BBR BRAZILIAN BUSINESS REVIEW
Vol. 4, No. 2
Vitória-ES, May – Aug 2007
p. 103-113 *ISSN 1808-2386* *DOI: http://dx.doi.org/10.15728/bbr.2007.4.2.2*

Analysts'Optimism and Selection Bias

Antônio Lopo Martinez*
Federal University of Bahia

ABSTRACT: This paper is an empirical examination, drawing on the Institutional Brokers Estimate System (I/B/E/S) database, of analysts' earnings forecast optimism for Brazilian companies. The study found that analysts were optimistic on average and performed poorly in terms of precision and accuracy. The study period was January 1995 to December 2002. The forecasting errors in one period were correlated with the errors of the following period. There was evidence of persistent consensus errors among analysts, with those who were persistently optimistic outweighing those who were persistently pessimistic. A possible explanation for this predominant overoptimism is selection bias. In order to adjust analysts' consensus forecasts, an optimization methodology is suggested, providing results that minimize the optimism bias. The evidence presented is relevant, especially for those using analysts' earnings forecasts as an input in their stock valuation models.

Keywords: optimism, earnings forecasts, selection bias, analysts.

Received in 05/13/2007; revised in 05/26/2007 accept in 06/03/2007.

Corresponding authors:
** Universidade Federal da Bahia.*
Salvador, BA - Brazil
Telefone: (71) 96096556
antoniolopo@terra.com.br

Editor's note: *This paper was accepted by Alexsandro Broedel Lopes.*

1. INTRODUÇÃO

An important part of market analysts' job is to project future earnings. With these predictions, analysts estimate a crucial variable for stock valuation models. Good projections of future earnings are prerequisites for adequately measuring the fair price of a stock.

In this context, it is important to analyze the characteristics (or properties) of these projections, generated by capital market analysts for Brazilian companies. Are they accurate? Is there some type of bias in these forecasts? Knowing the accuracy of analyst's predictions is an important aid to guide investors regarding the validity (or limitations) of these projections for their valuation models.

This study is focused on analysts' consensus. The consensus is the average of earnings predictions for a company in a determined period, and is known as the street consensus. The analysis of this consensus is based on the idea that a representation of the market expectations can be obtained by a measure of the central trend of the distribution of analysts' projections.

In general lines, the analyses performed here allow the conclusion that just as for analysts who price international companies, forecasts for Brazilian companies tend to be biased, being significantly optimistic.

In this paper, I discuss the methodological concepts and procedures followed for the study of analysts' forecasts about Brazilian companies. I identify the database used and the characteristics, such as accuracy, precision and bias for the distribution of forecasting errors.

I discuss selection bias as one of the possible explanations for the optimistic bias found for analysts' predictions. In this respect, after a careful assessment of the problem, I propose a method aiming to adjust the consensus forecasts to make them more accurate.

The article ends with a set of conclusions with implications for those who consider earnings forecasts in company valuation models.

2. REVIEW OF THE LITERATURE: WHAT IS KNOWN ABOUT ANALYSTS' EARNINGS FORECASTS?

The great majority of studies in the literature have concluded that analysts are optimistic. This optimistic bias is inferred from the finding that there is a systematic negative difference between the real and estimated profits. In other words, predicted earnings are predominantly above those actually obtained. This optimism has been documented using Value Line, I/B/E/S and Zacks. The estimates of analysts' optimism vary among the studies, partly because of differences in the methodologies used, the definition of the variables and the time periods analyzed.

LIM (1998), using the average of quarterly earnings estimates, found optimism of 0.94% for stock prices. The bias was considerably higher, 2.5%, for small firms, against 0.53% for companies with large market capitalization. The bias was predominant for the entire market and for all years studied. RICHARDSON et al. (1999) used the forecasts of individual analysts and their forecasting errors in each month. They also indicated that although bias continued to exist, there was a significant decrease in it, from a magnitude of 0.91% of the price to 0.09% of the price, whenever the forecasting horizon was reduced from a year to a month. In turn, BROWN (1998), studying more recent periods, observed that the bias appeared to be changing from optimistic to pessimistic, or at least to practically nil.

A possible explanation, commonly found in the literature on bias, is the existence of a high number of extreme observations, or outliers, that cause the distribution of forecasting errors to be asymmetric. GU & WU (2003) and ABARBANELL & LEHAVY (2003b) observed that a small number of forecasting errors contributed to the bias observed.

For analysts of Brazilian companies, an optimistic bias has also been documented, by DA SILVA (1998) and FRANCO (2000). Although they used different methods and databases, trying to solve distinct problems, they both reported the existence of optimism in the predictions of analysts of Brazilian firms.

3. METHODOLIGICAL ASPECTS

3.1. Database

To analyze the precision of the projected earnings of Brazilian companies, I used the data collected by the Institutional Brokers Estimate System (I//B/E/S). Since 1971, this system has been the most important source of earnings predictions for investment professionals throughout the world.

