Residual Income Valuation and B/P: factoring value creation and forecast errors on B3

Marcus E. Firmato
Sao Paulo School of Economics - FGV, Brazil
mefirmato@gmail.com

Abstract Investigating the performance of models based on fundamental signals that better reflect market expectations for future earnings and the influence of behavioral factors on the explanatory power of such determinants are a challenging quest in the literature. This paper examines the contribution of a value creation approach (RIV-EBO model) to explain stock returns relative to a simpler strategy based on book-to-market ratio (B/P) and possible behavioral biases impacting analysts' accuracy in an emerging market with lower informational efficiency such as the Brazilian capital market. Our results suggest a safety quality attribute in favor of the RIV-EBO model and once compared to the U.S. market, indicate the need for improvement in the quality of local analysts' earnings forecasts for enhancing its predictive power for stock returns.
Keywords: Fundamental Signals, Value Creation, Predictive power for stock returns, Analyst forecast errors.

1. Introduction

The relation between financial statement data or fundamental signals and firm's value has been the subject of academic studies, particularly the importance of using financial statement data in asset pricing and predicting returns, as well as attempts to reconcile the use of fundamental signals in empirical works with modern finance theory (Penman, 1992; Abarbanell and Bernard, 2000; Jorgensen et al., 2011; Nekrasov, 2016; Anesten et al., 2020). In this sense, the studies by Ohlson (1995) and Feltham and Ohlson (1995) played a relevant role in the development of capital market research by to some extent invigorating the residual income valuation (RIV) model.[1] This is due to a change in empirical research focus from the emphasis on explaining "stock price behavior" towards "predicting future earnings and book value growth", as emphasized by Bernard (1995, p. 734) and Abarbanell and Bushee (1997). Which is justified to the extent that the firm's value, even in a finite horizon scenario, can be measured by a composite function of earnings and book value predictions and the firm discount rate, assuming the consistency of these predictions to a clean surplus relation (CSR).[2] Another example of the use of fundamental signals is represented by the book-to-market ratio (B/P), an investment strategy that became notorious particularly from empirical studies of asset pricing such as Fama and French (1992), which reported a robust relation between B/P and the firm's market equity (ME) and the average stock returns in the U.S. market from 1963 to 1990.

In addition, the quality of analyst earnings forecasts plays a relevant role since it allows evaluating the reliability of fundamental values estimated from models that use analyst earnings forecasts in their estimates but particularly the existence of excessive optimism in analyst forecasts helps to explain the value/glamour effect[3] observed in La Porta et al. (1997), Dechow and Sloan (1997) and Piotroski and So (2012) and also documented in this study, consequently supporting the mispricing hypothesis.

---

[1] As Dechow et al. (1999) point out, the Ohlson (1995) model basically relies on three assumptions: (i) the stock price is equal to the present value of expected dividends; (ii) the application of the clean surplus relation assumption ($b_t = b_{t-1} + x_t - d_t$, where b_t= book value, x_t = net income, and d_t= dividends, all determined in period (t)): (a) allows future dividends to be expressed in terms of future net income and book value, and (b) combined with the discounted dividend model, implies that the stock price equals the sum of book value and the present value of future abnormal earnings; and (iii) abnormal earnings and a corrective variable representing "other information" satisfies an autoregressive model (information dynamics), meaning that lagged values of these two variables represent explanatory variables in the estimation of future abnormal earnings.

[2] Or that the change in book value over a given period is equal to net earnings minus dividends distributed during the period.

[3] Effect characterized by the superior performance of portfolios formed by firms with low prices relative to their fundamentals compared to portfolios formed by firms with high prices relative to their fundamentals.

Given the relevance of investigating the behavior of future return prediction models and the evidence of systematic errors in market expectations for future earnings associated with the value/glamour effect, this study relates to both a strand of literature (Frankel and Lee, 1998; Dechow et al., 1999; Piotroski, 2000; Dissanaike and Lim, 2010; Artikis, 2018) that empirically analyzes the explanatory power of specific factors or drivers, such as the value to price ratio *V/P* and book-to-market ratio *B/P*, as fundamental value proxies and (Lakonishok et al., 1994; La Porta, 1996; Piotroski and So, 2012) that empirically investigate the influence of behavioral factors on the explanatory power of such factors or drivers.

Specifically, the focus of this paper is twofold. First, we question whether the adoption of a more complete valuation approach, that incorporates projections of growth from invested capital and expectations of future earnings based on analyst forecasts in estimating firm's fundamental value (RIV-EBO model), can contribute to explaining returns and present significantly superior predictive power for stock returns relative to a simpler strategy based on *B/P* in the local market. Second, we investigate the quality of analyst earnings forecasts in the Brazilian market to (a) evaluate the reliability of RIV-EBO model fundamental values estimation based on analyst earnings forecasts and (b) assess whether analyst forecast errors and local market lower informational efficiency present a limitation for the RIV-EBO model to achieve significantly superior predictive power for stock returns relative to *B/P* in the local market.

In favor of choosing the RIV-EBO model are its solid theoretical framework, its relative parsimonious construction, and the fact that its residual income present value incorporates profitability-growing, safety, and payouts attributes. These attributes, in turn, not only signal the firm's performance but also drive market multiples, making the comparison between value to price ratio *V/P* and book-to-market ratio *B/P* relevant. Moreover, fundamental value drivers such as earnings and invested capital are regularly monitored by firm's managers and market analysts (Bernard, 1994; Lee, 2014).

Applying the discounted residual income model called Edwards-Bell-Ohlson (EBO) to estimate firm's fundamental value in the U.S. market, Frankel and Lee (1998) observed a significant relation between fundamental value (V) and contemporaneous stock price (P) and evidenced that the value to price ratio (V/P) is a reliable predictor of stock returns (effect not explained by B/P ratio, size (ME) and firm's market beta). In the European market Dissanaike and Lim (2010), when analyzing firms listed on the London Stock Exchange (LSE), observed that, except for cash flow to price (which had similar performance), sophisticated models such as Ohlson (1995) and a specific residual income model generally outperformed simple strategies commonly used in finance literature such as book-to-market ratio, cash flow to price, earnings to price and past long-term returns. Artikis (2018), in turn, using a residual income model to measure firm's intrinsic value, found in the German stock market that there is a persistent value to price ratio (V/P) effect in longer investment periods that is not explained by firm's systematic risk, size (ME) or B/P ratio.[4] Providing grounds for using fundamental signals in asset pricing in the Brazilian market, Ohlson and Lopes (2007) demonstrated that the firms' market value is a function of their expected earnings, and that it is ambiguous to affirm that expected free cash flow models necessarily reflect value creation. Additionally, Ferreira et al. (2008) compared the RIV (Ohlson, 1995), earnings growth valuation (AEG) and free cash flow models and observed that, despite the RIV model presented superior and equivalent explanatory power compared to the AEG along different periods of time, the free cash flow model presented the lowest explanatory power in all analyzed years. Examining 97 valuation reports of stocks public tender offers between 2000 and 2007, Almeida, Brito et al. (2012) observed no statistically significant differences between the values estimated by the RIV (Ohlson, 1995, 2005) and AEG models compared to discounted cash flow model (DCFM) estimates. Finally, Almeida, Lima et al. (2012) found that the RIV and AEG models estimated values cannot be considered statistically different, but that the RIV model is better adjusted to the stocks market value and that the AEG model shows higher standard deviation of estimated values compared to the RIV model.

In addition, evidence in the U.S. market suggests that analysts influence the informational efficiency of the capital market (Hong et al., 2000; Frankel et al., 2006). Besides, behavioral factors, particularly expectation errors, play an important

[4] In this study the author did not use market analysts' forecasts to estimate future return on equity (ROE), but rather a time series model based on historical ROEs.

role in explaining the superiority of returns of value stocks compared to glamour stocks. When evaluating the existence of systematic errors in analyst forecasts, La Porta (1996) showed that contrarian investment strategies obtain superior returns due to market's excessive optimism (pessimism) on higher (lower) expected earnings' growth rate portfolio. La Porta et al. (1997) also observed that a significant part of the difference in returns between value and glamour stocks is attributed to earnings surprises, which are systematically more positive for value stocks. In addition, Dechow and Sloan (1997) noted that, despite earnings growing at less than half the rate initially predicted by market analysts, stock prices appear initially to substantially reflect biased analyst forecasts for long-term earnings growth, and that investors' reliance on such predictions explains more than half of the subsequent higher returns observed when buying (selling) stocks of firms with low (high) prices relative to their fundamentals. Piotroski and So (2012), in turn, observed that the outperformance of value stocks over glamour stocks and *ex post* reviews to market expectations are concentrated (do not exist) in firms with *ex ante* biased (unbiased) market expectations.

There is also evidence, as in the U.S. market, that on average analysts are optimistic in the Brazilian market and that forecasting errors of a given period are correlated with those of the subsequent period (Martinez, 2007a). Beiruth et al. (2014) also suggest the existence of over-optimism in analyst forecasts, as one possible explanation for the greater similarity observed between the value of the stock and the AEG model estimation (Ohlson and Juettner-Nauroth, 2005) compared to the target price estimated by market analysts. The value/glamour effect has also been evidenced in the Brazilian market in different periods, with distinct results for risk assessment (Braga and Leal, 2000; Rostagno et al., 2006; Santos and Montezano, 2011).

By comparing the predictive power for stock returns of alternatives strategies, based on both a value creation approach and *B/P*, and at the same time analyzing the analyst forecast errors and the value/glamour effect from a RIV-EBO model perspective, this study fills a gap in the existing literature. For this purpose, we apply the RIV-EBO model presented in Frankel and Lee (1998) to 136 firms in the Brazilian stock market.

Furthermore, the application of the RIV-EBO model presented in Frankel and Lee (1998) in the Brazilian market allows for a brief but relevant comparison of the performance of the model between markets with different forms of informational efficiency, contributing to broaden the discussion of the local performance of a residual-income-model relative to a *B/P* trading strategy in the context of foreign literature (Dissanaike and Lim, 2010; Artikis, 2018) and to the understanding of possible behavioral biases impacting analysts' accuracy in an emerging market with lower informational efficiency such as the Brazilian capital market, as suggested by Lopes and Galdi (2006) and Gonçalves et al. (2022). This is particularly important to assess the limitations the application of the RIV-EBO model may have under these specific circumstances and the empirical suitability of the RIV-EBO model to the Brazilian market relative to a book-to-market ratio trading strategy. To our knowledge this is the first study to compare the explanatory power of value to price ratio *V/P* using the RIV-EBO model based on analyst earnings forecasts and the book-to-market ratio *B/P* in the local stock market between 1998 and 2018.

To examine the relative contribution of each strategy in predicting returns we construct uni-dimensional portfolios based on *V/P*, *B/P* and size (*ME*) and also use a multiple regression approach based on percentile and scaled decile ranks, fixed/random effects models and first differences. We also evaluate the quality of analyst earnings forecasts by analyzing the relationship of certain firm characteristics, book-to-market ratio (*B/P*), market value of equity (*ME*), cost of equity (*Ke*) and an analyst optimism measure derived from the RIV-EBO model (*OP*), with the analyst forecast errors. In other words, we assess to what extent analyst over-optimism or forecast errors can be explained by these characteristics, given their use by market analysts in valuing firms.

We observe that in general, the RIV-EBO model is well-suited to the Brazilian market. Using percentile and decile rank multiple regressions, we find that the RIV-EBO and B/P models show similar predictive power for stock returns in the short term, but when extending the investment horizon to 36 months there is a gain in RIV-EBO model and an apparent loss of B/P explanatory power for returns. Using first differences, the RIV-EBO model shows a more robust and stable contribution in predicting returns over 12, 24 and 36 months compared to the B/P strategy. Both models demonstrate to be robust predictors and show similar returns at the end of all investment periods in the uni-dimensional portfolio analysis, with the RIV-EBO model, however, less associated with higher differences in market risk. Overall, the results suggest that a more sophisticated value creation approach contributes to explaining returns relative to a simpler trading strategy in the local market.

We also find that local analysts' tendency to place excessive confidence in glamour firms and the higher forecast errors compared to the U.S. market help explain the RIV-EBO model not having shown significant superiority of returns in High-ratio portfolio over B/P. Overall, the results signal the mispricing hypothesis as an explanation for part of the observed value/glamour effect, suggesting that the lower information efficiency of the Brazilian market makes financial signals not necessarily have the same relevance as observed in the U.S. market in Frankel and Lee (1998) study. The findings suggest that the search for enhancing the predictive power for returns of the RIV-EBO model relative to B/P requires overall the improvement in the quality of analyst earnings forecasts in the local market and the adoption of additional predicted forecast error trading strategies potentially exploiting B/P and OP as explanatory variables.

The remainder of this article is divided into the following sections: (a) Section 2 describes the methodology, presenting the theoretical foundation of the RIV-EBO model and describing the data sample, portfolio formation and statistical methods used, (b) Section 3 examines and discuss the results of the tests conducted; and (d) Section 4 concludes.

2. Methodology

2.1 Model and estimation

The RIV-EBO model, as demonstrated in Lee (1996), has its theoretical development associated with the economic value added (EVA) model, in that both use the residual income concept linked to a previously invested capital, meaning that a firm, by deriving its value from existing assets and generating value in future activities, creates value for its shareholders by increasing its earnings using the same capital amount, reducing the employed capital while maintaining the earnings level, or reducing the cost of equity (Lee, 1996), as summarized below:

$$\mathrm{EVA}_t = earnings_t - r * capital_{t-1} \tag{1}$$

in which EVA_t = economic value added in period (t), $earnings_t$ = earnings or net income reported in period (t), r = cost of capital, $capital_{t-1}$ = capital employed at the beginning of period (t).

Rearranging equation 1 we have the EVA model as the difference between the realized return and the cost of equity from the capital employed in the firm:

$$\begin{aligned}\mathrm{EVA}_t &= NI_t - r_e * B_{t-1}\\ &= \left(\frac{NI_t}{B_{t-1}} - r_e\right) * B_{t-1}\\ &= (ROE_t - r_e) * B_{t-1}\end{aligned} \tag{2}$$

in which NI_t = net income at the end of period (t), r_e = cost of equity, B_{t-1}= book value at the beginning of period (t)[5] and ROE_t = return on equity at the end of period (t).

It can be inferred that the created wealth depends on the amount of employed capital (c_t) and that the firm creates value for investors if the return on equity (ROE) exceeds the cost of equity (r_e). By extending this concept to several periods, the firm's value can be estimated by summing up the invested capital and future activities represented by the present value of future EVAs (Lee, 1996):

[5] To make it easier to interpret the formulas we use the reference beginning of period (or year) (t) which in practice refers to the amount reported at the end of December in period (year) (t-1).

$$V_t = c_t + EVA_{pv} \tag{3}$$

in which V_t = firm's value in period (t), c_t= capital employed in period (t), EVA_{pv} = present value of future EVAs.

In addition, provided that future earnings and future book values are estimated consistently with the clean surplus relation (CSR) assumption, it is possible to demonstrate (Ohlson, 1991; Feltham and Ohlson, 1995) that a DDM estimation of a firm's fundamental value (equation 4) can be expressed as a function of reported book value and discounted residual income expectations (equation 5) (Bernard, 1994; Frankel and Lee, 1998) as follows:

$$V_t^* \equiv \sum_{i=1}^{\infty} \frac{\mathrm{E}_t(d_{t+i})}{(1+r_\mathrm{e})^i} \tag{4}$$

in which V_t^*= stock's fundamental value at time (t), $\mathrm{E}_t(d_{t+i})$ = expected future dividends for period (t+i) conditional on the information available at time (t), r_e = cost of equity based on information set at time (t).

$$\begin{aligned} V_t^* &= B_t + \sum_{i=1}^{\infty} EVA_{t+1} \\ &= B_t + \sum_{i=1}^{\infty} \frac{\mathrm{E}_t[(ROE_{t+i} - r_\mathrm{e})B_{t+i-1}]}{(1+r_\mathrm{e})^i} \\ &= B_t + \frac{\left(ROE_{(t+1)} - r_e\right)}{1+r_e} B_t + \frac{\left(ROE_{(t+2)} - r_e\right)}{(1+r_e)^2} B_{t+1} + \frac{\left(ROE_{(t+3)} - r_e\right)}{(1+r_e)^3} B_{t+2} + \ldots \end{aligned} \tag{5}$$

in which V_t^* ou P_t^* = stock's fundamental value at time (t), $B_{(t)}$= book value at time (t), $\mathrm{E}_t[\ ]$ = expectation based on information available at time (t), B_{t+i-1} = book value in period (t+i-1), ROE_{t+i} = return on equity for period (t+i), r_e = cost of equity.

