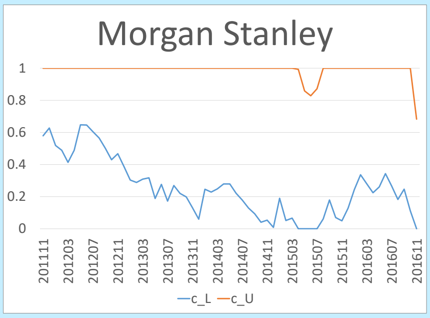
Figure 2
Bounds for c based on conditional probability estimates, Morgan Stanley analyst recommendation for Goldman stocks



hold=0



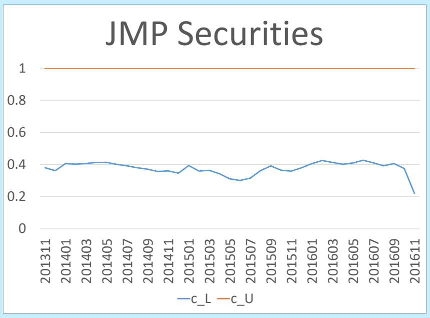
hold=1



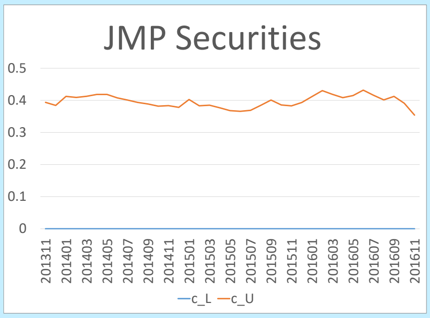
hold=previous

Notes: Bounds for c estimated using conditional logit regression. The forecast value is the binarized analyst recommendation for Goldman stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Hold recommendations are categorized as 0 (a), 1 (b), and imputed by the previous value (c). The actual value is one if price growth for Goldman stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of Goldman Sachs in t.

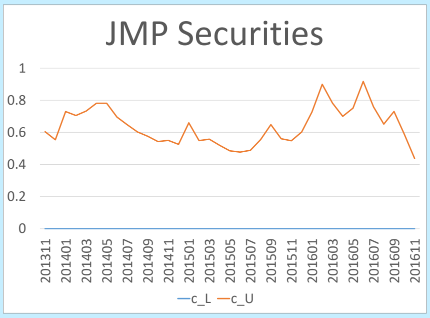
Figure 3
Bounds for c based on conditional probability estimates, JMP analyst recommendation for Goldman stocks



hold=0



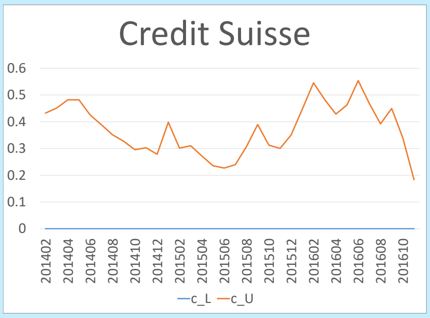
hold=1



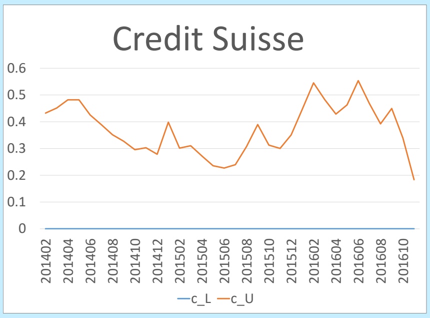
hold=previous

Notes: Bounds for c estimated using conditional logit regression. The forecast value is the binarized analyst recommendation for Goldman stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Hold recommendations are categorized as 0 (a), 1 (b), and imputed by the previous value (c). The actual value is one if price growth for Goldman stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of Goldman Sachs in t.

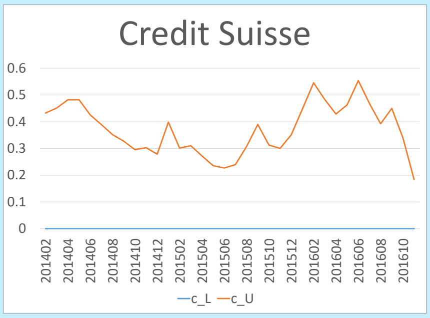
Figure 4
Bounds for c based on conditional probability estimates,Credit Suisse analyst recommendation for Goldman stocks



hold=0



hold=1



hold=previous

Notes: Bounds for c estimated using conditional logit regression. The forecast value is the binarized analyst recommendation for Goldman stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Hold recommendations are categorized as 0 (a), 1 (b), and imputed by the previous value (c). The actual value is one if price growth for Goldman stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of Goldman Sachs in t.