The study covered the period from January 1995 to December 2002. In this part of the study, I used all the companies for which the database had information, with no type of filtering. All told, there were 239 listed Brazilian corporations analyzed regarding analysts' future earnings projections.

As the first part of the analysis, I gathered information on analysts' consensus projections for **earnings per share (EPS)** for the next year. Among the various yardsticks available, I found that in EPS predictions, the current year was the forecast with the greatest number of observations. This is a key variable for valuations based on indicators of the price/earnings (P/E) type. Unlike in the United States, where the majority of predictions are aimed at quarterly results, in Brazil annual results predominate.

I collected the EPS projections for a determined year on a month-to-month basis. The I/B/E/S ascertains monthly the consensus of analysts for all the forecasts up to the month previous to the publication of the earnings results. Thus the system registers analysts' consensus for EPS of a determined year until the month before the release of the results.

3.2. Metric for the forecasting errors

As a metric to identify analysts' performance in their projections, I calculated the prediction error (*ErrPred*). This error was calculated by the difference between the real (observed) result and the result projected (estimated) by analysts. A negative prediction error means a negative surprise: the projected earnings were overestimated, i.e., they were higher than actually occurred. On the other hand, when the actual earnings turned out to be greater than the consensus estimates, the surprise is positive.

For the effect of comparability, I computed the forecasting errors in terms of actual EPS. Hence, the prediction error for this study is the actual earnings minus analysts' prediction, divided by the absolute value (modulus) of the actual result for the period:

$$Err\,\mathrm{Pr}\,ev = \frac{EPS_{real} - EPS_{\mathrm{Pr}\,ev}}{\left| \; EPS_{real} \; \right|}$$

where:

EPS_{real} Actual earnings per share in the period

EPS_{pred} Analysts' consensus (average) earnings per share

I used the absolute value in the denominator to capture with precision the direction of the prediction. Dividing the numerator by the absolute value of the real earnings allows comparability in percentage terms.

From a methodological standpoint, there are various other yardsticks to measure forecasting errors. Besides actual results (profit or loss verified), one can often find in the international literature prediction errors measured in terms of percentage of price per share. I believe that these prices would bring distortions because the prediction errors would be measured in terms of a factor over which the analysts have no control.

Similarly, I did not use total assets (or net equity) as a factor to deflate the forecasting errors. First, I believe that the assets could be correlated with the prediction errors in an undesirable manner. If working with figures deflated by assets, one would in essence be measuring an indicator of return on assets. Certain transactions have a greater return on assets (ROA) than others. This factor could compromise the comparability of the forecasting errors ascertained.

I recognize, however, that measuring the prediction error in terms of the real earnings result is not free from flaws and problems. For example, for companies recording earnings very near zero, the formula chosen produces exaggeratedly high prediction errors. Additionally, instances where earnings were effectively nil had to be discarded because the denominator was zero.

4. CHARACTERISTICS OF THE DISTRIBUTION OF ANALYSTS' FORECASTING ERRORS

In the international literature, certain authors use the median of the estimates as the analysts' consensus. I chose to use the average, because it more precisely reflects the magnitude of the estimates, not only their number.

I used the mean of the prediction errors (MPE) to verify the possible existence of **bias**. If the MPE indicates a negative value, it means that in average terms the forecasting errors are negative (negative surprise), meaning that the predictions were higher than the results actually observed. A negative and significant MPE thus is evidence of an optimistic bias in the forecasts. The MPE is calculated by the formula below, where n is the number of prediction errors (*ErrPred*).

$$MPE = \left(1/n \right) x \sum_{i=1}^{n} Err\,\Pr ev$$

The **precision** is estimated by the inverse ratio of the standard deviation of the prediction errors (*ErrPred*). Therefore, the smaller the standard deviation, the more precise analysts' forecasting errors are. In algebraic terms, the proxy for precision was calculated as follows:

$$D.P. = \sqrt{\frac{\sum_{i=1}^{n} \left(Err\,\Pr ev_i - MPE \right)^2}{(n-1)}}$$

In order to estimate the **accuracy**, I sought to consider the distribution of errors that, in absolute terms, were nearer to zero, i.e., treating the prediction error in the same way regardless of its being positive or negative. In evaluating the accuracy, all the errors are considered. To estimate the bias, positive errors offset negative ones of the same magnitude.

The variable used to determine accuracy was the mean of the absolute prediction errors (MAPE): the further from zero the value of MAPE is, the higher were the forecasting errors computed.

$$MAPE = \left(1/n \right) x \sum_{i=1}^{n} \left| Err\,\Pr ev \right|$$

5. HOW DO THE FORECASTING ERRORS BEHAVE OVER TIME?

5.1. Monthly trend of annual forecasting errors

Analysts revise their expectations of future earnings as new information becomes available. It is reasonable to assume that as a year progresses, analysts will be able to predict the outcome of a particular variable for that year more exactly. In other words, it is logical that the consensus of analysts' annual earnings estimates will be better in the later months of the year than in earlier ones.