The model, although based on a DDM, provides a clearer and more intuitive framework for analyzing the relation between firm's value and accounting numbers (Frankel and Lee, 1998), enabling the distinction between the wealth effectively created (measured by earnings) and the often-discretionary wealth distribution (through dividends) (Penman, 1992).[6] Besides, the model incorporates not only the invested capital represented by B_{t+i-1}, variable commonly used to measure future stock performance through the book-to-market ratio (Rosenberg et al., 1985; Fama and French, 1992; Piotroski, 2000), but also specific attributes of the residual income present value, such as profitability and growth (as future profitability and growth), safety (as the firm's risk associated with the discount rate used to estimate its residual income present value) and payouts (as capital distributed to shareholders), which are recognized in the literature (Lee, 2014).[7]

In turn, a superior performance of these attributes, as they translate into a higher residual income present value, signals an asset quality component. Asness et al. (2013), defining "quality" as stocks that investors should be willing to pay a higher price for when characterized (*ceteris paribus*) as safe, profitable, growing, and well-managed, reaffirm the significance of the accounting information in valuing firms.

[6] It is assumed that the eventual violation of the clean surplus on a per share basis is addressed by estimating the book value forecast per share as the sum of prior book value per share and earnings per share minus dividends per share (see Appendix A.1 equations 14 to 18), or that in expectation dirty surplus items are zero (Bernard, 1995; Ohlson, 2000; Penman, 2005).

[7] For supporting that firms with higher profitability (either current or future) obtain higher returns see Piotroski (2000), Mohanram (2005), Galdi (2008) and Villaschi et al. (2011), that safer firms earn higher returns see Black (1972), Ang et al. (2006), George and Hwang (2010) and Frazzini and Pedersen (2014), and that there is an overall positive relation between a firm's payout distribution to shareholders and future returns see Spiess and Affleck-Graves (1995), Affleck-Graves and Miller (2003) and Billett et al. (2006).

In conclusion, the model is capable to highlight relevant accounting metrics indicating the firm's performance and also driving multiples used by the market in valuing stocks (Lee, 2014). Following Frankel and Lee (1998), according to equation 5 we use the following three models to estimate the fundamental value based on analyst forecasts (Vf):

$$\hat{V}_t^1 = B_{t-1} + \frac{(FROE_t - r_e)}{(1 + r_e)} * B_{t-1} + \frac{(FROE_t - r_e)}{(1 + r_e)r_e} * B_{t-1}, \tag{6}$$

$$\hat{V}_t^2 = B_{t-1} + \frac{(FROE_t - r_e)}{(1 + r_e)} * B_{t-1} + \frac{(FROE_{t+1} - r_e)}{(1 + r_e)r_e} * B_t, \tag{7}$$

$$\hat{V}_t^3 = B_{t-1} + \frac{(FROE_t - r_e)}{(1 + r_e)} * B_{t-1} + \frac{(FROE_{t+1} - r_e)}{(1 + r_e)^2} * B_t + \frac{(FROE_{t+2} - r_e)}{(1 + r_e)^2 r_e} * B_{t+1}. \tag{8}$$

The first two models represent a two-period RIV-EBO model, with the first one assuming a one-year ahead ROE forecast earned in perpetuity, while the second model assumes a two-year ahead ROE forecast earned in perpetuity. The third model, on the other hand, represents a three-period model assuming a three-year ahead ROE forecast in perpetuity. Modeling with 3 alternatives is particularly relevant to evaluate the behavior of each model, considering not only the number of periods expansion, due to the possibility of losing predictive power over time, but also the ROE forecast used in the terminal value. It also allows us to choose the RIV-EBO model with greater predictive power to compare against *B/P*, to evaluate how analyst forecasts behave as the forecast period increases and to compare our results in the Brazilian market with Frankel and Lee's (1998) study in the U.S. market. The stock's fundamental values obtained from the 3 models (equations 6 to 8) based on analyst earnings forecasts are denoted as (*V1f*), (*V2f*) and (*V3f*), respectively (for details on the estimation of the future book value and the firms' future ROEs (FROEs) see Appendix A.1).

Given the relevance of estimating future ROEs in the RIV-EBO model and the need to compare alternative methods to apply the model, in addition to the approach based on analyst earnings forecasts, we use the reported or historical earnings as an alternative for estimating future ROEs. Moreover, like the analyst earnings forecasts, the historical earnings approach is also based on ex ante information. In addition to the analysis carried out based on market analyst earnings forecasts, the 3 (three) models are also applied to measure firms' fundamental value based on their historical earnings (Vh). For this purpose, all FROEs obtained from analyst earnings forecasts in equations 6 to 8, as well as in equations 18 and 20, for estimating future book values, are replaced by the return on equity (ROE_t) estimated as follows:

$$ROE_t = \frac{NI_{t-1}}{B_{t-1}} \tag{9}$$

in which NI_{t-1}= net income at the end of December in year (t-1), B_{t-1}= book value at the end of December in year (t-1).

The analysis using ROE_t aims to evaluate how the RIV-EBO model behaves in the local market using historical earnings relative to analyst earnings forecasts, to choose, as indicated, the RIV-EBO model with the greatest predictive power to confront *B/P*, to evaluate how the historical data used behaves as the period in equations 6 to 8 increases, and to assist estimating analysts optimist. The fundamental values using historical earnings (ROE_t) are denoted as ($V1h$), ($V2h$) and ($V3h$), respectively.

Finally, to estimate the discount rate of each observation following a capital asset pricing model we use the implied equity risk premium in Brazil during the sample period, the industry-based betas of the sample firms and the risk-free rate also observed during the sample period as follows:

$$r_e = RFR + (\beta * ERP) \tag{10}$$

in which r_e = firm's cost of equity, RFR = risk-free rate, β = firm's industry-based beta, $ERP =$ implied equity risk premium.

To investigate the quality of analyst earnings forecasts in the second part of our analysis, also following Frankel and Lee (1998), we estimate the analyst forecast errors (FErr) and the analyst optimism (OP) measure derived from the RIV-EBO model as follows:

$$FErr_{t+i} = FROE_{t+i} - ROE_{t+i} \tag{11}$$

in which $FErr_{t+i}$= analyst forecast error relative to year (t+i), $FROE_{t+i}$ = forecasted ROE based on analyst EPS forecasts for the end of year (t+i) according to equations 17, 19 and 21, ROE_{t+i} = actual reported ROE for the year (t+i).

$$OP = (V_f - V_h)/|V_h| \tag{12}$$

in which OP = analyst optimism measure derived using the RIV-EBO model, V_f = fundamental value based on analyst EPS forecasts estimated from equation 7 (two-period model), or *V2f*, V_h = fundamental value based on historical earnings (ROE_t) estimated from equations 6 and 9 (one-period model), or $V1h$.

Analyst forecast errors (FErr) represent the difference between the forecast of future ROE for year (t+i) estimated from the clean surplus relation (CSR) assumption and analyst earnings per share (EPS) forecasts, and the actual ROE reported for year (t+i).[8] In this sense, the analyst over-optimism (over-pessimism) relative to the actual reported ROE results in more positive (negative) values of FErr. The analyst optimism (OP), in turn, represents the difference between the fundamental values estimated from the RIV-EBO model based on analyst forecasts and on historical ROE.[9]

Significant divergences between the fundamental values based on analyst forecasts and on historical earnings though can signal an underweighting of historical data used as benchmarks in earnings forecasts. In this event, OP will be positively related to analyst optimism, causing analysts to expect that firms with higher (lower) OP will achieve higher (lower) growth in future ROEs. In turn, analyst over-optimism in forecasting future ROEs relative to the actual reported ROEs will result in higher forecast errors (Ferr) (Frankel and Lee, 1998).

In addition to the OP measure, analyst forecast errors are also evaluated based on their relationship with the book-to-market ratio (*B/P*), the market value of equity (*ME*) and the cost of equity (*Ke*), justified in that B/P allows for explaining the value/glamour effect, observed in this study, and that market analysts consider these firm characteristics in valuing firms, making it relevant to evaluate their influence on analyst forecast errors. Besides, in addition to the pursuit of a greater stability in earnings per share, size is among the variables documented by Holloway et al. (2013) that influence managers to hold a stock in their portfolios in the local market.

2.2 Data and sample description

The data used represent ordinary shares of all firms listed on B3, excluding financial firms, with data available between 1997 and 2018 and are obtained from Eikon (Refinitiv), I/B/E/S (Refinitiv) and the firms' financial statements database.

In order to estimate firms' fundamental values based on both analyst earnings forecasts and historical earnings data, fundamental ratios (Vf/P) and book-to-market ratios (B/P) firms needed accounting data available for common shareholder equity, net income available to common shareholders, cash dividends paid to common stockholders and total assets in year (t-

[8] Consistent with Frankel and Lee (1998), the use of ROE instead of EPS (used in Martinez (2007a, 2007b)) in the estimation of the forecast errors aims to mitigate eventual problems of stock split timing when comparing EPS forecasts and the actual reported EPS values.

[9] The choice of *V2f* and *V1h* in equation 12 is justified by (i) the superiority obtained from the *V2f* model relative to *V3f* revealed in Section 3.3.1, and by (ii) the superiority of the R^2 of *V1h* over *V2h* and *V3h* in the Spearman correlation tests reported in Table 16.

1), as well as total common shares outstanding both in year (t-1) and in June of year (t) (see Appendix A.2.1 for a detailed description of the data used). The year (t) represents the year in which (i) the estimates of fundamental values based on analyst earnings forecasts (Vf), fundamental values based on historical earnings (Vh), fundamental ratios (Vf/P), and book-to-market ratios (B/P) are carried out, specifically at the end of June, and (ii) returns are calculated, specifically from the beginning of July, for each observation that makes up the sample. Due to restrictions adopted in portfolios formation, the end of the accounting period coincides with the fiscal year-end in December of year (t-1) for all firms in the sample. In addition, the firms needed available accounting data on analysts' estimated earnings per share (EPS) for the end of years (t) (FY1), (t+1) (FY2) and (t+2) (FY3) to estimate their FROEs for the end of years (t), (t+1) and (t+2).

The firm's cost of equity estimation (Ke) is based on (i) the Equity Risk Premium (ERP) for the month of June of each year (t) estimated by CEQEF[10], (ii) the risk-free rate used in the ERP estimation, also for the month of June of each year (t) and (iii) each firm industry-based beta estimated by Nefin.[11] The estimated ERP represents an implied equity risk premium based on current Brazilian market prices and firm's fundamentals (return on equity and payout ratio) as inputs to the Gordon formula, incorporating therefore market expectations.[12] As a consequence, the significant variation of the ERP throughout the sample period observed in Appendix A.2.2 reflects the sensitivity and responsiveness of the ERP to current market conditions and risks. [13]

The industry-based betas estimated by Nefin comprise 6 economy industries (basic products, manufacturing, construction, energy, consumer and other) and represent historical CAPM Betas based on data between January 2001 and June 2022. To estimate the historical CAPM betas of each industry for the mentioned period, Nefin ran a regression of monthly industry excess returns on the monthly market risk factor, representing the regression coefficient, therefore, the beta for the respective industry.

The use of industry-based betas estimated by Nefin in the mentioned period is justified by the unavailability of local market data for estimating industry-based betas with ex ante information throughout the entire sample period (1998 to 2018). This procedure does not suggest to influence on the results as (i) they represent historical proxies for systematic risk throughout the entire period, (ii) preliminary tests using a common beta equal to 1 did not show significant differences in the results, (iii) the observed variation of the betas across the industries and their sample distribution in Appendix B.1 Table 11 compared to the observed variation of the equity risk premium (ERP) in Appendix A.2.2 suggests that the latter seems more relevant due to the large differences in ERP evidenced over the period in an emerging market like Brazil, and (iv) as indicated in the Model and estimation Section (2.1), despite the discount rate relevance, the strength of the model lies in incorporating growth projection from invested capital and future earnings forecasts in the firm's intrinsic value estimation.[14]

For each month in the sample period, the last closing price of each stock (price close) published by Eikon (Refinitiv) and the closing value of IBOV published by Economatica are collected. (see Appendix A.2.3 for detailed description of the closing price).

Following Frankel and Lee (1998), to avoid unreasonable ratios, ratios that do not allow for an economic interpretation and outliers, some restrictions and measures are applied to accounting data and analyst forecasts to form the

---

[10] Center of Quantitative Studies in Economics and Finance at Sao Paulo School of Economics - FGV. https://ceqef.fgv.br/bancos-de-dados.
[11] Center for Research in Financial Economics of the University of São Paulo. https://nefin.com.br.
[12] For detail on ERP estimation see Sanvicente (2017). To calculate the ERP the author uses the estimated discount rate of future dividends of B3 stocks (excluding financial institutions) based on the Gordon model by Gordon (1962), and the 10-year U.S. Treasury bond current yield to maturity, as the risk-free rate, obtained from Economatica website.
[13] See Sanvicente (2015) for justifying not adding up an additional and separate country risk factor in equation 10 at the risk of double counting and emphasizing that the Brazilian capital market is already developed for stock current prices to incorporate relevant risks information, including the country risk premium. See also Sanvicente and Carvalho (2020) for justifying the use of an international security as a risk-free asset, assuming that the Brazilian market is sufficiently integrated into the global capital market, as evidenced in Sanvicente et al. (2017). It is also demonstrated that, in addition to the country risk premium, changes in local interest rates (CDI rate), the U.S. market liquidity premium and the level of the S&P500 are determinants of changes in the ERP in Brazil.
[14] In addition, both Frankel and Lee (1998), who used constant interest rates (11%, 12%, and 13%), a single-factor and a three-factor industry-specific discount rates (derived by Fama and French, 1997), and Abarbanell and Bernard (2000), who allowed for firm-specific variations in discount rates, found that these variations had little effect on their results.

sample. Firms with negative book value and analysts' estimated EPS (*FY1*, *FY2*, *FY3*), as well as those with ROEs (return on equity) and k (dividend payout ratio) greater than 1, are excluded from the sample. For firms that reported negative net income in year (t-1), which represented 108 observations or 12.0% of the sample, a proxy (as normal earnings) equivalent to the average percentage of ROA (return on assets) for the entire sample (which is 7%) multiplied by the total assets reported by the respective firm in year (t-1) is used. As a result, both ROEs/FROEs and ks are restricted to the range of 0 to 100%.[15]

To investigate the quality of analyst earnings forecasts in the second part of our analysis, due to the existence of firms with negative book value, ROEs higher than 100%, lacking data related to book value and net income in the year(s) (t+1), (t+2), and/or (t+3), 13 (thirteen) observations from the original sample were excluded, totaling 889 observations. These data are analyzed for the purpose of determining the reported ROEs in the same periods and estimating $FErr_{t+i}$ according to equation 11.

2.3 Portfolio formation and statistical methods

After estimating the fundamental ratio (*Vf/P*) and book-to-market ratio (*B/P*) for each observation at the end of June of year (t), all observations in the sample, regardless of year (t) to which each observation refers, are assigned in descending order according to fundamental ratio (*Vf/P*), book-to-market ratio (*B/P*) and size (*ME*), and sorted into tertiles, that is, portfolios 1T (*Low ratio*), 2T and 3T (*High ratio*). The same procedure is used for investigating the quality of analyst earnings forecasts in the second part of our analysis on the basis of (i) analyst optimism (*OP*), (ii) book-to-market ratio (*B/P*), (iii) size (*ME*), and (iv) cost of equity (*Ke*).

The formation of each portfolio from all observations of the sample regardless of year (t) to which each observation refers (pooled data) is justified by the general limitation of the number of observations in each year (t) of the period sample, which made it unfeasible to form annual cross-sectional portfolios without significant observations loss. Furthermore, the formation of tertiles is also justified by the data unavailability, which made it unfeasible to segregate the sample into a greater number of portfolios and obtain a greater difference in sampling between extreme portfolios.

Following Fama and French (1992) and Frankel and Lee (1998), a conservative period of 6 months is adopted between the end of the fiscal year in December of year (t-1) and the ratios estimation (*Vf/P* and *B/P*) at the end of June of year (t). This is being done due to the disclosure of firms' results occurring in different months and the collection of analyst earnings forecasts through I/B/E/S occurring until the end of May of year (t). In turn, this allows the observations to be collected in a timely manner and be part of the sample at the end of June of each year (t) and ensures that analyst earnings forecasts correctly correspond to the fiscal year, in the case of the sample, at the end of December of year (t-1). In practical terms, the 6-month period aims to ensure that accounting data and analyst forecasts are known before calculating stock returns. Also consistent with Fama and French (1992) and Frankel and Lee (1998), we use the stock price (P) as of the end of December in year (t-1) in order to maintain time-related consistency between the variables that make up the Vf/P and B/P ratios. As to the firms' market value of equity calculation for portfolio formation based on size (*ME*), we use the stock price (P) and the number of shares outstanding as of the end of June in year (t).