Table 2
Bounds for the asymmetry parameter, 3M Co. stocks

	hold=0		hold=1		hold=prev	
analyst's firm	unconditional	conditional	unconditional	conditional	unconditional	conditional
Bernstein	[0.11, 1]	[0.33, 0.45]	[0, 0.38]	[0.39, 0.36]	[0, 0.41]	[0, 0.36]
RBC	[0.13, 1]	[0.31, 1]	[0.13, 1]	[0.31, 1]	[0.13, 1]	[0.31, 1]
Morgan Stanley	[0.15, 1]	[0.3, 0.4]	[0, 0.86]	[0.32, 0.34]	[0.05, 1]	[0.28, 0.4]
Barcleys	[0.31, 1]	[0.39, 1]	[0, 0.33]	[0, 0.29]	[0.31, 1]	[0.39, 1]
Goldman Sachs	[0.34, 1]	[0.48, 1]	[0, 0.35]	[0, 0.13]	[0.34, 1]	[0.48, 1]
William Blair	[0.39, 1]	[0.42, 0.65]	[0, 0.42]	[0, 0.4]	[0, 0.42]	[0, 0.4]
Edward Jones	[0, 0.5]	[0, 0.44]	[0, 0.46]	[0, 0.33]	[0, 0.46]	[0, 0.33]
Jefferies	[0, 0.32]	[0, 0.19]	[0, 0.32]	[0, 0.19]	[0, 0.32]	[0, 0.19]
Credit Suisse	[0, 0.33]	[0, 0.16]	[0, 0.29]	[0, 0.16]	[0, 0.33]	[0, 0.16]
J.P. Morgan	[0, 0.79]	[0, 0.61]	[0, 0.3]	[0, 0.26]	[0, 0.67]	[0, 0.52]

Notes: The forecast value is the binarized analyst recommendation for 3M Co. stocks made in t (strong buy, buy: 1; sell, strong sell: 0). Hold recommendations are categorized as 0 (column 1 and 2), 1 (column 3 and 4), and imputed by the previous value (columns 5 and 6). The actual value is one if price growth for 3M Co. stocks is positive and outperforms the DJI one month from making the forecast, and zero otherwise. The explanatory variable Z is the smooth P/E ratio of 3M Co. in t.

In columns 3 and 4 in Table 2, we see that apart from RBC, the highest values are ruled out for c. This is similar to what I have found for Goldman stocks. The result is in line with the argument for low risk aversion towards buy strategies. Column 5 and 6 show the results for hold=previous. Here, in four of the cases the lowest values are ruled out, while in the other six cases $\hat{c}$ is relatively low.

We can see that the results are highly sensitive to the categorization of hold recommendations. If we take the hold=previous specification as baseline, we find that in the majority of cases, the highest values for c are ruled out.

5. Conclusion

In a binary variable forecasting environment, I carry out an empirical analysis to estimate bounds for the parameter characterizing the forecaster's loss function. I use analyst stock recommendations as forecast data, and compare it to the one-month-ahead relative price performance of the analyzed stock. In the conditional logit regressions, I include a proxy for the publicly observed part of the forecaster's information set as an explanatory variable. Using a theoretical result from Lieli and Stinchcombe (2013), I set-identify the parameter that captures the analyst's cost of over- versus underpredicting the target (asymmetry parameter). Another novelty of this chapter is the derivation of confidence intervals for the bounds of the loss function asymmetry parameter introduced by Lieli and Stinchcombe (2013).

Previous research suggests that incorporating positive bias in stock analyst's forecasts is a rational action (Lim (2001)). It is also shown that controlling for accuracy, analysts who frequently issue optimistic forecasts are rewarded: they are much more likely to be offered higher prestige positions, with higher wages (Hong and Kubik (2003)). Therefore, we can expect analysts to issue overly optimistic forecasts more easily than pessimistic ones.

The reverse side of the argument can also be supported by intuitive claims. Consider that if an analyst issues a buy recommendation, then in the case of underperformance of the stock, her client will lose money for sure. However, if the analyst recommends a sell strategy, then her client might not even observe if the stock indeed outperforms the market. This suggests that analysts should avoid proposing overly optimistic recommendations.

I find that the results are highly sensitive to the categorization of hold recommendations. When we assume that 'hold' means 'sell', the estimated asymmetry parameters are relatively high. This suggests that analysts are not very reluctant to propose a 'sell'. However, when categorizing 'hold' into the buy category, the reverse is found: in almost all cases the highest possible values for the asymmetry parameter are ruled out. When imputing 'hold' with the previous recommendation, again the highest values are ruled out in more than half of the cases. Developing additional empirical applications (i.e. other binary forecasting problems) for the identification of the loss function's asymmetry parameter would be an interesting area for further research.