To investigate this hypothesis, I stratified the observations of analysts' EPS forecasting errors into months and years. **Table 1, Panel A** shows the information and the calculation of the mean and Student's *t* statistic, which checks the null hypothesis that the mean is different than zero.

At this point it can be noted that for the total values, as well as for all the months in general, there is a perceptible decrease in the prediction errors as the year progresses. This confirms the hypothesis that analysts, on average, appear to revise their forecasts, adjusting them nearer to the real outcomes.

Notwithstanding the significant fall in the forecasting errors during the year, the mean of these errors for all the Decembers analyzed is near -0.88 (t = -4.26). Thus, there is still a clear optimistic outlook among analysts in the last month of the year for which the prediction is being made. An exception was December 2000, when there was actually a small increase in the negative prediction errors. However, according to the *t* statistic, these values are not significantly different from zero. This analysis makes it clear that consensus forecast in the last month of years studied was the least skewed one.

The indicators of precision (DP) and accuracy (MAPE) did not show any significant improvement. In reality, these measures did not follow the same continuing falling pattern as the MPE. In other words, although the optimistic bias went down, the precision and accuracy did not evolve in the same direction as the end of the year approached. Therefore, despite the reduction in the bias, the accuracy continued to be compromised due to the low precision of the estimates.

5.2 Yearly trend of annual forecasting errors

Still in the temporal plane, I examined the behavior of the forecasting errors over the years. The analysts' earnings predictions had marked variability, on average, over the eight years studied. The yearly figures were always significantly negative, although these were greater in some years and smaller in others.

It would be interesting to clarify whether the analysts' forecasts become less biased with passing time. This point has been demonstrated in the international literature. An analysis of the means did not reveal a clear trend for declining bias, but there also was no way to speak of an increase in overoptimism.

Unquestionably, 1995 was the year with the greatest prediction errors, when analysts were the most overoptimistic. On the other hand, the forecasts were nearest zero in 1999. In reality, in various months of that year the values were not significantly different than zero (indicating an absence of bias). The last year covered, 2002, also had peculiar characteristics: in some months the analysts were pessimistic in their forecasts, especially in the second half of the year (although the values were never significantly positive).

Analysts' earnings projections are very sensitive to the market expectations, the economic circumstances and even to political factors, which wind up significantly influencing the profile of optimism or pessimism with which analysts see the future.

To empirically test the trend of the behavior of the forecasting errors over the years and months, I ran some regressions. The models estimated have the following form: $MPE_t = \beta_1 + \beta_2\, t + \varepsilon_t$ and $MAPE_t = \varphi_1 + \varphi_2\, t + \varepsilon_t$, where t is the unit of time, which can be a year or month, depending on the trend that is being verified.

According to GUAJARATI (1995: 171), this type of model is called a **linear trend model**. Trend is defined as a variable's sustained increasing or decreasing movement. If the slope coefficient (β_2) is positive, there is an increasing MPE (or MAPE) trend; if it is negative, the trend in either case is declining.

The results of these regressions are shown in **Table 1**, **Panel B**. From a statistical standpoint, there is a marked trend for the forecasting errors to improve over the months of the year (0.073). The positive sign indicates that as the months pass, the optimistic bias is reduced. In other words, as the end of the year approaches, analysts get a better perception of the firm's result.

Regarding the behavior over the years, the tendency was positive. However, in statistical terms, this value was not significant at the traditional standards (t = 1.583 and adjusted R^2 = 17.70%). Hence, it was not clear whether there was an improving trend in terms of annual forecasting errors, but there certainly was no worsening trend.

Additionally, I performed some other regressions to check whether there was a fall in the dispersion of the errors, but the results were not statistically satisfactory (and are not shown in the table).

It can be argued that each year has its particularities: the bias and accuracy depend on the facts that occurred. To test this hypothesis, I used the KRUSKAL-WALLIS test (not reported in the table), which is very useful to decide if k independent samples (k>2) come from populations with equal means. This test is a functional alternative to variance analysis.

In this test, I computed χ^2 of 11.206 (*sig* 0.130) for the MPE and χ^2 of 22.924 (*sig* 0.002) for the MAPE. The test indicated that the years are significantly different in terms of accuracy, but are not regarding bias. Therefore, this test reinforces the idea that the particularities of each year may be explaining a higher or lower level of accuracy (MAPE). However, regarding the bias (MPE), it cannot be categorically affirmed that the years are different, indicating that the magnitude of bias can be a general characteristic of all the samples, regardless of the year.