After forming portfolios according to *Vf/P*, *B/P* and *ME*, the equally weighted average return is calculated for the observations (pooled data) that make up each portfolio. The cumulative return for each observation is estimated from the beginning of July in year (t), in which the observation is collected to compose the sample, for years (t+1), (t+2) and (t+3), thus determining the cumulative return for 12, 24 and 36 months of each observation from:

[15] After applying the proxy to the sample, due to the mentioned restrictions and measures, 33 observations were excluded due to negative analysts' estimated EPS (*FY1*, *FY2*, *FY3*), 121 were excluded for having a dividend payout ratio greater than 1, and 9 observations were excluded for having a return on equity greater than 1, resulting in a 902-observation sample.

$$R_{t+i} = \frac{P_{t+i}}{P_t} - 1 \tag{13}$$

in which R_{t+i} = cumulative return in year (t+i), P_{t+i}= closing price at the end of June in year (t+i), I = number of months (12, 24 and 36 months), P_t= closing price at the end of June in year (t).

For investigating the quality of analyst earnings forecasts in the second part of our analysis, once portfolios are formed based on *OP*, *B/P*, *ME*, and *Ke*, the analyst forecast errors (FErr) and the return of each portfolio are estimated based on an equally weighted average of all (pooled) observations that make up each portfolio. The analyst forecast errors (FErr) for each observation are estimated for years (t+1), (t+2) and (t+3) according to equation 11.

The estimation of the cumulative return does not account for tax and transaction costs and liquidity risk, and there is no rebalancing in constructing the portfolios, as they are formed from a pool of observations (pooled data) regardless of year (t) to which each observation refers, using a buy-and-hold strategy. Additionally, we estimate the 12, 24 and 36-month returns for all observations in the sample, allowing for the evaluation of portfolio performance over these investment horizons to be based on the same number of observations. As measures of systematic risk and risk adjusted return, post-ranking market betas and Sharpe ratio are estimated as described in Appendix A.3.1.

Finally, to assess the statistical significance of the differences in the results of the T3 (High ratio) and T1 (Low ratio) portfolios (T3-T1) both in the uni-dimensional and bi-dimensional analyses the Monte Carlo technique is used for portfolio simulation as detailed in Appendix A.3.2. In the regression analysis, we initially use ordinary least squares and pooled data (pooled OLS) for estimating the coefficients of multiple and simple regressions with independent variables expressed in terms of percentile and scaled decile ranks. To consider possible effects of heterogeneity, we apply fixed and random effects models to cross-sectional units or firms, and to also evaluate potential within cross-section serial correlation, we apply first differences and cross-section clustering.

3. Results and discussion

We now show and discuss the results of the comparison between the predictive power for stock returns of the *Vf/P* and *B/P* strategies and analyst forecast errors analysis. We introduce with the descriptive statistics of the sample to evidence the consistency of the overall sample parameters and the Spearman correlation test for all variations of the RIV-EBO model (for information related to the data distribution by industry and year see Appendix B.1). We then analyze the results on portfolio returns based on *V2f/P*, *B/P* and *ME* and of the regressions used to evaluate the predictive power of the models based on coefficients interpreted as the difference in returns between top and bottom ranked percentile firms and as returns obtained from a long and short investment portfolio, and also while controlling for heterogeneity and serial correlation. Finally, we examine the results of portfolios formed based on *OP*, *B/P*, *ME* and *Ke* and of regressions used to evaluate the explanatory power of these firm characteristics, expressed in terms of their percentile ranks, for forecast errors.

3.1 Descriptive statistics

Table 1 reports a book value per share mean of 46.58 Brazilian reais and a market equity (ME) mean of 9113.99 (nine billion one hundred thirteen million). These values are significant due to the RIV-EBO model used, which requires that the sample firms be covered by market analysts, and the sample period being relatively long (1998-2018), compared to the restrictive nature of samples observed in local stock market literature. Additionally, the significant difference observed between *min.* and *max.* values for book value per share is influenced by construction (real estate) and energy industries which usually have considerable amounts of fixed assets. We observe an average ROE of 17.07%[16], a higher percentage compared to the average ROE in the

[16] The significant value observed for ROE in the local market is not due to the application of the proxy applied to observations with negative net profits (average % of ROA x total assets), since the average ROE of observations with positive net income is 16.19%.

U.S. market observed by Frankel and Lee (1998) during the period between 1975 and 1993 (13%) and disclosed by Damodaran between 2016 and 2021 (13.45%).[17] The average ROE obtained in the sample is associated not only with the coverage of firms by market analysts but also with the prevailing equity risk premium over the sample period in the Brazilian market presented in Appendix A.2.2 and as a consequence with the average cost of equity (*Ke*) of 13.01% (ranging from 7.64% to 25.76%) of the sample. It is evidenced, as stated by the EVA model (equation 2), that value creation depends on obtaining ROEs that exceed the applied cost of equity.

Table 1: Descriptive Statistics.

Descriptive Statistics						
	Avg.	*Min.*	*Max.*	*Median*	*Stdv.*	*Total*
B(t-1)	46.58	0.10	4827.72	7.35	276.06	-
ROE(t-1)	0.1707	0.0015	0.9245	0.1446	0.1247	-
k	0.3188	0.0000	0.9998	0.2753	0.2599	-
B/P	1.12	0.01	24.51	0.75	1.42	-
ROA	0.0700	0.0002	0.3619	0.0591	0.0511	-
ME	9113.99	86.94	321204.32	2375.54	28279.76	-
Ke	0.1301	0.0764	0.2576	0.1273	0.0281	-
Beta Nefin	1.10	0.72	1.59	1.09	0.21	-
Obs.	-	-	-	-	-	902
N. Firms	-	-	-	-	-	136

Table values denoted as *Avg.* represent equally weighted average of all sample observations for all years (t) between 1998 and 2018. B(t-1) *Avg*. represents the average book value per share of the firms in the sample at the end of December in year (t-1). *ROE Avg* is the average return on equity of the sample, calculated as net income divided by book value, both at the end of December in year (t-1). *k Avg* is the average dividend payout ratio of the firms in the sample, calculated as dividends paid to ordinary shareholders divided by net income, both at the end of December in year (t-1). As for ROE and *k*, for firms that reported negative net income in December of year (t-1), a proxy was used for net income, equivalent to the percentage average of ROA (return on assets) of the entire sample (equivalent to 7%) multiplied by the total assets reported by the respective firm at the end of December in year (t-1). *B/P Avg* represents the average book-to-market ratio of the sample at the end of December of year (t-1), where B represents book value per share and P represents the price of the firm's share, both at the end of December in year (t-1). *ROA Avg* is the average return on assets of the sample, calculated as net income divided by total assets, both at the end of December in year (t-1). *ME Avg* is the average market value of equity of the firms in the sample in millions of Brazilian reais, calculated as shares outstanding multiplied by the share price, both determined at the end of June in year (t). *Ke Avg* is the average cost of equity of the firms in the sample. *Beta Nefin Avg* is the average of the industry-based betas of the sample firms estimated by Nefin. *Min.* and *Max.* represent the minimum and maximum values observed for each parameter in the sample. *Median* and *Stdv.* represent the median and the standard deviation values for each parameter in the sample. N. Firms is the total number of firms that make up the sample. Obs. is the total number of observations in the sample.

Taking also into account a median value of 14.46%, the observed average ROE seems well adjusted to the RIV-EBO model for conducting empirical tests. When applying industry-based betas to each firm, an average of 1.10 is observed for the sample. This value is close to the equally weighted average of 1.09 obtained from the 6 industry-based betas indicated in Table 11 (Appendix B.1) and the simple average beta for emerging markets between 2016 and 2021 equivalent to 1.11 (considering the entire market except to financials) disclosed by Damodaran.[18] As to the values of 31.88% for dividend payout (k) and 7% for ROA, they are close to the averages observed by Frankel and Lee (1998) in the U.S. market, which are 27% and 6%, respectively. The average of 1.12 obtained for B/P suggests the possible existence of a value effect in the sample, which is further investigated in more detail. Assuming competitive equilibrium, not only would the ROEs be close to their cost of equity (*Ke*), but also the book-to-market ratio would be close to 1 (Frankel and Lee, 1998). In general, the results described appear to reflect a consistent sample for conducting empirical tests.[19]

[17] Return on Equity decomposition by industry. Total Market (without financials). ROE (unadjusted) Aggregated net income divided by aggregated book value of equity, across all firms in the group**.** https://pages.stern.nyu.edu/~adamodar/New_Home_Page/dataarchived.html.
[18] Beta Regression Coefficients by Industry. Weighted average of 2-year and 5-year weekly return regression betas with 2-year betas weighted 2/3rds). https://pages.stern.nyu.edu/~adamodar/New_Home_Page/ dataarchived.html.
[19] As indicated in the data and sample description section, in order to conduct the analyst forecast errors analysis there was a small change in the original sample represented by the exclusion of only 13 observations. Since there are no significant or material changes in the results of descriptive statistics, data distributions and Spearman correlation tests that justify new comments and conclusions, particularly in terms of its consistency for carrying out empirical testing see Appendix B.2 for the Descriptive Statistics, Data distribution by industry, Data distribution by year and the Spearman correlation of the sample related to the analyst forecast errors.

3.2 Choosing RIV-EBO model best alternatives

In order to compare alternative metric values derived from the RIV-EBO model and choose the best alternative to confront the *B/P* model, Table 2 reports the results for Spearman correlation tests performed between both the stock price at the end of December in year (t-1) ($P_{(t-1)}$) and at the end of June in year (t) ($P_{(t)}$) and (i) the book value per share *B(t-1)*, (ii) the 3 fundamental values based on analyst earnings forecasts (Vf) and (iii) the 3 fundamental values based on historical earnings (Vh).[20] Based on evidence in the literature (Brown et al., 1987a, 1987b; O'Brien, 1988; Frankel and Lee, 1998) that observed the superiority of analyst earnings forecasts over random-walk models, the I/B/E/S analyst earnings forecasts should represent a better proxy for market earnings expectations compared to historical earnings (Frankel and Lee, 1998). However, despite the results from foreign literature, Gatsios et al. (2021), when revisiting the topic in the Brazilian market, argued that there is no unrestricted superiority, showing that the subject remains open for empirical studies.

Table 2: Spearman correlation.

	Spearman correlation						
	B(t-1)	*V1h*	*V2h*	*V3h*	*V1f*	*V2f*	*V3f*
$\boldsymbol{P_{(t-1)}}$							
Corr	0.57	0.63	0.63	0.62	0.74	0.77	0.78
R^2	0.323	0.401	0.395	0.389	0.544	0.599	0.608
$\boldsymbol{P_{(t)}}$							
Corr	0.55	0.64	0.63	0.63	0.73	0.76	0.77
R^2	0.304	0.403	0.399	0.394	0.531	0.583	0.591
Obs.	902	902	902	902	902	902	902

This Table presents the correlation coefficients and respective R^2 (i) between the stock price at the end of December of year (t-1) ($P_{(t-1)}$) and at the end of June of year (t) ($P_{(t)}$) and B*(t-1)*, (ii) between $P_{(t-1)}$ and $P_{(t)}$ and each of the 3 (three) fundamental values based on analyst earnings forecasts (*Vf*) and (iii) between $P_{(t-1)}$ and $P_{(t)}$ and each of the 3 (three) fundamental values based on historical values (*Vh*) for the sample comprising all years (t) from 1998 to 2018. *B(t-1)* represents the book value per share of the firms in the sample at the end of December of year (t-1). *V1f, V2f* and *V3f* and *V1h, V2h* and *V3h* represent fundamental values based on analyst forecasts and historical earnings, respectively. Obs. represents the total number of observations in the sample.

We find for $P_{(t-1)}$ that the book value *B(t-1)* is able to explain about 32% of the price variation in the sample. Both fundamental values based on analyst forecasts and historical earnings (*Vfs* and *Vhs*) obtain a correlation with stock prices higher than *B(t-1)*. *Vh* is able to explain around 40% of the variation in stock prices, and as the number of periods in the models (*V1h*, *V2h*, *V3h*) increases, the explanatory power of *Vh* tends to decrease, possibly indicating a decay in the predictive power of the historical data ROE_t as time passes. As observed in the U.S. market by Frankel and Lee (1998), *Vh*, by using historical data ROE_t for fundamental value estimation, appears to contain relevant information not captured by *B(t-1)*, and the highest correlation coefficients among all models are obtained from using analyst earnings forecasts (*Vfs*). *V2f* and *V3f* are able to explain about 60% of the variation in stock prices, and it is observed that, unlike *Vh*, as the number of periods and forecast horizons increases (*V1f*, *V2f*, *V3f*), the explanatory power of *Vf* tends to increase, even if to a small extent. This indicates that, when using analyst earnings forecasts, the RIV-EBO model is not only able to capture more relevant information than that reflected in historical ROEs, but also potentially better reflects market expectations for future earnings. For comparison, the correlation coefficient (Corr) results obtained by Frankel and Lee (1998) in the U.S. market based on the average of annual cross-section correlations relative to the stock price at the end of June in year (t) were 0.60, 0.70, 0.69, 0.69, 0.80, 0.81 and 0.82, respectively. Despite slightly higher values compared to the local market, the conclusions of the results for both markets

[20] Despite Dissanaike and Lim (2010) pointing out the possibility of the existence of a January effect, in which abnormal returns would be higher in January due to the stocks sale with losses in December by investors seeking to offset previous capital gains (tax loss selling), our tests to assess the correlation between (i) $P_{(t-1)}$ and B(t-1) and (ii) $P_{(t-1)}$ and Vf, show superior results compared to tests conducted with stock prices in June of year (t).

are similar. Based on the results of Table 2, we select *V2f* and *V3f* as the RIV-EBO models which will be comparing results with portfolios formed based on *B/P* ratio and size (*ME*).

3.3 Analyzing the predictive power for stock returns: uni-dimensional portfolios

After evaluating the sample and selecting the RIV-EBO models, we analyze the results of portfolios formed based on V2f, V3f, *B/P* and *ME*. To assess the statistical significance of the differences in the results between the portfolios T3 (*High ratio*) and T1 (*Low ratio*) (T3-T1) we use the Monte Carlo technique for portfolio simulation. This technique is similar to those discussed in Barber and Lyon (1997), Kothari and Warner (1997), Frankel and Lee (1998) and Lyon et al. (1999). As Frankel and Lee (1998) point out, since the investment strategies, both in the sample and reference portfolios, are made from the same set of observations and the returns are calculated in the same way and at the same period of time, i.e., from a buy-and-hold strategy over the same investment horizon, the use of randomization in forming simulated (reference) portfolios helps address potential biases identified in the literature such as new listing and survivor, rebalancing, and skewness biases.

The return calculation of buy-and-hold strategy over the same time horizon, argue Frankel and Lee (1998), also helps to adjust for potential serial correlation of returns arising from overlapping holding periods. It should be noted that all observations that make up the sample have available data for calculating returns of 12, 24 and 36 months, thus also avoiding delisting events. Furthermore, p-values estimated from the distribution of the difference between *High-Low* portfolio ratios results, obtained in the simulation procedure, represent an appropriate and commonly used approach to deal with skewness bias (Frankel and Lee, 1998). Note that Lyon et al. (1999) also support the use of p-values estimated from the distribution of long-term abnormal returns mean obtained from the pseudo-portfolios simulation procedure to deal with skewness bias and to assess the statistical significance of the sample results.

Table 3: Uni-dimensional Portfolios.