References

Asness, C. S., T. J. Moskowitz and L. H. Pedersen (2013). Value and momentum everywhere. *The Journal of Finance, 68*(3), 929-985.

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.

Bradshaw, M. T., M. S. Drake, J. N. Myers and L. A. Myers (2012). A re-examination of analysts' superiority over time-series forecasts of annual earnings. *Review of Accounting Studies, 17*, 944–968.

Brown, L. D. and M. S. Rozeff (1978). The superiority of analyst forecasts as measures of expecta-tions: Evidence from earnings. *The Journal of Finance, 33*(1), 1-16.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *The Review of Financial Studies, 21*(4), 1509-1531.

Green, T. C. (2006). The value of client access to analyst recommendations. *Journal of Fi-nancial and Quantitative Analysis, 41*(01), 1-24.

Grolmusz, V. M. (2016). Recovering Stock Analysts' Loss Functions from Buy/Sell Recommendations. *Spring Wind II. Conference Volume*, 386-399.

Hong, H. and J. D. Kubik (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance, 58*(1), 313-351.

Hong, H., J. D. Kubik and A. Solomon (2000). Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics*, 121-144.

Jegadeesh, N., J. Kim, S. D. Krische and C. Lee (2004). Analyzing the analysts: When do recom-mendations add value?. *The Journal of Finance, 59*(3), 1083-1124.

Lieli, R. P. and M. B. Stinchcombe (2013). On the Recoverability of Forecasters' Preferences. *Econometric Theory, 29*(03), 517-544.

Lim, T. (2001). Rationality and analysts' forecast bias. *The Journal of Finance, 56*(1), 369-385.

Loh, R. K. and G. M. Mian (2006). Do accurate earnings forecasts facilitate superior investment recommendations?. *Journal of Financial Economics, 80*(2), 455-483.

Ramnath, S., S. Rock and P. Shane (2008). The financial analyst forecasting literature: A taxo-nomy with suggestions for further research. *International Journal of Forecasting, 24*(1), 34-75.

Rapoza, K. (2013, March 13). How Much Do Wall Streeters Really Earn?. *Forbes In-vesting*, Retrieved from http://www.forbes.com/sites/kenrapoza/2013/03/13/how-much-do-wall-streeters-really-earn/#3a92a01c7f08

Womack, K. L. (1996). Do brokerage analysts' recommendations have investment value?. *The Journal of Finance, 51*(1), 137-167.

Appendix A Sketch of proof of theorem 1

a) Using the delta-method, we can write $\widehat{U}_t$ in the following linear form (assuming that the second and higher order parts of the Taylor-expansion are zero):
$\sqrt{T}(\widehat{U}_T - U) \approx \frac{1}{q}\sqrt{T}(\widehat{p}_T - p) - \frac{p}{q^2}\sqrt{T}(\widehat{q}_T - q)$.
The central limit theorems for the univariate iid series $\widehat{p}_t$ and $\widehat{q}_t$ are the following:
$E(\widehat{p}_t) = p \quad Var(\widehat{p}_t) = p(1-p) < \infty$, then $\sqrt{T}(\widehat{p}_t - p) \xrightarrow{d} N(0, p(1-p))$
$E(\widehat{q}_t) = q \quad Var(\widehat{q}_t) = q(1-q) < \infty$, then $\sqrt{T}(\widehat{q}_t - q) \xrightarrow{d} N(0, q(1-q))$.
The Cramer-Wold theorem states that $X_n \xrightarrow{d} X$ if and only if $a'X_n \xrightarrow{d} a'X$ for all $a \in \mathbb{R}^k$. Let $pq \xrightarrow{d} N_k(0, \Sigma)$ then we can take any vector $a \in \mathbb{R}^k$;(k=2 in this case) and show: $a'\left[\sqrt{T}\widehat{p}_t\widehat{q}_t - pq\right] \xrightarrow{d} a'pq$. In the case of the upper bound, a), $a = \lambda_U$.

b) Using the delta-method, we can write $\widehat{L}_t$ in the following linear form (assuming that the second and higher order parts of the Taylor-expansion are zero):
$\sqrt{T}(\widehat{L}_T - L) \approx \sqrt{T}(\widehat{p}_T - p) - \frac{p-1}{(1-q)^2}\sqrt{T}(\widehat{q}_T - q)$.
Then, we use the Cramer-Wold device as in point a) for the upper bound, but now $a = \lambda_L$.

MNB Working Papers 2023/4

Recovering Stock Analysts' Loss Functions from Buy/Sell Recommendations

Budapest, December 2023

mnb.hu

1013 BUDAPEST, KRISZTINA KÖRÚT 55.