Table 1 – Forecasting errors of annual EPS and their trend, analysts' consensus, in the months of April to December between 1995 and 2002

This table documents the statistics on the forecasting errors of analysts' consensus for EPS and their trend, computed from I/B/E/S data. Panel A documents the mean of the forecasting errors, standard deviation, number of observations and the Student's t-statistic, which tests the differences of the mean from zero. The left column shows the months in which the consensus was determined. The rows are the years included in the study period. Panel B shows the trend of the mean of the forecasting errors (β_2) and the *t* statistic of the linear regression of the forecasting errors over time. The adjusted R^2 refers to these regressions**.**

Panel A: Distribution of the forecasting errors by months and years

Month/Year	STAT.	1995	1996	1997	1998	1999	2000	2001	2002	TOTAL	N.OBS.
ABR	M.P.E	-3,20	-1,17	-0,91	-1,05	-1,02	-1,42	-1,63	-1,12	**-1,50**	879
	D.P.	12,34	3,04	2,75	3,66	4,92	5,63	6,73	3,21	**6,43**	
	M.A.P.E	3,45	1,50	1,19	1,54	2,10	1,73	1,84	1,32	**1,89**	
MAY	M.P.E	-3,22	-1,09	-0,80	-1,07	-1,02	-1,51	-1,42	-1,04	**-1,44**	973
	D.P.	12,83	2,95	2,60	3,60	4,87	5,45	5,20	3,23	**6,34**	
	M.A.P.E	3,48	1,43	1,24	1,54	2,01	1,78	1,62	1,26	**1,85**	
JUN	M.P.E	-3,15	-1,00	-0,75	-1,02	-0,71[#]	-1,32	-1,35	-1,13	**-1,34**	998
	D.P.	12,54	3,15	2,42	3,44	5,94	5,22	5,30	3,40	**6,32**	
	M.A.P.E	3,40	1,52	1,15	1,50	2,24	1,67	1,56	1,34	**1,85**	
JUL	M.P.E	-3,10	-1,00	-0,74	-0,90	-0,32[#]	-1,25	-1,41	-0,73	**-1,24**	996
	D.P.	12,45	2,96	2,42	3,08	6,81	4,86	5,56	1,67	**6,34**	
	M.A.P.E	3,35	1,35	1,15	1,37	2,17	1,61	1,60	0,96	**1,77**	
AUG	M.P.E	-2,85	-0,96	-0,76	-1,60	0,24[#]	-1,04	-1,25	-0,73	**-1,20**	1007
	D.P.	11,33	2,80	2,37	9,19	7,56	3,22	5,47	1,63	**6,84**	
	M.A.P.E	3,06	1,29	1,14	2,05	1,82	1,35	1,48	0,95	**1,73**	
SEPT	M.P.E	-2,91	-0,97	-0,76	-1,59	0,40[#]	-1,01	-1,17	0,25[#]	**-1,10**	1010
	D.P.	11,05	2,71	2,54	9,25	7,51	3,26	5,61	7,73	**7,04**	
	M.A.P.E	3,13	1,28	1,16	2,04	1,80	1,34	1,43	1,82	**1,78**	
OCT	M.P.E	-2,61	-0,86	-0,72	-1,53	-0,17[#]	-0,73	-0,91	0,20[#]	**-1,04**	1007
	D.P.	10,06	2,44	2,22	9,29	2,52	2,34	3,82	7,69	**6,10**	
	M.A.P.E	2,82	1,15	0,96	1,98	1,12	1,04	1,19	1,83	**1,53**	
NOV	M.P.E	-2,17	-0,80	-0,68	-1,43	-0,50	-0,69	-0,83	-1,21	**-1,07**	1016
	D.P.	9,18	2,34	2,24	8,77	3,74	2,28	3,67	4,64	**5,59**	
	M.A.P.E	2,49	1,11	0,93	1,88	1,46	1,00	1,22	1,36	**1,47**	
DEC	M.P.E	-1,53	-0,55	-0,66	-1,19	-0,40[#]	-1,47[#]	-0,78	0,25[#]	**-0,88**	998
	D.P.	9,62	2,29	2,16	7,24	3,74	9,82	3,65	7,05	**6,51**	
	M.A.P.E	2,41	0,84	0,87	1,63	1,42	1,79	1,16	1,60	**1,49**	
TOT. YEAR	M.E.P	**-2,74**	**-0,94**	**-0,75**	**-1,27**	**-0,37**	**-1,16**	**-1,19**	**-0,59**	**-1,20**	
	D.P.	**11,34**	**2,77**	**2,41**	**7,04**	**5,57**	**5,23**	**5,10**	**5,06**	**6,40**	
	M.A.P.E	**2,78**	**1,20**	**1,07**	**1,73**	**1,71**	**1,43**	**1,42**	**1,26**	**1,63**	
N.OBS.	\	1393	1360	1322	1306	1103	962	836	602		8884

Panel B: Definition of the trend of the forecasting errors (months and years)

$MPE_t = \beta_1 + \beta_2 + \varepsilon_t$ $MAPE_t = \alpha_1 + \alpha_2 t + \varepsilon_t$

	β_2	*t*-stat	sig	adjust R^2
MEP (Years	0.161	1.583	0.165	17.70%
MEP (Months)	0.073	13.717	0.000	95.90%

	α_2	*t*-stat	sig	adjust R^2
MEP (Years	0.115	1.278	0.249	8.30%
MEP (Months)	0.056	6.866	0.000	85.20%

6. HOW DO THE ERRORS BETWEEN DIFFERENT PERIODS CORRELATE?

To identify the trend of the magnitude of the forecasting errors, I used regression analyses to explain the forecasting error in a period *t* by the forecasting error in the previous period *t*-1. I estimated a statistical relation aiming at determining whether the errors were correlated over time. I implemented this process with the suitable adjustments for autocorrelation for each type of prediction error (positive or negative) and for the combined total of the errors.