Table 3.a: V2f/P Portfolios

	T1 (*Low Vf/P*)	T2	T3 (*High Vf/P*)	All obs. (*Avg.*)	T3-T1 (*Dif.*)
V2f/P	0.51	0.95	2.20	1.22	1.69***
B/P	0.70	0.91	1.75	1.12	1.05***
ME	9867.98	9240.74	8233.66	9114.13	-1634.32
Beta	0.83	0.95	0.78	0.85	-0.05
Ret12m	0.056	0.077	0.205	0.113	0.148***
Ret24m	0.221	0.195	0.404	0.273	0.183***
Ret36m	0.355	0.546	0.681	0.527	0.326**
Ret36m p.a.	0.106	0.156	0.189	0.152	
Sharpe 36m	-0.12	0.06	0.19	-	
Stdv Ret36m p.a.	0.26	0.31	0.26	0.28	
Ke	13.1%	13.0%	13.0%	13.0%	
Obs.	301	300	301	902	

Table 3.b: V3f/P Portfolios

	T1 (*Low Vf/P*)	T2	T3 (*High Vf/P*)	All obs. (*Avg.*)	T3-T1 (*Dif.*)
V3f/P	0.56	1.02	2.38	1.32	1.82***
B/P	0.69	0.93	1.74	1.12	1.05***
ME	9778.11	8807.08	8755.74	9113.65	-1022.37
Beta	0.85	0.92	0.78	0.85	-0.07*
Ret12m	0.068	0.075	0.195	0.113	0.126***

Ret24m	0.218	0.228	0.374	0.273	0.156***
Ret36m	0.407	0.502	0.671	0.527	0.264**
Ret36m p.a.	0.121	0.145	0.187	0.152	
Sharpe 36m	-0.06	0.03	0.18	-	
Stdv Ret36m p.a.	0.28	0.29	0.27	0.28	
Ke	13.1%	13.0%	12.9%	13.0%	
Obs.	301	300	301	902	

Table 3.c: B/P Portfolios

	T1	T2	T3	All obs.	T3-T1
	(*Low B/P*)		(*High B/P*)	(*Avg.*)	(*Dif.*)
B/P	0.29	0.77	2.30	1.12	2.00***
V2f/P	0.72	1.12	1.82	1.22	1.11***
V3f/P	0.81	1.18	1.97	1.32	1.17***
ME	12909.56	7573.33	6853.95	9112.28	-6055.61***
Beta	0.84	0.85	0.87	0.85	0.03
Ret12m	0.054	0.099	0.184	0.113	0.130***
Ret24m	0.162	0.260	0.397	0.273	0.235***
Ret36m	0.348	0.579	0.653	0.527	0.305**
Ret36m p.a.	0.105	0.165	0.182	0.152	
Sharpe 36m	-0.12	0.09	0.16	-	
Stdv Ret36m p.a.	0.27	0.29	0.28	0.28	
Ke	12.5%	13.1%	13.4%	13.0%	
Obs.	301	300	301	902	

Table 3.d: ME Portfolios

	T1	T2	T3	All obs.	T3-T1
	(*Low ME*)		(*High ME*)	(*Avg.*)	(*Dif.*)
ME	758.71	2610.27	23951.38	9106.78	23192.67***
B/P	1.63	0.99	0.73	1.12	-0.90***
V2f/P	1.49	1.19	0.98	1.22	-0.51***
V3f/P	1.61	1.28	1.07	1.32	-0.54***
Beta	0.89	0.84	0.81	0.85	-0.08**
Ret12m	0.133	0.080	0.125	0.113	-0.008
Ret24m	0.357	0.198	0.265	0.273	-0.092*
Ret36m	0.721	0.344	0.515	0.527	-0.205*
Ret36m p.a.	0.198	0.104	0.149	0.151	
Sharpe 36m	0.18	-0.14	0.04	-	
Stdv Ret36m p.a.	0.35	0.24	0.25	0.28	
Ke	14.0%	13.3%	11.8%	13.0%	
Obs.	301	300	301	902	

The results for the portfolios formed from the entire sample, covering all years (t) from 1998 to 2018, are presented based on *fundamental value ratio* (*V2f/P* and *V3f/P*), book-to-market ratio (*B/P*) and size (*ME*). Except for Ret36m p.a., Sharpe 36m and Stdv Ret36m p.a., the results for each tertile represent equally weighted averages of all observations (pooled data) that make up each tertile, regardless of year (t) to which the observation refers. The return of each observation is cumulative and estimated from the beginning of July of year (t) in which the observation is collected to compose the sample. T1, T2, and T3 represent each of the tertiles into which the entire sample is sorted. Except for Obs., which represents the total number of observations in the sample, All obs. *(Avg.)* represent the simple average of the results from T1, T2 and T3. T3-T1 is the difference between the results obtained in the highest (High) and lowest (Low) tertiles. *V2f/P* and *V3f/P* represent the fundamental value ratio obtained from the estimation of fundamental value based on analyst forecasts: two-year ahead ROE forecast and three-year ahead ROE forecast, respectively. B/P is the book-to-market ratio relative to the year (t-1). B represents the book value per share at the end of December in year (t-1). P represents the stock price determined at the end of December in year (t-1). *ME* represents the market value of equity of the firms in the sample in millions of Brazilian reais, calculated as the total number of shares outstanding multiplied by the stock price, both values determined at the end of June in year (t). Beta represents the estimated beta for each observation calculated from monthly returns starting from July of year (t) over a period of 36 months, considering Ibovespa as the estimate of market return. *Ret12m*, *Ret24m* and *Ret36m* are the average returns of each portfolio based on the buy-and-hold strategy for the periods of 12, 24 and 36 months, respectively, calculated from the beginning of July of year (t) in which the observation is collected to compose the sample. Ret36m p.a. is the annualized *Ret36m*. Sharpe 36m is the Sharpe ratio calculated based on three components: (i) Ret36m p.a., (ii) the annualized return of Ibovespa derived from the equally weighted average of the 36-month cumulative return of each year between 1998 and 2018, replacing the risk-free rate, and (iii) Stdv Ret36m p.a., the standard deviation of equally weighted annualized returns derived from the 36-month cumulative return of the observations that make up the portfolio. *Ke* represents the cost of equity of the firms in the sample estimated in accordance with equation 10. Obs. Represents the total number of observations in the sample that make up the respective portfolio. The statistical significance of the differences is based on Monte Carlo simulation tests. The simulated portfolios are formed by random allocation (without replacement) of the entire sample into portfolios (T1, T2, and T3). The notation ***, **, * indicates that the observed difference between T3-T1 is significantly different from the differences of the simulated portfolios at the 1%, 5%, and 10% levels respectively (two-tailed).

3.3.1 Choosing the best RIV-EBO model

Initially, when analyzing the results of the V2f/P and V3f/P portfolios, we find that in both models the High-ratio portfolio (T3) obtains significantly higher returns than the Low-ratio (T1) for 12, 24 and 36 months. Moreover, the differences T3-T1 (High-Low) of V2f/P in these periods (14.8%, 18.3% and 32.6%) are higher compared to those of the V3f/P model (12.6%, 15.6% and 26.4%). Except for the T2 portfolio in 24 months in the V2f/P model, the relation of V2f/P and V3f/P with their respective returns show to be positive and monotonic across the tertiles in all investment horizons. Despite T3 higher returns, this portfolio has in both models the lowest beta in the entire sample (0.78), suggesting that the V2f/P and V3f/P returns are not necessarily associated with a higher variation of market risk, as corroborated when comparing V2f/P and B/P results. Due to the (slight) superiority of the V2f/P model relative to V3f/P for the returns, Sharpe ratio, standard deviation of annualized 36-month returns and especially in the difference obtained between the T3-T1 portfolios in all returns, we choose the V2f/P model to confront B/P and compare their predictive power for returns.

3.3.2 Revealing RIV-EBO model safety attribute once compared against B/P

Comparing *V2f/P* and *B/P* results, we find that, as in *V2f/P*, the returns on *B/P High-ratio* portfolio for 12, 24 and 36 months are considerably higher when compared to *B/P Low-ratio* portfolio, suggesting the existence of a value effect.[21] Furthermore, the relation between *B/P* and its respective returns are shown to be positive and monotonic across the portfolios in all investment horizons. The T3-T1 (*High-Low*) differences of the *B/P* model (13% and 30.5%) for 12 and 36-month investment horizons are slightly lower than those of *V2f/P* (14.8% and 32.6%), but higher at 24 months (23.5% vs. 18.3%). The returns of *B/P* model *High-ratio* portfolio for 12, 24 and 36 months (18.4%, 39.7%, and 65.3%) are also slightly lower than *V2f/P* (20.5%, 40.4%, and 68.1%). However, unlike *V2f/P*, which has a beta of 0.78 for the T3 portfolio (the lowest beta in the entire sample), the return of the *B/P* model in T3 is obtained at the expense of a higher beta (0.87, the highest beta among *B/P* portfolios).[22]

We also observe a lower Sharpe ratio for the *B/P* model compared to *V2f/P* over 36 months (0.16 vs. 0.19), a higher standard deviation of annualized 36-month returns (0.28 vs. 0.26) and a higher *Ke* (13.4 vs. 13.0), evidencing *B/P* stronger association of obtaining higher returns with higher differences in market risk when compared to the RIV-EBO model. Additionally, the variation of the cost of equity (*Ke*) over T1 and T3 for the *B/P* model (12.5% - 13.4%) is slightly higher than that of *V2f/P* (13.1% - 13.0%). Furthermore, we find that not only the difference in market equity (ME) between T3-T1 in the *B/P* model (-6055.61 million) is relevant and statistically significant at 1%, but also that the average market equity (ME) of the T3 portfolio in the *B/P* model is lower than that of *V2f/P* (6853.95 vs. 8233.66 million), and that the latter market value of equity is closer to the sample average (9113.99 million). These results, combined with the higher beta of *B/P* T3 portfolio, suggest a greater influence of size effect in the *B/P* model.

The *V2f/P* and *B/P* models also show a robust positive relation between them. Just as *V2f/P* exhibits a significant B/P ratio (1.75) in portfolio T3, the *B/P* model reveals a considerable V2f/P ratio (1.82) in the same portfolio as well, helping to

[21] Result also observed in the U.S. market by Fama and French (1992), Lakonishok et al. (1994), Dechow and Sloan (1997), Frankel and Lee (1998) and Chen et al. (2008) and in the Brazilian market by Braga and Leal (2000), Rostagno et al. (2006) and Santos and Montezano (2011).

[22] Since the portfolio cumulative returns are not estimated based on annual cross-section data, but pooled data, note for reference (not benchmark purpose) that an annual equally weighted average of monthly cumulative returns over 12, 24 and 36 months between 1998 and 2018, estimated from the beginning of July of the year (t) in which IBOV index is considered, resulted in 14.3%, 30.3% and 47.3%, respectively. I.e., the 36-month cumulative return of 47.3% represents the equally weighted average of all cumulative returns obtained in the 36th month of each year (t) between 1998 and 2018.

explain the small difference in cumulative return obtained by the two models at the end of 12 and 36 months. In order to contextualize this relation, we present in Table 4 the results of Frankel and Lee (1998) for the U.S. market and replicate the results of *V2f/P* and *B/P High* and *Low-ratio* portfolios from Tables 3.a and 3.c for the local market.[23] Due to the use of the *V3f/P* model (three-year ahead forecast) in the U.S. market, we refer to the RIV-EBO model applied in the U.S. and Brazilian markets as *Vf/P* throughout the text when comparing them against the *B/P* model in these markets.

Table 4: Vf/P and B/P Portfolios: Average return USA vs. Brazil: 12, 24 and 36 months.
Long strategy (buy-and-hold)[24]

USA			
	High ratio	*Low ratio*	*High – Low*
Vf/P	16.9%, 36.9% and 63.7%	13.8%, 21.7% and 33.1%	3.1%, 15.2% and 30.6%
B/P	18.6%, 33.3% and 55.8%	13.7%, 25.1% and 40.7%	4.9%, 8.2% and 15.1%
Brazil			
	High ratio	*Low ratio*	*High – Low*
Vf/P	20.5%, 40.4% and 68.1%	5.6%, 22.1% and 35.5%	14.8%, 18.3% and 32.6%
B/P	18.4%, 39.7% and 65.3%	5.4%, 16.2% and 34.8%	13.0%, 23.5% and 30.5%

Except for the low return observed in the Brazilian market, relative to the U.S. market, in the *Low-ratio* portfolio at 12 months in both models, in general, there is consistency in the results between the markets over the investment horizons. In addition, unlike the Brazilian market, the *B/P* model achieves a higher return than *Vf/P* in the *High-ratio* portfolio in the first 12 months, evidencing the strength of the market multiple in predicting returns in the short term in the U.S. market. However, we observe that *Vf/P* average return becomes superior (inferior) at the end of 24 and 36 months in the *High* (*Low*) *ratio* portfolio compared to *B/P* average return, resulting in more significant differences between *High-Low* portfolios in favor of *Vf/P* relative to *B/P* at 24 and 36 months (15.2% vs. 8.2% and 30.6% vs. 15.1%) in contrast to the local market (18.3% vs. 23.5% and 32.6% vs. 30.5%).

In summary, when comparing the results of the *B/P* and *Vf/P* models at the end of 24 and 36 months between the two markets, a more resilient positive relation is observed between the two models in the Brazilian market, possibly limiting the RIV-EBO model from achieving significantly superior returns compared to *B/P* in the local market. This is because, unlike the Brazilian market, the difference between *High* and *Low-ratio* portfolios obtained by the *Vf/P* model is approximately twice that obtained by *B/P* in the U.S. market (for a more detailed view on the robust relation between these models see bi-dimensional portfolio analyses in Appendix B.3. One possible explanation for the greater superiority of the RIV-EBO model over *B/P* in the U.S. market is U.S. analysts' lower forecast errors, particularly in the 24 and 36-month investment horizons, compared to the Brazilian ones, as further detailed in our analyst forecast error analysis. Additionally, the results therein suggest that when

[23] The returns presented in the study for the U.S. market are cumulative and represent averages of portfolio (quintile) results formed annually (annual cross-sectional portfolios) from a buy-and-hold strategy over 12, 24 and 36 months for a sample of approximately 18,100 observations in the period between 1975 and 1993. That study used the *V3f* model (three-year ahead ROE forecast).

[24] Except for the *High-Low* result for 36 months in the Brazilian market, all other *High-Low* differences were significant at 1%, indicating a statistically significant difference between the two portfolios.

forming portfolios based on *B/P*, analysts in the local market tend to make fewer (more) forecast errors in the *High* (*Low*) ratio portfolio, formed by undervalued (overvalued) firms, which in turn generates higher (lower) return.[25]

Resuming analyzing the *V2f/P* and *B/P* models in the Brazilian market, we find that both models prove to be robust predictors of returns for *High-ratio* portfolio in a long-position and show similar returns at the end of 12, 24 and 36 months. Despite its simplicity compared to the greater sophistication of the *V2f/P* model, *B/P* evidences robustness assessing value from undervalued firms, yet it is observed that obtaining higher returns is associated with higher differences in market risk. On the other hand, the value creation approach is able to capture relevant information when selecting undervalued firms in *High-ratio* portfolio, without necessarily incurring higher betas and resorting to firms with lower market equity, signaling the RIV-EBO model usefulness in capturing, along with the profitability, the safety quality attribute pointed out in Asness et al. (2013) and Lee (2014).

3.3.3 Taking into account size

Examining the results of *ME* relative to *V2f/P* and *B/P* models, we observe that *ME Low-ratio* portfolio (T1), with an average market equity value equivalent to only 758.71 million and the highest B/P and V2f/P ratios (1.63 and 1.49) compared to T2 and T3, obtains the highest return in the sample for 36 months, a slightly higher value than the *V2f/P* model (72.1% vs 68.1%). However, this return is achieved at the expense of a beta of 0.89 (higher than the *V2f/P* and *B/P* models) and a standard deviation of 0.35 (the highest in the sample), resulting in a Sharpe ratio of 0.18 (slightly lower than the *V2f/P* model (0.19)). The superiority of the T1 portfolio (Low *ME*) over T3 (High *ME*) is significant for 24 and 36 months, showing the presence of a small firm effect[26] in the medium and long term, as inferred from Fama and French (1995) and Chen and Zhang (1998), and consequently a higher market risk. The differences between T3-T1 in these periods (9.2% and 20.5% in absolute value) are smaller than those of the *V2f/P* and *B/P* models, indicating *ME* low robustness in differentiating *High* and *Low-ratio* portfolios in all investment horizons. In addition, we find that the variation of the cost of equity (*Ke*) throughout T1 and T3 (14.0% - 11.8%) is also the highest in the entire sample. Overall, the RIV-EBO model safety attribute is also present when size (ME) is considered, raising the value creation approach as the best risk-adjusted return predictor alternative in the long term.

3.4 Analyzing the predictive power for stock returns: regression analysis

In order to evaluate *V2f/P*, *B/P* and *ME* portfolio results, mitigating potential biases from using simulated portfolios[27], we apply additional formal and robust statistical methods based on the entire sample without segregation of data into tertiles. First, we run regressions using percentile and decile ranks of *V2f/P*, *B/P* and *ME* values and, in a second step, we use fixed/random effects models and first differences to consider possible effects of heterogeneity and serial correlation also based on *V2f/P*, *B/P* and *ME* values.

[25] Potential differences in earnings (ROEs), cost of equity (*Ke*) and dividend payout ratio (*k*) between these markets do not seem to explain such superiority in our case, since they seem consistent assuming a competitive market in which the ROE would adjust to the applicable cost of equity: average ROEs observed in the sample (17.07%) vs. approx. 13% observed by Frankel and Lee (1998) and Damodaran in the U.S. market (see Descriptive statistics section 3.1); *Ke* observed by the same authors ranging from 8.23% to 16.46% and in the last 7 years, as reported by Damodaran, has an average of 8.26% vs. 7.64% to 25.76% in the sample with an average of 13.01%; and average *k* observed by the authors of 27% vs. 31.88% of the sample. The 8.26% average represent a simple average between the years 2016 and 2022 considering the entire market excluding financials, as disclosed by Damodaran in his "Cost of Capital by Industries" report. https://pages.stern.nyu.edu/~adamodar/New_Home_Page/dataarchived.html.

[26] For further evidence of small firm effect in the Brazilian market. See Miralles-Quiros et al. (2017)

[27] For discussion of the need to complement Monte Carlo simulation procedures with the use of additional robust statistical methods see Frankel and Lee (1998). This is due to that allocating firms randomly into quintiles result in the covariance of the reference (simulated) portfolios formed being equal to the average covariance across the sample returns. Nonetheless, in case the actual correlation of the reference portfolios significantly differs from this average covariance, the p-values estimated in the simulations may be potentially biased.

3.4.1 Shedding light with top and bottom firms and long & short investment portfolios

Given the relevance of the existing association between the strength of the predictive power of a given model and the achievement of greater differences in returns between *High* and *Low-ratio* portfolios, particularly identified in the U.S. market (Table 4), it becomes important to evaluate coefficients obtained as the difference in returns between top and bottom firms and as returns from a long and short investment portfolio.

To do so, first we regress the cumulative return of each observation (as a dependent variable) on *V2f/P*, *B/P* and *ME* expressed in terms of their percentile ranks (as independent variables). This estimation procedure in addition to reducing the influence of possible outliers allows us to interpret the coefficients obtained from the regression as the difference in returns between top and bottom ranked percentile firms and compare them (Frankel and Lee, 1998).

Second, as Dechow and Sloan (1997) clarify, the use of scaled decile ranks in the independent variables allows interpreting the coefficients obtained as estimates of the return to a zero-investment portfolio with a long position in the firms in the highest decile and a short position in the firms in the lowest decile. With respect to *V2f/P* and *B/P*, Dechow et al. (1999) point out that the lower (higher) decile consists of overpriced (underpriced) firms relative to their intrinsic value, expecting therefore lower (higher) future returns from them. Similarly to Dechow and Sloan (1997), Bernard and Thomas (1990) and Frankel and Lee (1998) also used scaled decile ranks in the independent variables. To obtain scaled decile ranks, we sort *V2f/P*, *B/P* and *ME* values into deciles (from 1 to 10) and each decile is subtracted by 1 and divided by 9 so that they vary from 0 (lowest decile) to 1 (highest decile).

In Table 5.a we report the results of pooled OLS regressions of the cumulative return of each observation, calculated from the beginning of July of the year (t) in which the observation is collected to compose the sample, adopting a buy-and-hold strategy in the periods of 12, 24 and 36 months, on the values of *V2f/P*, *B/P* and *ME* of each observation expressed in terms of their percentile ranks (percentile ranks or (PRK(.)).

Table 5.a: Percentile Rank Analysis.

Percentile Ranks (multiple regressions)				Percentile Ranks (simple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**	***Variable***	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	-0.047	0.137	0.491	Intercept	0.007	0.134	0.276
P-Value	0.373	0.166	0.056	P-Value	0.810	0.009	0.000
t-stat	-0.891	1.387	1.917	t-stat	0.240	2.616	3.548
				PRK(V2f/P)	**0.211**	**0.279**	**0.502**
PRK(V2f/P)	**0.156**	**0.141**	**0.375**	P-Value	0.0001	0.002	0.001
P-Value	0.015	0.201	0.050	t-stat	4.032	3.078	3.465
t-stat	2.429	1.281	1.961	R^2	0.018	0.010	0.007
				DW-stat	1.521	1.717	1.937
PRK(B/P)	**0.119**	**0.210**	**0.059**	Intercept	0.019	0.112	0.313
P-Value	0.085	0.083	0.832	P-Value	0.506	0.031	0.004
t-stat	1.723	1.733	0.212	t-stat	0.666	2.166	2.857
				PRK(B/P)	**0.187**	**0.323**	**0.428**
PRK(ME)	**0.043**	**-0.078**	**-0.361**	P-Value	0.0003	0.0004	0.009
P-Value	0.455	0.477	0.208	t-stat	3.594	3.544	2.637
t-stat	0.748	-0.711	-1.260	R^2	0.014	0.014	0.005
				DW-stat	1.529	1.738	1.955
R^2	0.022	0.016	0.011	Intercept	0.137	0.376	0.769
Adj. R^2	0.019	0.013	0.007	P-Value	0.0E+00	0.0E+00	0.0E+00
F-stat	6.686	4.906	3.199	t-stat	4.117	5.779	4.768

Prob(F-stat)	0.0002	0.002	0.023	**PRK(ME)**	**-0.049**	**-0.206**	**-0.484**
DW-stat	1.538	1.733	1.943	P-Value	0.360	0.042	0.041
Obs.	902	902	902	t-stat	-0.916	-2.038	-2.046
				R^2	0.001	0.006	0.006
				DW-stat	1.489	1.704	1.943
				Obs.	902	902	902

This Table shows the coefficients and pooled OLS regression results for the entire sample, covering the year (t) from 1998 to 2018. The dependent variable is the cumulative return of each observation, calculated from a buy-and-hold strategy for periods of 12, 24 and 36 months starting in July of year (t) in which the observation is included in the sample (Ret12m, Ret24m and Ret36m). The independent variable is the percentile rank of *V2f/P*, *B/P* and *ME* (PRK(.)). The left column shows the results of 3 multiple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and PRK(V2f/P), PRK(B/P) and PRK(ME) as explanatory variables for each regression. The right column shows the results of 09 simple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and PRK(V2f/P), PRK(B/P) and PRK(ME) as a single explanatory variable for each regression. *Percentile Ranks* or PRK(.) represent the values of V2f/P, B/P, and ME expressed in terms of their percentile ranks. V2f/P, B/P, B, P, and ME have the same definitions as in Table 3. DW-stat represents the Durbin-Watson statistic. Obs. is the total number of observations in the sample. *Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance*.

In Table 5.b we report the results of pooled OLS regressions of cumulative return (dependent variable), calculated in the same way described in the percentile rank procedure, on *V2f/P*, *B/P*, and *ME* values expressed in scaled decile ranks or (DRK(.) (independent variables).

Table 5.b: Decile Rank Analysis.

Decile Ranks (multiple regressions)				Decile Ranks (simple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**	***Variable***	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	-0.030	0.148	0.496	Intercept	0.020	0.154	0.319
P-Value	0.525	0.105	0.044	P-Value	0.446	0.001	0.0E+00
t-stat	-0.636	1.623	2.018	t-stat	0.763	3.264	4.234
				DRK(V2f/P)	**0.186**	**0.239**	**0.416**
DRK(V2f/P)	**0.135**	**0.111**	**0.298**	P-Value	0.0001	0.003	0.001
P-Value	0.016	0.256	0.070	t-stat	3.975	2.949	3.303
t-stat	2.404	1.137	1.818	R^2	0.017	0.009	0.006
				DW-stat	1.520	1.714	1.936
DRK(B/P)	**0.110**	**0.202**	**0.071**	Intercept	0.029	0.128	0.342
P-Value	0.073	0.067	0.779	P-Value	0.264	0.008	0.002
t-stat	1.798	1.831	0.281	t-stat	1.117	2.652	3.171
				DRK(B/P)	**0.167**	**0.291**	**0.370**
DRK(ME)	**0.041**	**-0.063**	**-0.308**	P-Value	0.0004	0.001	0.015
P-Value	0.444	0.530	0.242	t-stat	3.524	3.482	2.427
t-stat	0.765	-0.628	-1.171	R^2	0.014	0.014	0.005
				DW-stat	1.529	1.739	1.954
R^2	0.021	0.015	0.009	Intercept	0.133	0.363	0.734
Adj. R^2	0.018	0.012	0.006	P-Value	0.0E+00	0.0E+00	0.0E+00
F-stat	6.452	4.708	2.803	t-stat	4.248	5.964	4.929
Prob(F-stat)	0.0003	0.003	0.039	**DRK(ME)**	**-0.042**	**-0.179**	**-0.415**
DW-stat	1.539	1.732	1.943	P-Value	0.398	0.053	0.052
Obs.	902	902	902	t-stat	-0.845	-1.940	-1.945
				R^2	0.001	0.005	0.006
				DW-stat	1.489	1.704	1.944
				Obs.	902	902	902

This Table presents the coefficients and of pooled OLS regression results of the entire sample, comprising the year (t) from 1998 to 2018, with the cumulative return of each observation, calculated from a buy-and-hold strategy over periods of 12, 24 and 36 months from the beginning of July in year (t) in which the observation is collected for the sample (Ret12m, Ret24m and Ret36m), as the dependent variable and the scaled decile ranks of *V2f/P*, *B/P* and *ME* (DRK(.)) as the independent variable. The left column shows the results of 3 multiple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and DRK(V2f/P), DRK(B/P) and DRK(ME) as explanatory variables for each regression. The right column shows the results of 09 simple regressions with Ret12m, Ret24m and

Ret36m as the dependent variable and DRK(V2f/P), DRK(B/P) and DRK(ME) as a single explanatory variable for each regression. *Scaled decile ranks* or DRK(.) represent the values of *V2f/P*, *B/P* and *ME* expressed in scaled decile ranks, i.e., the *V2f/P*, *B/P* and *ME* criteria are ranked in deciles (from 1 to 10), and each decile is subtracted by 1 and divided by 9, so that they range from 0 (lowest decile) to 1 (highest decile). *V2f/P*, *B/P*, B, P and *ME* have the same definitions as in Table 3. DW-stat represents the Durbin-Watson statistic. Obs. is the total number of observations that make up the sample. *Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance.*

Controlling for independent variables, we find that both *B/P* and *V2f/P* models have robust and similar predictive power (with a small superiority of the latter) for short-term return. Nonetheless, despite the *B/P* model being able to maintain the robustness of its predictive power up to 24 months (unlike *V2f/P* which does not show statistically significant result), by using analyst earnings forecasts to proxy for market expectations and incorporating growth projection from invested capital in the model *V2f/P* showed, to some extent, greater predictive power for returns in the long term, contributing to explaining returns in the 36-month period.

Providing a context for a Decile Rank analysis comparison, Frankel and Lee (1998), controlling for *B/P* and *ME*, observed in the U.S. market, like in the Brazilian market, similarity in returns between *Vf/P* and *B/P* models in the short term, with a small superiority of *B/P* (3.1% significant at 5% vs. 3.9% significant at 1%) in 12 months. In the 36-month horizon, however, the *Vf/P* model in the U.S. market was superior to *B/P* (35.2% significant at 1% vs. 5.1% not statistically significant). These results are also consistent with those obtained in the local market, which as revealed also show the superiority of *Vf/P* over *B/P* in the long term.

All results for 12, 24 and 36 months are not statistically significant for *ME*, supporting the lower power of differentiation of returns across extreme portfolios and the lower statistical significance of the T3-T1 results obtained by *ME* formed portfolios (-1%, -9% and -21%* - Table 3.d) compared to *V2f/P* and *B/P*.

In summary, the RIV-EBO model reveals to be alike to *B/P* in explaining the overall sample firms' returns in the short term, with the RIV-EBO model signaling a greater predictive power when extending the investment horizon to 36 months. These results are similar to those observed in the U.S. market in Frankel and Lee (1998) study, suggesting a relevant pattern, to be potentially exploited by trading strategies, that reflects RIV-EBO model gain and *B/P* apparent loss of explanatory power for returns in the long term.

3.4.2 Taking into account heterogeneity and serial correlation

By pooling observations without considering the firm and year (t) to which the observation refers, the pooled OLS method adopted in the percentile and decile rank regressions does not allow considering possible heterogeneity or individuality among the observations that make up the sample, potentially resulting in correlation between the error term and the independent variables and eventually biased coefficients. To consider possible heterogeneity effects and observe *V2f/P* and *B/P* predictive power for returns from their ratio values, we apply the fixed or random effects models to the sample, conditioned to the result of Hausman tests performed for all independent variables (using multiple and simple regressions) and investment horizons.

Wooldridge (2013) emphasizes the importance of considering unobserved time-constant factors that can affect the dependent variable (unobserved effects). In addition, when using the fixed effects model, in practice a different intercept is allowed for each cross-sectional unit or firm (Wooldridge, 2013). The use of the random effects model, on the other hand, is justified when unobserved effects are not correlated with the independent variables, requiring carrying out the Hausman test previously (Wooldridge, 2013).

After conducting Hausman tests, except for *V2f/P* 36-month return simple regression for which we use cross-section random effect (RE), we regress, by panel least squares method and fixed effects for cross-sectional units or firms (dummy variables) (FE), the cumulative return of each observation from a buy-and-hold strategy in the periods of 12, 24 and 36 months (dependent variable) on *V2f/P*, *B/P* and *ME* values (independent variables).

In contrast to the percentile and decile rank analysis, in Table 6.a we use the values of the *V2f/P* and *B/P* ratios and the market value of equity for *ME* expressed in billions of Brazilian reais, which allows the coefficients obtained to be interpreted as the average change in the sample return when increasing one unit of these ratios and *ME* value, while keeping the other variables constant. *ME* scaling adjustment allows for closer result values to those of *V2f/P* and *B/P*, making the interpretation of *ME* results easier.

Table 6.a: Fixed/Random Effects Analysis.

Fixed Effects (multiple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	-0.005	0.117	0.352
P-Value	0.913	0.176	0.011
t-stat	-0.109	1.355	2.563
V2f/P	**0.105**	**0.134**	**0.132**
P-Value	0.002	0.017	0.095
t-stat	3.163	2.398	1.674
B/P	**0.025**	**0.071**	**0.145**
P-Value	0.254	0.122	0.083
t-stat	1.142	1.548	1.737
ME	**-0.004**	**-0.010**	**-0.016**
P-Value	0.007	0.006	0.013
t-stat	-2.716	-2.787	-2.492
R^2	0.216	0.304	0.361
Adj. R^2	0.074	0.178	0.245
F-stat	1.524	2.417	3.118
Prob(F-stat)	0.0003	0.0E+00	0.0E+00
DW-stat	2.048	1.604	1.670
Obs.	902	902	902

Fixed Effects/Random Effects (simple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	-0.039	0.050	0.284
P-Value	0.321	0.429	0.013
t-stat	-0.992	0.792	2.481
V2f/P	**0.124**	**0.182**	**0.182**
P-Value	0.0001	0.001	0.001
t-stat	3.888	3.491	3.234
R^2	0.207	0.286	0.008
DW-stat	2.070	1.609	1.380
Intercept	0.059	0.155	0.318
P-Value	0.088	0.018	0.004
t-stat	1.708	2.375	2.862
B/P	**0.048**	**0.105**	**0.186**
P-Value	0.105	0.069	0.051
t-stat	1.625	1.823	1.957
R^2	0.188	0.283	0.352
DW-stat	2.077	1.621	1.675
Intercept	0.166	0.386	0.713
P-Value	0.0E+00	0.0E+00	0.0E+00
t-stat	7.695	9.093	7.907
ME	**-0.006**	**-0.012**	**-0.020**
P-Value	0.001	0.001	0.003
t-stat	-3.487	-3.404	-2.944
R^2	0.188	0.281	0.350
DW-stat	2.082	1.605	1.662
Obs.	902	902	902

This Table presents the coefficients and regression results for the entire sample from 1998 to 2018, using the panel least squares method with fixed effects (FE) for cross-sectional units or firms (dummy variables) for all variables and periods, except for *V2f/P* simple regression return for 36 months, for which the random effects model is used for cross-sectional units or firms (RE). The dependent variable is the cumulative return of each observation from a buy-and-hold strategy for periods of 12, 24 and 36 months, calculated from the beginning of July of year (t) in which the observation is collected to compose the sample (Ret12m, Ret24m and Ret36m) and the independent variable is the value of *V2f/P*, *B/P* and *ME* for each observation. The left column shows the results of 3 multiple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and V2f/P, B/P and ME as explanatory variables for each regression. The right column shows the results of 09 simple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and V2f/P, B/P and ME as a single explanatory variable for each regression. *V2f/P*, *B/P*, B and P have the same definitions as in Table 3. *ME* represents the market value of equity of the firms in the sample in billions, calculated as shares outstanding multiplied by the stock price, both values determined at the end of June of year (t). DW-stat represents the Durbin-Watson statistic. Obs. is the total number of observations that make up the sample. White (diagonal) standard errors and covariance (d.f. corrected). *Heteroskedasticity robust*.

Using multiple regressions, we find that both the V2f/P and B/P models show long-term predictive power for returns, as an increase in their respective ratios lead to an average increase in the overall sample return (13.2% and 14.5%). Nonetheless, despite the average return of 24 months being close to that of 36 months (13.4% and 13.2%), the *V2f/P* model shows more stable and statistically significant average return increases over the 12, 24 and 36-month investment horizons compared to *B/P*, which yields non statistically significant results for 12 and 24 months. As for *ME*, all results are significant, at 1% for 12 and 24 months and 5% for 36 months, also identifying the influence of a small firm effect, given the negative coefficient sign.

Running simple regressions, the *V2f/P* model also shows greater consistency and statistically significance across all investment horizons compared to *B/P* (12.4%, 18.2% and 18.2% significant at 1% vs. 4.8%, 10.5% and 18.6% not statistically significant for 12 months and significant at 10% for 24 and 36 months). When controlling for heterogeneity, the results suggest greater consistency and statistical significance in predicting returns over time in favor of *V2f/P* over *B/P*.