The regression equation was as follows:

$$(\boldsymbol{ErrPred})_{\mathbf{t}} = \boldsymbol{\delta_1} + \boldsymbol{\delta_2}\,(\boldsymbol{ErrPred})_{\mathbf{t-1}} + \boldsymbol{\varepsilon_t}$$
where $\varepsilon_t \sim N(0,\sigma^2)$; $E(\varepsilon_{i,}\ \varepsilon_j) = 0$ for $\forall i \neq j$

This regression can be interpreted as follows: δ_1 is the mean of the prediction errors for the period. The coefficient δ_2 can be interpreted as the mean percentage change of the prediction error between one period and the other. Hence, following this logic, if the value of δ_2 is positive and significant, the errors are positively correlated with the errors of the preceding period.

Given the inevitable presence of autocorrelation of the residuals of this regression, I applied **COCHARAN-ORCUTT** transformations to rectify this problem.

Table 2 indicates that analysts' forecasting errors are positively correlated with the errors from the previous period (δ_2). This means that the analysts, although they had committed forecasting errors in a determined period, continued to commit the same errors in the following period.

When analyzed based on subgroups, in the subgroup of positive errors although δ_2 is positive, it is not statistically significant. For the subgroup of negative errors, the persistence of the errors is significant. So, it can be inferred that in scale of forecasting errors, the street consensus can be persistently optimistic over different periods, but this persistence phenomenon is not found for the pessimistic consensuses.

Regarding the intercepts, δ_1, they were highly significant, indicating that analysts tend to be optimistic in their forecasts. The intercept for the negative error subgroup is much greater than that for the positive subgroup, reaffirming a trend in the errors and a trend in the optimistic consensus of analysts. These observations lend support to the hypothesis that the size and the trend of the consensus prediction errors are optimistic.

Table 2 - Results of the regressions, on the trend of the analysts' forecasting errors – consensus of December 1995 – 2002

Metric	δ_1	*t*-stat	sig	δ_2	*t*-stat	sig
All Erros						
Err Pred	-0.622	*-3.428*	[0.0006]	0.054	1.723	[0.0852]
Positive Erros						
Err Pred	0.687	*3.526*	[0.0004]	0.064	0.064	[0.4991]
Negative Erros						
Err Pred	-1.669	-5.921	[0.0000]	0.087	1.794	[0.0735]

7. ADJUSTMENT OF THE CONSENSUS FOR SELECTION BIAS

7.1. Selection bias

The studies carried out permit the conclusion that analysts' forecasts are optimistic in average terms. One of the explanations for this optimistic bias in analysts' consensus is known in the literature as selection bias.

The idea behind selection bias in this context is that each analyst is truthful about his or her expectation of a firm's performance. However, those who believe that the particular firm will perform poorly tend not to announce their estimates, and hence the overall street consensus will be skewed toward better expectations than if all analysts announced their true feelings.

The phenomenon of selection bias regarding the companies for which earnings forecasts are announced is a plausible explanation for part of the excess optimism found in analysts' consensus. Even though they are preparing their projections ex ante without any bias, analysts can still produce an overly optimistic consensus.

Some defenders of this explanation, such as HAYES & LEVINE (2000), suggest that this selection bias is clearly associated with the incentives analysts have to obtain stock trading commissions. If the perspectives on a firm are good, it is worthwhile to produce a report. But if the scenario is unfavorable, it is not worth the effort to produce a report on the firm.

In a very elucidating example, suppose a teacher wants to compute the level of success of students in passing his or her class by the grades on the final exam. However, on the day of the final, the five worst students in the class simply decide not to show up, judging it would be a waste of time. If the teacher determines the success rate only with the students who actually take the exam, the estimate will be biased in favor of passage. In other words, the average pass rate of the students who take the test will certainly be higher than if all the students had taken the test.