Finally, due to the possible influence of firms' results between subsequent periods throughout the sample, it is also relevant to evaluate for within-cross section serial correlation. For this purpose, we use first differences accompanied by the White period (cross-section cluster) method, whose estimator aims to accommodate heteroscedasticity and serial correlation for the sample data. Wooldridge (2013) argues that inferences from fixed effect estimators could be potentially more sensitive to non-normality, heteroscedasticity, and serial correlation in idiosyncratic errors and also emphasizes that for first differences, in general, the approach to obtaining "fully robust standard errors and test statistics" is known as clustering, which is justified in panel data when N is substantially larger than T (Wooldridge (2013, pp. 483, 511) which is our case.

As in the fixed/random effects analysis, we use in Table 6.b. the values of *V2f/P* and *B/P* ratios and the market value of equity expressed in billions of reais for *ME*.

Table 6.b: First Difference Analysis.

First Difference (multiple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	0.031	0.017	0.065
P-Value	0.009	0.335	0.042
t-stat	2.648	0.968	2.053
V2f/P	**0.115**	**0.128**	**0.147**
P-Value	0.018	0.067	0.078
t-stat	2.396	1.850	1.777
B/P	**0.058**	**0.069**	**0.108**
P-Value	0.116	0.197	0.158
t-stat	1.582	1.297	1.421
ME	**-0.013**	**-0.020**	**-0.027**
P-Value	0.093	0.047	0.157
t-stat	-1.695	-2.012	-1.424
R^2	0.091	0.068	0.027
Adj. R^2	0.088	0.064	0.023
F-stat	25.560	18.438	7.131
Prob(F-stat)	0.0E+00	0.0E+00	0.0001
DW-stat	2.494	2.160	2.238
Obs.	766	766	766

First Difference (simple regressions)			
Variable	**Ret12m**	**Ret24m**	**Ret36m**
Intercept	0.025	0.009	0.053
P-Value	0.012	0.574	0.056
t-stat	2.563	0.564	1.932
V2f/P	**0.153**	**0.176**	**0.219**
P-Value	0.001	0.013	0.010
t-stat	3.342	2.529	2.636
R^2	0.052	0.032	0.011
DW-stat	2.562	2.172	2.235
Intercept	0.022	0.005	0.049
P-Value	0.040	0.754	0.079
t-stat	2.081	0.314	1.772
B/P	**0.094**	**0.111**	**0.157**
P-Value	0.085	0.132	0.113
t-stat	1.735	1.518	1.597
R^2	0.033	0.021	0.010
DW-stat	2.549	2.174	2.231
Intercept	0.027	0.013	0.060
P-Value	0.024	0.478	0.063
t-stat	2.279	0.712	1.874
ME	**-0.015**	**-0.022**	**-0.030**
P-Value	0.078	0.041	0.135
t-stat	-1.780	-2.065	-1.505
R^2	0.037	0.035	0.015
DW-stat	2.550	2.161	2.241
Obs.	766	766	766

This Table shows the coefficients and regression results of a sample comprising the period (t) from 1998 to 2018, using panel least squares, first differences (FD) and cross-section clustering method. The regression in question has the cumulative return of each observation as the dependent variable, calculated from a buy-and-hold strategy for the periods of 12, 24 and 36 months starting from July of year (t) in which the observation is collected to compose the sample (Ret12m, Ret24m and Ret36m) and the independent variable is the value of *V2f/P*, *B/P* and *ME* for each observation. The left column shows the results of 3 multiple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and V2f/P, B/P and ME as explanatory variables for each regression. The right column shows the results of 09 simple regressions with Ret12m, Ret24m and Ret36m as the dependent variable and V2f/P, B/P and ME as a single explanatory variable for each regression. *V2f/P*, *B/P*, B and P have the same definitions as in Table 3. *ME* represents the market value of equity of the sample firms in billions, calculated as shares outstanding multiplied by the stock price, both values determined at the end of June of year (t). DW-stat is the Durbin-Watson statistic. Obs. is the total number of observations in the sample. White period (cross-section cluster) standard errors & covariance (d.f. corrected). *Heteroskedasticity and within cross section serial correlation robust*.

While also controlling for serial correlation, the results of *V2f/P* not only show more stable and statistically significant average in return compared to *B/P*, but also a positive and monotonic relation with its returns over the investment horizons, evidencing an even more robust incremental contribution of the model in explaining returns obtained over 12, 24 and 36 months (11.5%, 12.8% and 14.7%, statistically significant in all investment horizons), which do not occur with the *B/P* model (results not statistically significant for all investment horizons). By controlling for serial correlation, *B/P* not only loses the statistical significance of the result obtained for 36 months controlling for other variables, but also the significance of the simple regression results obtained for 24 and 36 months reported in Table 6.a (fixed/random effects), with only the simple regression result for 12 months remaining statistically significant in Table 6.b (9.4% significant at 10%). Running simple regressions, the *V2f/P* model reports average returns of 15.3%, 17.6% and 21.9% (significant at 1%, 5% and 1%) for 12, 24 and 36 months. We may conclude that, while controlling for both heterogeneity and serial correlation, the RIV-EBO model exhibits a more robust and stable contribution in predicting returns over 12, 24 and 36 months, compared to the *B/P* model.

Overall, the results of this first part of the analysis reveal that, in general, the RIV-EBO model seems to fit well in the Brazilian market, even though Brazilian accounting may be considered uninformative (Ohlson and Lopes, 2007), and that a more complete valuation approach contributes to explaining returns in the Brazilian market relative to a book-to-market ratio trading strategy, finding also witnessed by Frankel and Lee (1998) and Dissanaike and Lim (2010) in the U.S. and British markets, respectively.

3.5 Analyzing analyst forecast errors: OP, B/P, ME and Ke portfolios

In this second part of our analysis we investigate analyst forecast errors from uni-dimensional portfolios based on analyst optimism *OP*, book-to-market ratio *B/P*, size *ME* and cost of equity *Ke*. To assess the statistical significance of the differences in the results between portfolios T3 (*High ratio*) and T1 (*Low ratio*) (T3-T1), we also use the Monte Carlo technique for portfolio simulation.

Table 7: OP, B/P, ME and Ke Portfolios.

Table 7.a: OP Portfolios

	T1 *(Low OP)*	T2	T3 *(High OP)*	All obs. *(Avg.)*	T3-T1 *(Dif.)*
OP	-0.30	0.33	3.07	1.03	3.38***
B/P	1.52	0.87	1.00	1.13	-0.52***
ME	8716.32	11997.24	6785.62	9166.40	-1930.70
Ke	13.03%	13.04%	13.05%	13.04%	0.02%
V2f/P	1.17	1.20	1.29	1.22	0.12
Beta	0.87	0.79	0.88	0.85	0.01
ROE(t-1)	0.202	0.199	0.103	0.168	-0.099***
FROE(t)	0.123	0.229	0.198	0.183	0.075***
FROE(t+1)	0.135	0.249	0.236	0.207	0.101***
FROE(t+2)	0.146	0.250	0.242	0.213	0.096***
FErr(t)	0.042	0.066	0.098	0.069	0.056***
FErr(t+1)	0.068	0.111	0.168	0.116	0.100***
FErr(t+2)	0.112	0.134	0.207	0.151	0.094***
Ret12m	0.114	0.104	0.130	0.116	0.015

Ret24m	0.253	0.253	0.329	0.279	0.076
Ret36m	0.477	0.495	0.620	0.530	0.143
Ret36m p.a.	0.139	0.143	0.174	0.152	
Stdv Ret36m p.a.	0.273	0.255	0.304	0.277	
Sharpe 36m	0.003	0.022	0.121		
Obs.	297	296	296	889	

Table 7.b: B/P Portfolios

	T1 *(Low B/P)*	T2	T3 *(High B/P)*	All obs. *(Avg.)*	T3-T1 *(Dif.)*
B/P	0.30	0.78	2.31	1.13	2.01***
OP	1.31	1.43	0.36	1.03	-0.95***
ME	13013.13	7552.46	6919.09	9161.56	-6094.04***
Ke	12.57%	13.18%	13.38%	13.04%	0.81%***
V2f/P	0.71	1.12	1.83	1.22	1.11***
Beta	0.82	0.85	0.87	0.85	0.05
ROE(t-1)	0.220	0.152	0.132	0.168	-0.088***
FROE(t)	0.277	0.164	0.107	0.183	-0.170***
FROE(t+1)	0.316	0.183	0.120	0.206	-0.196***
FROE(t+2)	0.325	0.185	0.127	0.212	-0.199***
FErr(t)	0.090	0.063	0.054	0.069	-0.036***
FErr(t+1)	0.169	0.100	0.079	0.116	-0.090***
FErr(t+2)	0.232	0.112	0.109	0.151	-0.122***
Ret12m	0.066	0.092	0.190	0.116	0.124***
Ret24m	0.192	0.239	0.405	0.279	0.213***
Ret36m	0.390	0.537	0.664	0.531	0.275**
Ret36m p.a.	0.116	0.154	0.185	0.152	
Stdv Ret36m p.a.	0.269	0.279	0.282	0.277	
Sharpe 36m	-0.081	0.059	0.168		
Obs.	297	296	296	889	

Table 7.c: ME Portfolios

	T1 *(Low ME)*	T2	T3 *(High ME)*	All obs. *(Avg.)*	T3-T1 *(Dif.)*
ME	755.73	2608.40	24161.95	9175.36	23406.22***
OP	0.78	1.22	1.10	1.03	0.32
B/P	1.65	0.99	0.73	1.13	-0.92***
Ke	14.01%	13.28%	11.83%	13.04%	-2.18%***
V2f/P	1.50	1.20	0.96	1.22	-0.54***
Beta	0.89	0.84	0.81	0.85	-0.08**
ROE(t-1)	0.149	0.162	0.193	0.168	0.044***
FROE(t)	0.149	0.185	0.214	0.183	0.065***
FROE(t+1)	0.169	0.212	0.239	0.207	0.070***
FROE(t+2)	0.175	0.216	0.247	0.213	0.072***
FErr(t)	0.092	0.059	0.055	0.069	-0.037***
FErr(t+1)	0.136	0.109	0.102	0.116	-0.034**
FErr(t+2)	0.156	0.156	0.140	0.151	-0.016
Ret12m	0.139	0.083	0.126	0.116	-0.013

Ret24m	0.370	0.203	0.263	0.278	-0.107*
Ret36m	0.741	0.352	0.498	0.530	-0.243**
Ret36m p.a.	0.203	0.106	0.144	0.151	
Stdv Ret36m p.a.	0.341	0.236	0.244	0.273	
Sharpe 36m	0.191	-0.136	0.026		
Obs.	297	296	296	889	

Table 7.d: Ke Portfolios

	T1 *(Low Ke)*	T2	T3 *(High Ke)*	All obs. *(Avg.)*	T3-T1 *(Dif.)*
Ke	10.45%	12.55%	16.14%	13.04%	5.69%***
OP	1.55	0.67	0.87	1.03	-0.68**
B/P	1.05	1.02	1.31	1.13	0.26**
ME	13865.77	10885.97	2730.05	9160.60	-11135.72***
V2f/P	1.28	1.18	1.20	1.22	-0.08
Beta	0.92	0.75	0.87	0.85	-0.05
ROE(t-1)	0.169	0.175	0.160	0.168	-0.009
FROE(t)	0.171	0.196	0.181	0.183	0.010
FROE(t+1)	0.192	0.220	0.208	0.207	0.016
FROE(t+2)	0.200	0.227	0.211	0.212	0.011
FErr(t)	0.070	0.067	0.068	0.069	-0.002
FErr(t+1)	0.104	0.124	0.120	0.116	0.016
FErr(t+2)	0.148	0.167	0.138	0.151	-0.010
Ret12	0.125	0.137	0.086	0.116	-0.039
Ret24	0.259	0.312	0.265	0.279	0.006
Ret36	0.485	0.655	0.451	0.530	-0.035
Ret36m p.a.	0.141	0.183	0.132	0.152	
Stdv Ret36m p.a.	0.282	0.285	0.265	0.277	
Sharpe 36m	0.011	0.158	-0.022		
Obs.	297	296	296	889	

The results for the portfolios formed from the entire sample, covering all years (t) from 1998 to 2018, are presented based on analyst optimism (*OP*), book-to-market ratio (*B/P*), size (*ME*) and cost of equity (*Ke*). Except for Ret36m p.a., Sharpe 36m and Stdv Ret36m p.a., as defined below, the results for each tertile represent equally weighted averages of all observations (pool) that make up each tertile, regardless of year (t) to which the observation refers. The return of each observation is estimated from the beginning of July of year (t) in which the observation is collected to compose the sample (cumulative return). T1, T2, and T3 represent each of the tertiles into which the entire sample is sorted. Except for Obs., which represents the total number of observations in the sample, All obs. *(Avg.)* represent the simple average of the results from T1, T2, and T3. T3-T1 is the difference between the results obtained in the highest (High) and lowest (Low) tertiles. *OP* represents the measure derived from the RIV-EBO model that reflects the analyst optimism estimated in terms of the equation 12. B/P, B, P, *ME*, *Ke*, V2f/P, Beta, *Ret12m*, *Ret24m*, *Ret36m*, Ret36m p.a., Stdv Ret36m p.a. and Sharpe 36m have the same definitions as in Table 3. ROE(t-1) represents the return on equity calculated as net income divided by the book value, both reported at the end of December in year (t-1). FROE(t), FROE(t+1) and FROE(t+2) represent the forecast ROE, based on the market analyst EPS forecasts for the end of years (t), (t+1) and (t+2) respectively. FErr(t), FErr(t+1) and FErr(t+2) represent the market analyst forecast errors for years (t), (t+1) and (t+2) respectively. Obs. represents the total number of observations in the sample that make up the respective portfolio. The statistical significance of the differences is based on Monte Carlo simulation tests. The simulated portfolios are formed by random allocation (without replacement) of the entire sample into portfolios (T1, T2, and T3). The notation ***, **, * indicates that the observed difference between T3-T1 is significantly different from the differences of the simulated portfolios at the 1%, 5%, and 10% levels respectively (two-tailed).

3.5.1 Revealing analysts' optimism

Initially, we examine the results of portfolios formed using the measure derived from the RIV-EBO model which reflects analyst optimism (*OP*) estimated from equation 12. Given that the *OP* measure captures analyst optimism relative to historical ROEs, the observed difference in the OP ratio between T3 (*High OP*) and T1 (*Low OP*) (3.07 and -0.30) signals optimism for the former and pessimism for the latter. For comparison, Frankel and Lee (1998) observed in the U.S. market values of 5.23 for the highest quintile (*High OP*) and -0.018 for the lowest quintile (*Low OP*), evidencing analysts' greater optimism and less pessimism in the U.S. market compared to the Brazilian market.

A non-monotonic but positive relationship is also observed between *OP* and future ROEs forecasts (FROEs) in 1, 2 and 3 years across the tertiles, despite the reported ROE in T3 being almost half of T1 (0.103 vs. 0.202), reflecting the high confidence of analysts in the improvement of the performance of firms that did not necessarily report such significant results. Analyst optimism for future ROEs (FROEs) in 1, 2, and 3-year forecasts, in turn, reflects in an increase in analyst forecast errors across the tertiles, establishing a monotonic and positive relationship between *OP* and FErr(t). In short, despite lower reported ROEs, higher optimism (*OP*s) is associated with higher future ROE forecasts (FROEs) and higher forecast errors (FErr). That is, the divergence between firm's fundamental values based on analyst earnings forecasts and historical earnings is positively related to higher forecast errors at all forecast horizons, revealing an overall excess of analyst optimism in the local market, as observed in Martinez (2007a) and Beiruth et al. (2014).

When comparing the results of the *OP* portfolios in the Brazilian market with the results of Frankel and Lee (1998) in the U.S. market, Table 8 shows that, in general, Brazilian market analysts make on average more forecast errors, especially in the 2 and 3-year forecast periods, compared to U.S. market analysts.

Table 8: OP Portfolio: Forecast Errors (FErr) years (t), (t+1) and (t+2)[28] U.S. vs. Brazilian market.

USA			
	High OP	*Low OP*	*High–- Low*
FErr(t)	0.081	0.025	0.056
FErr(t+1)	0.126	0.050	0.076
FErr(t+2)	0.114	0.062	0.052
Brazil			
	High OP	*Low OP*	*High–- Low*
FErr(t)	0.098	0.042	0.056
FErr(t+1)	0.168	0.068	0.100
FErr(t+2)	0.207	0.112	0.094

That is, despite the higher optimism and lower pessimism in the U.S. market, as evidenced by comparing *OP* results for T3 and T1 portfolios in these markets, U.S. market analysts make fewer forecast errors compared to local market analysts. These lower forecast errors are evidenced particularly in FErr (t+1) and FErr (t+2) in both the Low *OP* and High *OP* portfolios. The results suggest that analyst optimism, or the discrepancy between firm's fundamental values based on analyst earnings forecasts and historical earnings (*OP*), in that market is "better justified" than in the local market and help to explain the greater differences between *High-Low* portfolios in favor of *Vf/P* compared to *B/P* in the U.S. market, in contrast to the Brazilian market, particularly in the 24 and 36-month periods reported in Table 4.

In other words, a possible explanation for the RIV-EBO model to achieve superior results compared to B/P in the U.S. market, compared to the local market, is attributed to analysts' greater accuracy and lower future ROE forecast errors in that market when using *Vf/P*, particularly in the 24 and 36-month investment horizons, compared to the Brazilian market. This is consistent with the results observed in Gonçalves et al. (2022), who found that optimistic analyst reports in the U.S. market are related to greater accuracy in earnings forecasts compared to analysts' reports of the same nature in the Brazilian market, and in Martinez (2008), who observed differences in market reaction between such markets. While the Brazilian market appears to be more sensitive to bad news, the U.S. market would respond in a balanced way to both good and bad news. Overall, Lima and Almeida (2015) observed lower consistency in sell-side analysts' predictions in the Brazilian market compared to international markets studies.

[28] All *High-Low* differences were significant at 1% in both studies.

Furthermore, the considerable increase in forecast error and in over-optimism in the local market in the High *OP* portfolio between FErr (t) and FErr (t+1) (from 6.9% to 11.6% - a variation of 71.4%) and in the T3-T1 difference also between FErr (t) and FErr (t+1) (from 5.6% to 10% - a variation of 78.6%) helps to explain the inferior result of the RIV-EBO model in T3-T1 difference in the 24-month investment horizon compared to *B/P*, documented in Tables 3.a and 3.c. Consistent with Martinez (2007b), this result suggests that local analysts tend to significantly increase forecast error when extending the forecast horizon to 24 months, an over-optimism that possibly also reflects in part the compounding effect of the previous year's forecast error (Frankel and Lee, 1998).

We find that, in addition to the strong and resilient relationship between *B/P* and the RIV-EBO model predicting long term returns (evidenced in Section 3.3.2), a possible explanation for the RIV-EBO model not having shown significant superiority of returns in *High-ratio* portfolio over the *B/P* strategy (Tables 3.a and 3.c) is local analysts' significant forecast errors, particularly in 24 and 36-month periods in the *High OP* portfolio. This finding is line with Beiruth et al. (2014), who highlight the impact of analyst forecast errors on target price estimation in the local market.

3.5.2 Unveiling the value/glamour effect

When forming portfolios based on *B/P*, a higher optimism (*OP*) in T1 (Low *B/P*) compared to T3 (High *B/P*) (1.31 vs. 0.36), followed by a negative and monotonic relationship between *B/P* and future ROEs forecasts (FROEs) and between *B/P* and the ROE reported in year (t-1) are evidenced.

However, despite showing a higher reported ROE (0.220 vs. 0.132), higher market value of equity (13013.13 and 6919.09), and higher optimism regarding future ROEs forecasts (FROEs), the Low *B/P* portfolio not only has a higher forecast error (FErr) but also lower return at all horizons compared to the High *B/P* portfolio.

The results highlight a negative and monotonic relationship between *B/P* and FErr (t), FErr (t+1) and FErr (t+2) and a positive and monotonic relationship between *B/P* and *Ret12*, *Ret24* and *Ret36*, supporting the value/glamour effect documented in La Porta et al. (1997), Dechow and Sloan (1997) and Piotroski and So (2012) in the U.S. market, and in Braga and Leal (2000), Rostagno et al. (2006) and Santos and Montezano (2011) in the Brazilian market.

In short, higher reported ROEs(t-1) are associated with higher optimism (*OPs*) and higher future ROE forecasts (FROEs), which in turn are associated with higher forecast errors (FErr) and lower returns (*Ret*).

It is worth noting that Fama and French (1992), Lakonishok et al. (1994) and Chen et al. (2008) also observed the outperformance of value stocks over glamour stocks in the U.S. market. The latter study found a stable value premium (difference between the average returns of value and glamour stocks) of around 6.1% per annum between 1945 and 2005 in that market. The authors reinforce that the results support Fama and French (1998) and Davis et al. (2000) (apud Chen et al., 2008), as they also concluded that such outperformance accounts for a real premium, being unlikely to result from a statistical bias. The results also suggest that the low profitability of value investment strategies in the 1990s, contrary to a permanent decline, reflects cyclical movements of the premium, and it is not possible to eliminate the mispricing hypothesis pointed out in Lakonishok et al. (1994). Yet in the Brazilian market Santos and Montezano (2011) using the price/book value (P/B) parameter observed a value premium of 18.34% (compound annual return) between 1989 and 2008.

Indeed, the cause of the value/glamour effect is subject of debate and attributed to the compensation for risk (risk premium) and mispricing hypotheses. While Fama and French (1992) associated the B/P ratio with financial distress and attributed it to compensation for assumed risk, Lakonishok et al. (1994) attributed it to the fact that value stocks are underpriced relative to their risk and return characteristics.

On one hand, the risk premium hypothesis is supported by the evidence of a negative relationship between B/P and return on book value (ROE), as well as the association of risk measures such as financial leverage, earning uncertainty and distress with value stocks returns (Fama and French, 1995; Chen and Zhang, 1998). On the other hand, the mispricing hypothesis is supported by evidence of systematic errors in investors' expectations of returns for glamour (value) stocks. This is explained by the excessive reliance on past good (bad) performance, leading to excessively optimistic (pessimistic) predictions and creating an incentive to favor (neglect) stocks with robust (weak) reported results, believing they are less (more) risky investments. Nevertheless, investors are subsequently surprised by negative (positive) returns (Lakonishok et al. 1994; La Porta et al., 1997; Piotroski, 2000; Piotroski and So, 2012). Furthermore, Piotroski (2000) reinforces that while value stocks align more with a financial statement data-based analysis, glamour stocks' valuations tend to rely on cash flow forecasts and long-term sales, causing investors to heavily rely on non-financial information, showing the momentum factor influence in glamour stocks' valuation. As Asness (1997, p. 34) states when evaluating the interaction between value and momentum strategies: "Value works, in general, but largely fails for firms with strong momentum. Momentum works, in general, but is particularly strong for expensive firms."

In line with Lakonishok et al. (1994), La Porta et al. (1997) and Piotroski and So (2012), the value/glamour effect observed in the local market indicates a tendency for analysts to be more optimistic about overvalued firms (Low *B/P*), with higher reported ROE and market equity and less associated with risk, that is, with lower cost of equity (12.57% vs. 13.38%) and lower standard deviation of annualized 36-month returns (0.269 vs. 0.282) compared to High *B/P* firms.[29]

Nonetheless, T1 (Low *B/P*) shows a higher forecast error (Ferr) and lower returns in all periods, reaffirming not only the value/glamour effect in the sample but also suggesting the mispricing hypothesis as an explanation for part of the value/glamour effect. This is a similar result to that observed by Frankel and Lee (1998) in the U.S. market for the 36-month period. Furthermore, it should be noted that the difference in betas between the High *B/P* and Low *B/P* portfolios observed in the Brazilian market (0.82 and 0.87) is not considerable or statistically significant, reinforcing the mispricing hypothesis suggested in La Porta (1996).

One possible explanation for this result, as pointed out by La Porta et al. (1997), is that large differences in price-to-earnings ratios between value and glamour stocks reflect expectations that differences in past growth rates will persist for much longer than is expected from a reliable forecast based on past data. As a result, value (glamour) stocks tend to obtain higher (lower) returns as the market takes some time to realize that earnings growth is actually higher (lower) than initially expected. In the Brazilian market, accounting information also takes some time to be incorporated into prices, some of them, representing the local economic reality, seem to be incorporated even more slowly into prices than in the U.S. market (Lopes and Galdi, 2006).

In addition, when comparing *B/P* portfolios results in the local market with Frankel and Lee (1998) results in the U.S. market, Table 9 shows that on average Brazilian market analysts make considerably more forecast errors in the portfolio formed by overvalued firms (Low *B/P*) (between 3 and 2 times more in (t+1) and (t+2)) compared to the U.S. market, while in the undervalued portfolio (High *B/P*) the forecast errors between the two markets are relatively close.

Table 9: B/P Portfolio: Forecast Errors (Ferr) years (t), (t+1) and (t+2).[30] U.S. vs. Brazilian market.

USA			
	High B/P	*Low B/P*	*High – Low*
Ferr(t)	0.060	0.037	0.023
Ferr(t+1)	0.073	0.054	0.019
Ferr(t+2)	0.074	0.113	-0.039

[29] See Lakonishok et al. (1994, p. 1576) reinforcing that institutional investors tend to prefer glamour to value stocks since they are considered "prudent investments", making it easier to justify to their clients investments in firms that have had positive results in the past and with less likelihood of financial distress.

[30] The differences between *High-Low* are significant at 1% level in both studies.

Brazil			
	High B/P	*Low B/P*	*High – Low*
Ferr(t)	0.054	0.090	-0.036
Ferr(t+1)	0.079	0.169	-0.090
Ferr(t+2)	0.109	0.232	-0.122

These results suggest that local market analysts place excessive confidence and optimism in glamour firms compared to the U.S. market and help to explain the strength of *B/P* strategy in predicting returns in the High *B/P* portfolio (Section 1.4.5.2). This is because local market analysts tend to make fewer forecast errors in the High *B/P* portfolio, which in turn is composed of undervalued firms that generate higher returns as also suggested in Dechow and Sloan (1997) and Piotroski and So (2012) in the U.S. market and Braga and Leal (2000), Rostagno et al. (2006) and Santos and Montezano (2011) in the Brazilian market.

Rostagno et al. (2006) and Varga and Brito (2016) point out that the higher return of value stocks compared to glamour stocks observed in the Brazilian market could reflect in part a compensation for the lower liquidity of the former compared to the latter. On the other hand, Galdi (2008) controlling for liquidity observed that firms with high B/P ratios still explain future returns.

Further, the substantial evidence of prediction errors in the portfolio formed by glamour firms (*Low B/P*) in the Brazilian market, possibly reflecting in part the momentum factor pointed out in Asness (1997), Piotroski (2000) and Varga and Brito (2016), signal the existence of mispricing as an explanation for part of the observed value/glamour effect. Reinforcing the mispricing hypothesis in the local market, Galdi and Lima (2017) observed an increase in returns by combining strategies that explore the value/glamour effect and positive surprises when reporting firm earnings. It should be also noted that there is no consensus on whether the Brazilian capital market has a weak or a semi-strong form of informational efficiency (Camargos and Barbosa, 2006).

Overall, the empirical findings suggest that Brazilian market lower information efficiency makes financial signals not necessarily have the same relevance as observed in the U.S. market in Frankel and Lee (1998) study (for additional discussion of such relevance see also Lopes and Galdi, 2006).

3.5.3 Taking into account size and cost of equity

Finally, we examine portfolios based on market value of equity (*ME*) and cost of equity (Ke). In the *ME* portfolios we find a positive and monotonic relationship between *ME* and reported ROE(t-1), and *ME* and future ROE forecasts (FROEs), but a negative and monotonic relationship between *ME* and FErr(t) and FErr(t+1). The results reveal that analysts tend to make fewer forecast errors for firms with higher market value of equity, and that FErr(t+1) and FErr(t+2) T3-T1 difference in absolute terms (result not significant for FErr(t+2)) is much smaller in the *ME* portfolios (3.4% and 1.6%) compared to the *B/P* portfolio (9.0% and 12.2%), suggesting that market equity is a less relevant factor in explaining analyst forecast errors compared to *B/P*.

In fact, Low *ME* higher long-term return, compared to High *ME* (0.741 vs. 0.498), seems to reconcile with a risk premium explanation due to the higher beta (0.89 and 0.81), higher cost of equity (14.01% and 11.83%) and higher standard deviation (0.341 and 0.244) observed in Low *ME* compared to High *ME*. That is, while the mispricing hypothesis seems to explain part of the observed value/glamour effect in the *B/P* portfolios, the risk premium hypothesis seems to explain the small firm effect observed in the *ME* portfolios.

Examining the portfolios formed based on cost of equity *(Ke*), despite the significant difference in Ke (10.45% vs. 16.14%) and optimism (*OP*) (1.55 vs. 0.87) between the T1 and T3 portfolios, we find that most of the remaining variables are not statistically significant, suggesting that the cost of equity does not seem to be a relevant factor that influences analyst forecast errors. These results may be explained by the relative high concentration of Ke values (standard deviation of 0.028) around the mean (13,04%) in spite of the range from 7.64% to 25.76% observed in the sample and reported in Table 13 Appendix B.2.

3.6 Analyzing analyst forecast errors: percentile rank regression

As carried out in the first part of our analysis, we report the results of pooled OLS regressions of market analyst forecast errors for years (t), (t+1), and (t+2) (FErr(t+i)), on *OP*, *B/P*, *ME* and *Ke* values expressed in terms of their percentile ranks (percentile ranks or (PRK(.)).

Table 10: Percentile Rank Analysis.

Percentile Ranks (multiple regressions)			
Variable	**FErr(t)**	**FErr(t+1)**	**FErr(t+2)**
Intercept	0.160	0.228	0.343
P-Value	0.000	0.000	0.000
t-stat	3.786	5.394	5.276
PRK(OP)	**0.058**	**0.111**	**0.065**
P-Value	0.004	0.002	0.100
t-stat	2.900	3.045	1.646
PRK(B/P)	**-0.092**	**-0.175**	**-0.250**
P-Value	0.000	0.000	0.000
t-stat	-4.034	-6.441	-4.062
PRK(ME)	**-0.113**	**-0.142**	**-0.143**
P-Value	0.000	0.000	0.000
t-stat	-3.625	-3.635	-3.809
PRK(Ke)	**-0.036**	**-0.019**	**-0.058**
P-Value	0.162	0.480	0.135
t-stat	-1.398	-0.706	-1.498
R^2	0.059	0.078	0.038
Adj. R^2	0.055	0.074	0.034
F-stat	13.966	18.797	8.697
Prob(F-stat)	0.000	0.000	0.000
DW-stat	2.082	1.692	2.063
Obs.	889	889	889

Percentile Ranks (simple regressions)			
Variable	**FErr(t)**	**FErr(t+1)**	**FErr(t+2)**
Intercept	0.030	0.040	0.088
P-Value	0.007	0.010	0.000
t-stat	2.729	2.583	4.015
PRK(OP)	**0.076**	**0.152**	**0.126**
P-Value	0.000	0.000	0.005
t-stat	4.005	4.366	2.807
R^2	0.021	0.037	0.009
DW-stat	2.053	1.641	2.051
Intercept	0.100	0.189	0.257
P-Value	0.000	0.000	0.000
t-stat	10.236	13.373	6.675
PRK(B/P)	**-0.062**	**-0.145**	**-0.211**
P-Value	0.000	0.000	0.000
t-stat	-3.998	-6.502	-3.564
R^2	0.014	0.034	0.027
DW-stat	2.034	1.658	2.048
Intercept	0.096	0.141	0.154
P-Value	0.000	0.000	0.000
t-stat	7.786	7.027	8.008
PRK(ME)	**-0.055**	**-0.050**	**-0.006**
P-Value	0.004	0.088	0.870
t-stat	-2.886	-1.707	-0.163
R^2	0.011	0.004	0.000
DW-stat	2.061	1.665	2.040
Intercept	0.069	0.106	0.165
P-Value	0.000	0.000	0.000
t-stat	5.581	7.840	6.408
PRK(Ke)	**-0.002**	**0.020**	**-0.029**
P-Value	0.936	0.365	0.435
t-stat	-0.081	0.907	-0.781
R^2	0.000	0.001	0.001
DW-stat	2.037	1.650	2.043
Obs.	889	889	889

This Table presents the coefficients and pooled OLS regression results for the entire sample, covering the year (t) from 1998 to 2018, with market analyst forecast errors for years (t), (t+1) and (t+2) estimated according to equation 11 (FErr(t+i)) as the dependent variable, and percentile ranks of *OP*, *B/P*, *ME* and *Ke* (PRK(.)) as the independent variable. The left column presents the results of 3 multiple regressions with FErr(t), FErr(t+1), FErr(t+2) as the dependent variable and PRK(OP), PRK(B/P), PRK(ME) and PRK(Ke) as explanatory variables for each regression. The right column presents the results of 12 simple regressions with FErr(t), FErr(t+1), FErr(t+2) as the dependent variable and PRK(OP), PRK(B/P), PRK(ME) and PRK(Ke) as a single explanatory variable for each regression. Percentile Ranks or PRK(.) represent the *OP*, *B/P*, *ME*, and *Ke* values expressed in terms of their percentile ranks. *OP*, *B/P*, *ME* and *Ke* have the same definitions as in Table 7. DW-stat represents the Durbin-Watson statistic. Obs. is the total number of observations in the sample. *Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance*.