7.2. What are the effects of this selection bias?

To consider the implications of this type of selection bias, I sought to use a simple model. Suppose the forecast of analyst *j* for firm *i* at time *t* can be represented by the following behavior:

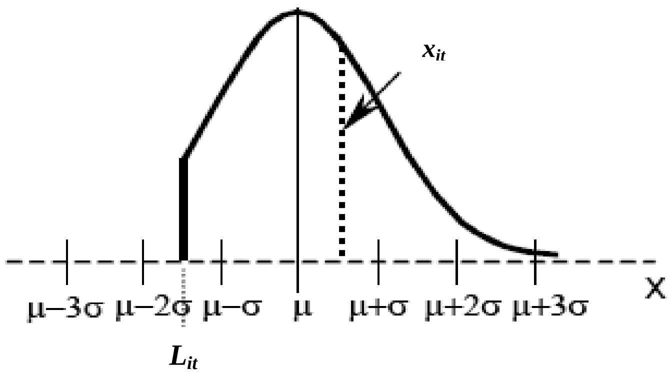
$$x_{it}^{j} = \mu_{it} + \varepsilon_{it}^{j} \qquad (1)$$

The term ε denotes the prediction error, which has behavior given by $N(0, \sigma_{it}^2)$. In this model, it is assumed that the information of the analysts follows a normal distribution. By hypothesis, it is considered that analysts do not disclose their predictions if these are below a lower limit L_{it}. In other words, this limit acts as a cutoff point. At values above this limit, the forecasts are disclosed, and at values below it they are omitted, supposing that analysts have a normal distribution of expectations regarding the firm.

So, the distribution of observed forecasts will be a truncated normal curve, as shown in **Figure 1**.

The value x_{it} corresponds to the mean of the verified observations. The value μ represents the expected mean if the curve had not been truncated. The difference between x_{it} and μ corresponds to the bias. Below I present the estimate of this bias in more formal terms.

Figure 1 – Truncated normal curve



Using the properties of a truncated distribution, it is possible to represent the expected value of x_{it} as follows:

$$E[x_{it}] = \mu_{it} + \sigma_{it}\left[\frac{\phi(\theta)}{1-\Phi(\theta)}\right] \quad (2)$$

where $\theta = \dfrac{L_{it} - \mu_{it}}{\sigma_{it}}$ and $\phi(\theta)$ e $\Phi(\theta)$ are the probability density function and the accumulated density function of the standard normal distribution, respectively.

The value x_{it} represents the mean of the observations that are available and μ_{it} is the effective mean of the population considering the part not observed.

Therefore, the value of $\sigma_{it}\left[\dfrac{\phi(\theta)}{1-\Phi(\theta)}\right]$ in reality represents a bias.

Transferring this discussion to the sphere of analysts, this bias is exactly what makes their consensus overly optimistic.

Assuming that the analysts' predictions are normally distributed, it is possible to estimate the bias.

If the point L_{it} is known, it is possible to calculate $\left[\dfrac{\phi(\theta)}{1-\Phi(\theta)}\right]$ and multiply it by the dispersion to estimate the selection bias.

There are various methods that can be used to estimate this bias, but the simplest is to use the maximum likelihood method (MLM).

The MLM of the mean μ_{it} can be written as:

$$\mu_{it}^{MMV} \in \arg\max \mu_{it} \left[cons\tan te - n\log\sigma_{it} - \sum_{j=1}^{n} \frac{\left(x_{it}^{j} - \mu_{it}\right)^2}{2\sigma_{it}^2} - \sum_{j=1}^{n} \log\left(1 - \Phi\left(\frac{L_{it} - \mu_{it}}{\sigma_{it}}\right)\right)\right] \quad (3)$$

In turn, the MLM of the dispersion σ_{it} can be represented as:

$$\sigma_{it}^{MMV} \in \arg\max \sigma_{it} \left[cons\tan te - n\log\sigma_{it} - \sum_{j=1}^{n} \frac{\left(x_{it}^{j} - \mu_{it}\right)^2}{2\sigma_{it}^2} - \sum_{j=1}^{n} \log\left(1 - \Phi\left(\frac{L_{it} - \mu_{it}}{\sigma_{it}}\right)\right)\right] \quad (4)$$

Considering that L_{it} is the minimum forecast reported by the analysts, the solution of the problem consists of an optimization process aiming to calculate simultaneously a mean and dispersion that maximize functions (3) and (4).

7.3. Is there selection bias among analysts of Brazilian companies?

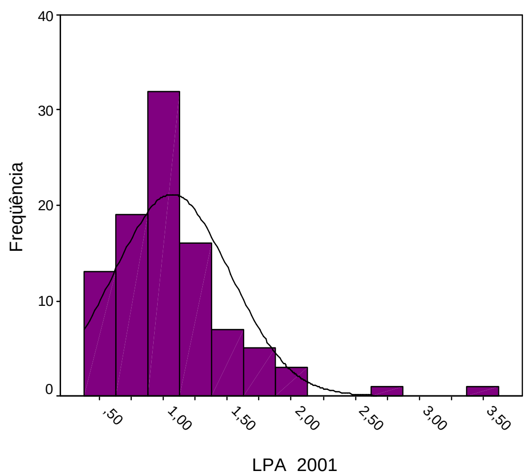
A question that deserves investigation is whether analysts of Brazilian companies select the companies for which they produce their reports. Good sense appears to say they do, if only from a perspective of rationalizing time. Their focus will tend to be only to analyze those companies that offer good prospects of generating future trades.

Thus, the assumption is that when an analyst imagines that the earnings will be below a minimum level, he or she will simply opt not to make any forecast. Since proving this phenomenon is no simple task, I preferred to present the distribution histograms of analysts' predictions for certain firms.

With this purpose, in **Figure 3** I show the histograms of analysts' forecasts for the EPS of four companies shown in **Table 3** below.

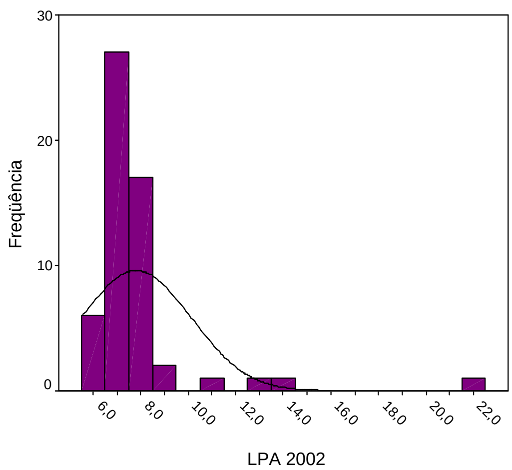
Figure 3 - Histograms of analysts' EPS estimates for four Brazilian companies

BRTP4 - Brasil Telecom Par PN



Mean of the Estimates: R$ 1.08 per lot
SD : 0.46
N = 97 estimates

UBBR4 – Unibanco PN

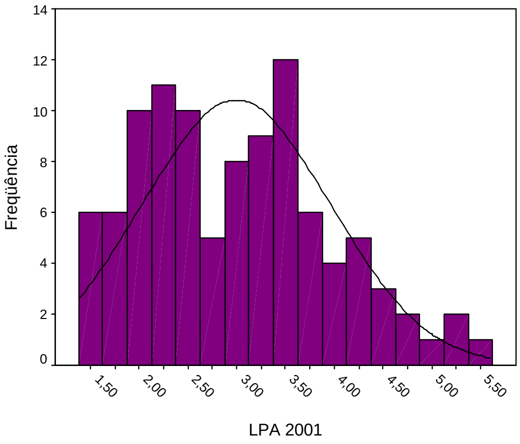


Mean: R$ 7.80 per lot
SD. : 2.32
N = 56 estimates

Real Earnings: R$ 0.75 per lot

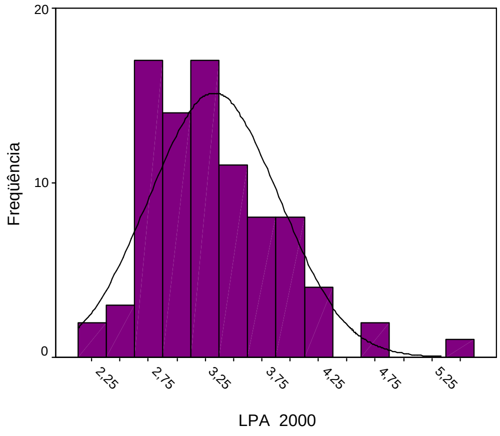
Real Earnings: R$ 7.17 per lot

PCAR4 - Pão de Açúcar PN



Mean: R$ 3.00 per lot
SD : 0,97
N = 101 estimates
Real Earnings: R$ 2.23 per lot

CMIG4 - Cemig PN



Mean: R$ 3.30 per lot
SD: 0.57
N = 87 estimates
Real Earnings: R$ 2.61 per lot

Note: The analysts' forecasts were accumulated during the respective years. The forecasts shown in the histograms are in R$ per lot of 1000 shares. The vertical axis is the total of the estimates per interval. The graphs also show the normal curve inferred from the observations, along with the descriptive statistics for each of the companies analyzed.

For each company the mean of the estimates and the actual result are shown. In all four examples, the analysts' consensus was optimistic in relation to the observed outcome.

Using the methodology discussed earlier, I determined the level of bias by an optimization process. I used **Excel Solver** to identify the mean and dispersion that would maximize equations (3) and (4). By applying the suggested procedure to calculate the bias, and deducting it from the consensus values, the adjusted consensus was obtained.

The values found were mainly pessimistic in relation to the observed EPS. The results are as follows:

Table 3: Results of Adjusting for Selection Bias for Four Brazilian Companies

Company	Consensus EPS	Bias	Adjusted Consensus EPS	Observed EPS
Brasil Telecom Participações	1.08	0.918	0.162	0.75
Unibanco	7.80	0.269	7.531	7.17
Pão de Açúcar	3.00	2.268	0.732	2.23
Cemig	3.30	1.422	1.878	2.61

Values in R$ for a lot of 1000 shares

The only exception was UNIBANCO, where the adjusted consensus continued being optimistic. In general lines, the existing optimistic bias, considering the adjusted consensuses, disappeared, in reality becoming a pessimistic bias. In terms of accuracy, the procedure was also not satisfactory, because the total of the prediction errors increased.

An important aspect of the analysis was the fact that, by hypothesis, it was assumed the previous forecasts were normally distributed, and perhaps this was not exactly so.

Another strong hypothesis was a homogenous cutoff point for presenting a forecast. There is no way to ensure that all the analysts have the same threshold below which they decide not to announce their predictions about a certain company.

Although the results were not exceptional in terms of accuracy and correcting for bias, I believe that a reflection on selection bias is opportune any time considering the consensus of analysts, be it earnings predictions or stock recommendations.

8. CONCLUSIONS AND FINAL CONSIDERATIONS

The analyses carried out in this article, considering the consensus forecasting errors of market analysts of Brazilian companies in the period from 1995 to 2002, indicate:

i. These analysts were optimistic on average. This is demonstrated by a significantly negative mean prediction error. Additionally, on average they performed poorly regarding accuracy and precision.

ii. The prediction errors in one period are correlated with those in the subsequent period. There was a clear persistence of analysts' consensus to err, with those who were persistently optimistic predominating over those who were persistently pessimistic.

iii. Considering, by hypothesis, that the distribution of observed forecasts by the analysts for a determined firm is a truncated normal curve, a simple optimization procedure allows calculating the selection bias and adjusting the consensus, eliminating the apparent optimism.

The results of this study bring some important implications that should be borne in mind when considering the behavior of analysts of Brazilian companies. It is particularly important when analyzing consensus information to understand that there can be an optimistic bias. The evaluation of the magnitude of this bias will depend on the specific characteristics of each firm, its size, conditions in which the consensus prediction was reached, number of estimates and their dispersion.

The forecasts of analysts of Brazilian companies should not be disregarded. They only need to be placed in context, allowing greater reliance on those that have conditions to be more effective. Understanding how these professionals operate, in average terms, is relevant both from an academic and practical standpoint. The observations reported here can open the way for future research, leading to a better understanding, in its precise acceptation, of the meaning of market expectations, or the street consensus, improving valuation models and optimizing the calculation of variables, such as the cost of capital.

REFERENCES

Abarbanell, J.; Lehavy, R. *Can stock recommendations predict earnings management and analysts' earnings forecast errors*? Journal of Accounting Research. Rochester, 41 (1): 1-31. 2003a.

Abarbanell, J.;Lehavy, R. *An explanation for why prior stock returns and analysts' earnings forecast revisions predict earnings management and forecast errors.* Working Paper 2003b, May. Available at: <http://webuser.bus.umich.edu/rlehavy/ALBiasExplanation05-23-03.pdf> Accessed on December 31, 2003.

Brown, L. *Analysts forecasting errors and their Implications for security analysis: An alternative perspective*. Financial Analysts Journal, pp. 40-46. 1996

Da Silva, Henrique R. *A capacidade previsionária no mercado acionário brasileiro – Um estudo focado nas previsões dos analistas de investimentos*. ENANPAD-Finanças. 1998

Das, S.; Levine, C.; Sivaramakrishnan, K. *Earnings predictability and bias in analysts' earnings forecasts*. The Accounting Review. Sarasota 73: pp. 277-294. 1998

Debondt, W.;Thaler, R. *Does the stock market overreact*? Journal of Finance. Chicago: 40: pp. 793-805. 1985.

Debondt, W.;Thaler, R. *Further evidence of investor overreaction and stock market seasonality*. Journal of Finance. Chicago: 42: pp. 557-581. 1987.

Franco, D. *Projeções de lucros: há evidências de exageros sistemáticos?* Revista da Bovespa, pp. 9-11. 2000.

Givoly, D. & Lakonishok, J. *Properties of analysts' forecast of earnings: A better surrogate for earnings expectations*. Journal of Accounting and Economics, Rochester. Pp. 85-107. 1984.

Guajarati, Damodar. Basic econometrics. New York: McGraw-Hill. 1996.

Gu, Z. & Wu, J. *Earnings skewness and analyst forecast bias*. Journal of Accounting and Economics. Rochester, 35 (1): pp. 5-29. 2003.

Hayes, R. & Levine, C. *An approach to adjusting analysts' consensus forecast for selection bias*. Contemporary Accounting Research. Toronto, 17 (1): pp. 61-83. 2000.

Kothari, S. P. *Capital market research in accounting*. Journal of Accounting and Economics. 31 (1): pp. 105-231.2001.

Lim, T. *Are analysts' forecasts optimistically biased*? Working Paper. Dartmouth University. 1998.

Lin, H.; Mcnichols, M. *Underwriting relationship and analysts' earnings forecast and investment recommendations*. Journal of Accounting and Economics. Rochester, 25: pp. 101-127. 1998.

Martins, G. A. Estatística Geral e Aplicada. São Paulo: Atlas.2002.

Richardson, S.; Teoh, S.; Wysocki, P. *Tracking analysts' forecasts over the annual earnings horizon: are analysts' forecasts optimistic or pessimistic*? Working Paper. University of Michigan. 1999.

Trueman, Brett. *Analysts' forecasts and herding behavior*. The Review of Financial Studies. Oxford, 7 (1): pp. 97-124. 1994.