Controlling for independent variables, we observe a significant increase in the differences of forecast errors between top and bottom ranked firms in the 24-month compared to the 12-month horizon for *OP* (5.8% - 11.1%). This result, consistent with the observed increase in FErr in the same periods in the uni-dimensional analysis (5,6% and 10,0% - Table 7.a), suggests that local analysts tend to significantly increase forecast errors over 24 months compared to the 12-month forecast horizon. This, in turn, also helps explain the lack of statistical significance of the RIV-EBO model return over 24-month holding period in the percentile and decile rank regressions (Tables 5.a and 5.b).

An interesting pattern, nonetheless, is revealed since the difference of forecast errors in FErr(t+2) returned to the same levels as in the 12-month horizon (around 6%), contrary to Martinez (2007b). This result reinforces analysts' overconfidence in predicting earnings over the 24-month horizon and suggests that the three-year-ahead biases did not reflect the compounding effect of the previous year's forecast error. In the analysis of simple regression, the same forecast error increase in 24 months and (less pronounced) decline in 36 months are observed.

Based on *B/P*, due to the negative sign, we find that analysts make more forecast errors in overvalued firms (Low *B/P*) compared to undervalued ones (High *B/P*) in all prediction horizons. Furthermore, the difference in forecast errors between higher and lower ranking firms is positively associated in absolute terms with the prediction horizon, resulting in the inferiority of forecast errors of undervalued firms compared to overvalued ones in 9.2%, 17.5%, and 25%. In other words, the analysts' overconfidence in favor of glamour stocks over value stocks, as suggested in Lakonishok et al. (1994), is noticeable. This result is similar to that observed in the simple regressions approach.

The negative sign of *ME* in the regressions, as in *ME* portfolios (Table 7.c), shows that analysts make more forecast errors in firms with lower market equity in all forecast horizons, with no considerable differences among the coefficients in FErr (t), (t+1) and (t+2), especially when analyzing the results of the simple regressions, statistically significant only for FErr(t) and FErr(t+1) (5.5% and 5.0%). These results support the lower relevance of market equity in explaining analyst forecast errors compared to *B/P*.

Finally, the results show that the cost of equity (*Ke*) does not contribute in general to explaining analyst forecast errors in the sample, as also evidenced analyzing *Ke* portfolios (Table 7.d). All the results, in both multiple and simple regressions analysis, are found to be not statistically significant for FErr(t), FErr(t+1), and FErr(t+2).

Overall, while controlling for the other independent variables, we find that *B/P* provides stronger and superior explanatory power for analyst forecast errors compared to the other firm characteristics. This supports *B/P* portfolios analysis carried out in Section 3.5.2 and signals the particular relevance of *B/P* as an analyst forecast error driver in the Brazilian market which, along with analyst optimism (*OP*), may be used as variables to compute firm's predicted forecast error, potentially incorporated in trading strategies taking short position in predicted forecast error high ranked firms, aiming to improve the predictive power for returns of the RIV-EBO model strategy.

Given the reliance of analyst-based RIV-EBO model to analysts' accuracy in forecasting earnings in different investment horizons, our findings suggest that the search for significantly enhancing its predictive power for returns relative to *B/P* in the Brazilian market, and likely the weakening of these two models resilient and positive relationship in predicting returns, require overall a material improvement in the quality of analyst earnings forecasts in the local market due to their influence on capital market informational efficiency, as also suggested by Frankel et al. (2006), and the adoption of additional predicted forecast error trading strategies incorporating *B/P* and *OP* as explanatory variables (Frankel and Lee, 1998).

This finding is consistent with Gonçalves et al. (2022), who, based on observed differences in behavioral biases impacting analysts' accuracy in different environments, particularly between Brazil and the USA, suggest that these biases should be considered in calibrating the models that predict analysts' accuracy. In addition, it should also not be disregarded the presence of analysts' correction (or learning) behavior over forecast errors derived from excessive optimistic/pessimistic predictions in the Brazilian market, unlike other markets, observed by Lima and Almeida (2015), which is also positive for increasing local market informational efficiency.

4. Conclusion

Investigating the performance of models based on fundamental signals that better reflect market expectations for future earnings and also the influence of behavioral factors on the explanatory power of such factors or drivers are a relevant and challenging quest in the literature. Examining the contribution of a value creation approach to explain stock returns relative to a B/P strategy, our findings suggest a safety quality attribute, when portfolio sorting is used, and a greater predictive power for returns in the long term, for the overall sample firms' returns, in favor of the RIV-EBO model, meaning that, overall, a more sophisticated value creation approach contributes to explaining returns in the Brazilian market. Nonetheless, due to observed local analysts' higher forecast errors compared to the U.S. market, the search for enhancing the predictive power for returns of the RIV-EBO model in the Brazilian market requires the improvement in the quality of analyst earnings forecasts and the use of B/P and analyst optimism (OP) as explanatory variables for additional predicted forecast error trading strategies.

A suggestion for future research is investigating the superiority of *V2f/P* (a two-year-ahead ROE forecast model) over *V3f/P* (a three-year-ahead ROE forecast model) in the 36-month investment horizon in light of the results herein. Since we document the predictive power inferiority of *V2f/P* compared to the *B/P* in the 24-month investment horizon, and also observe (i) significant increase forecast errors over the 24-month prediction horizon compared to the 12-month horizon and (ii) the return of three-year-ahead analyst forecast errors to the same levels as in the 12-month horizon, one explanation to be explored is the possibility that overly optimistic two-year-ahead analyst forecasts would be realized in a 36-month period, meaning that by incorporating the two-year analyst biases the *V2f/P* model would be capturing the 36-month period realized returns.

Finally, it is important to highlight the limitation of data availability in the Brazilian market, which made it unfeasible to distribute the sample into quintiles and form annual cross-sectional portfolios, preventing a greater difference in sample between extreme portfolios and the observation of results in annual periods. In addition, existing differences in accounting treatment and segregation of information for common and preferred shares may cause occasional distortions in the estimation of the residual income and the forecasted ROE. Transaction costs and liquidity risk may represent obstacles to the implementation of trading strategies in the local market, such as those indicated here, particularly for value stocks.

References

Abarbanell, J. S. and Bernard, V. (2000). Is the U.S. stock market myopic? *Journal of Accounting Research, 38*(2), 221–242.

Abarbanell, J. S. and Bushee, B. J. (1997, Spring) Fundamental Analysis, Future Earnings, and Stock Prices. *Journal of Accounting Research, 35*(1), 1-24.

Affleck-Graves, J. and Miller, R. (2003). The information content of calls of debt: evidence from long-run stock returns. *Journal of Financial Research, 26*(4), 421–447.

Almeida, J. E. F. and Dalmacio, F. Z. (2015). The Effects of Corporate Governance and Product Market Competition on Analysts forecasts: Evidence from the Brazilian Capital Market. *The International Journal of Accounting, 50*(3), 316–339.

Almeida, J. E. F., Brito, G. A. S., Batistella, F. D., and Martins, E. (2012). Análise dos modelos de avaliação residual income valuation, abnormal earnings growth e fluxo de caixa descontado aplicados às ofertas públicas de aquisição de ações no Brasil. *Revista de Contabilidade e Organizações, 6*(16), 3–19.

Almeida, J. E. F., Lima, G. A. S. F., Lima, I. S., Securato, J. R. (2012). Analysis of the Residual Income Valuation and Abnormal Earnings Growth Models: A Practical Approach using Analysts' Forecasts. *Revista de Contabilidade e Controladoria*, *4*(1), 1–16.

Anesten, S., Möller, N., Skogsvik, K., and Skogsvik, S. (2020). The pricing accuracy of alternative equity valuation models: Scandinavian evidence. *Journal of International Financial Management & Accounting, 31*:5–34.

Ang, A., Hodrick, R., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*(1), 259–299.

Artikis, P. (2018). Market efficiency and fundamental valuation. Evidence from the German stock exchange. *The Journal of Prediction Markets, 12*(2), 26–46.

Asness, C. S. (1997, March/April). The Interaction of Value and Momentum Strategies. *Financial Analysts Journal*. Association for Investment Management and Research.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2013, October 9). Quality minus junk. *Working Paper*. AQR Capital Management and New York University.

Barber, B. M. and Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics, 43*(3), 341–372.

Beiruth, A. X., Galdi, F. C., Lima, G. A. S. F., and Almeida, J. E. F. (2014, May/August). Comparação da acurácia de analistas com o modelo de *Ohlson-Juettner* (OJ) no mercado brasileiro. *Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ, 19*(2), 79–92.

Bernard, V. L. (1994, January). Accounting-based valuation methods, determinants of book-to-market ratios, and implications for financial statement analysis. *Working paper*. University of Michigan.

Bernard, V. L. (1995). The Feltham-Ohlson framework: implications for empiricists. *Contemporary Accounting Research, 11*(2), 733–747.

Bernard, V. L. and Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implication of current earnings for future earnings. *Journal of Accounting and Economics, 13*(4), 305–340.

Billett, M., Flannery, M., and Garfinkel, J. (2006). Are bank loans special? Evidence on the post-announcement performance of bank borrowers. *Journal of Financial and Quantitative Analysis, 41*(4), 733–751.

Black, F. (1972, July). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business, 45*(3), 444–455.

Braga, C. A. B. M. and Leal, R. P. C. (2000). Ações de Valor e de Crescimento nos anos 90. UFRJ; COPPEAD.

Brown, L. D., Griffin, P. A., Hagerman, R. L., and Zmijewski, M. E. (1987a). Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics, 9*(1), 61–87.

Brown, L. D., Richardson, G. D., and Schwager, S. J. (1987b). An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research, 25*(1), 49–67.

Camargos, M. A. and Barbosa, F. V. (2006, jan./mar.). Eficiência informacional do mercado de capitais brasileiro pós-Plano Real: um estudo de eventos dos anúncios de fusões e aquisições. *Revista de Administração, 41*(1), 43–58.

Campbell, J., Hilscher, J., and Szilagyi, J. (2008). In search of distress risk. *Journal of Finance, 63*(6), 2899–2939.

Chen, N.-F. and Zhang, F. (1998, October). Risk and Return of Value Stocks. *The Journal of Business, 71*(4), 501–535.

Chen, N.-F., Petkova, R., and Zhang, L. (2008). The expected value premium. *Journal of Financial Economics, 87*: 269–280.

Davis, J. L., Fama, E. F., and French, K. R. (2000). Characteristics, covariances, and average returns: 1929–1997. *Journal of Finance, 55*(1), 389–406.

Dechow, P. M. and Sloan, R. G. (1997). Returns to contrarian investment: tests of theIe expectations hypothesis. *Journal of Financial Economics, 43*(1), 3–27.

Dechow, P., Hutton, A., and Sloan, R. (1999). An empirical assessment of the residual income valuation model. *Journal of Accounting Economics, 26*(1–3), 1–34.

Dissanaike, G. and Lim, K. M. (2010). The Sophisticated and the Simple: The Profitability of Contrarian Strategies. *European Financial Management, 16*(2), 229–255.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.

Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance, 50*(1), 131–155.

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics, 43*:153–193.

Fama, E. F. and French, K. R. (1998). Value versus growth: the international evidence. *Journal of Finance, 53*: 1975–1999.

Feltham, G. A., Ohlson, J. A. (1995, Spring). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research, 11*(2), 689–731.

Ferreira, E. S., Nossa, V., Ledo, B. C. A., Teixeira, A. M. C., and Lopes, A. B. (2008, May/August). Comparação Entre Os Modelos Residual Income Valuation (RIV), Abnormal Earnings Growth (AEG) e Fluxo de Caixa Livre (FCF): Um estudo empírico no mercado de capitais brasileiro. *Brazilian Business Review, 5*(2), 152–175.

Frankel, R. and Lee, C. M. C. (1998). Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns. *Journal of Accounting and Economics, 25*(3), 283–319.

Frankel, R., Kothari, S., and Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics, 41*(1–2), 29–54.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics, 11*(1), 1–25.

Galdi, F. C. (2008). Estratégias de investimento em ações baseadas na análise de demonstrações contábeis: é possível prever o sucesso? [Doctoral dissertation, Universidade de São Paulo].

Galdi, F. C., Lima, V. S.-M. (2017). Value & Growth Investing e PEAD no Brasil. *Revista Brasileira de Finanças, 14*(4), 551–577.

Galdi, F. C., Teixeira, A. J. C., and Lopes, A. B. (2008, May/August). Análise empírica de modelos de valuation no ambiente brasileiro: fluxo de caixa descontado versus Modelo de Ohlson (RIV). *Revista Contabilidade & Finanças, 19*(47), 31–43.

Gatsios, R. C., Lima, F. G., Gaio, L. E., and Pimenta, T., Jr. (2021). Re-examining analyst superiority in forecasting results of publicly-traded Brazilian companies. *RAM, 22*(1), 1–30.

George, T. J. and Hwang, C.Y. (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics, 96*(1), 56–79.

Gonçalves, P. G., Silva, J. C., Cazetta, O. B., and Nardi, P. C. C. (2022, July 27-29). Acurácia das previsões de lucro tem relação com os vieses presentes nos relatórios dos analistas? Um estudo no Brasil e EUA. In: *19º Congresso USP de Iniciação Científica em Contabilidade* (1–20). USP.

Gordon, M. J. (1962). *The investment, financing and valuation of the Corporation*. Irwin.

Holloway, P., Rochman, R., and Laes, M. (2013). Factors Influencing Brazilian Value Investing Portfolios. *Journal of Economics, Finance and Administrative Science, 18*(Special Issue), 18–23.

Hong, H., Lim, T., and Stein, J. (2000, February). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *American Finance Association, 55*(1), 265–295.

Jorgensen, B., Lee, Y., and Yoo, Y. (2011), The Valuation Accuracy of Equity Value Estimates Inferred from Conventional Empirical Implementations of the Abnormal Earnings Growth Model: US Evidence. *Journal of Business Finance and Accounting,* 38(93–4), 446–471.

Kothari, S. P. and Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics, 43*(3), 301-340.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *Journal of Finance, 51*(5), 1715-1742.

La Porta, R., Lakonishok, J., Shleifer, A., and Vishny, R. (1997, June). Good news for value stocks: further evidence on market efficiency. *Journal of Finance, 52*(2), 859-874.

Lakonishok, J., Shleifer, A., and Vishny, R. (1994, December). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance, 49*(5), 1541–1578.

Lee, C. M. C. (1996, April). Measuring wealth. *CA Magazine, 129*(3), 32–37.

Lee, C. M. C. (2001). Market efficiency and accounting research: a discussion of 'capital research in accounting' by S. P. Kothari. *Journal of Accounting and Economics, 31*(1–3), 233–253.

Lee, C. M. C. (2014). Performance measurement: an in'estor's perspective. *Accounting and Business Research, 44*(4), 383–406.

Lima, M. P. and Almeida, V. S. (2015, July). Os analistas sell-side fazem boas previsões de preços-alvo no Brasil? *Revista Brasileira de Finanças, 13*(3), 365–393.

Lopes, A. B.; Galdi, F. C. (2006). Financial Statement Analysis also Separate Winners from Losers in Brazil. In: Seminário da EPGE (pp. 1–21). FGV.

Lyon, J. D., Barber, B. M., and Tsai, C.-L. (1999, February 1). Improved methods for tests of long-run abnormal stock returns. *Journal of Finance, 54*(1), 165–201.

Martinez, A. L.ª(2007a, May/August). Otimismo e Viés de Seleção dos Analistas. *Brazilian Business Review, 4*(2), 104–118.

Martinez, A. L. (2007b, July/December). Determinantes da Acurácia das Previsões dos Analistas do Mercado de Capitais. *UnB Contábil, 10*(2), 69–96.

Martinez, A. L. (2008, May/August). The Effect of Earnings Projection Revisions on Stock Returns in Brazil. *Brazilian Business Review, 5*(2), 121–135.

Miralles-Quiros, M. M., Miralles-Quiros, J. L., and Gonçalves, L. M. (2017, July/August). Revisiting the size effect in the Bovespa. *RAE – Revista de Administração de Empresas, 57*(4), 317–329.

Mohanram, P. (2005). Separating winners from losers among low book-to-market stocks using financial statement analysis. *Review of Accounting Studies, 10*(2–3), 133–170.

Myers, J. N. (1999, January). Implementing Residual Income Valuation with Linear Information Dynamics. *The Accounting Review, 74*(1), 1–28.

Nekrasov, A. (2016). Equity Value as a Function of (eps1, eps2, dps1, bvps, beta), Concepts and Realities. Discussion of Ohlson and Johannesson. *Abacus – A Journal of Accounting, Finance and Business Studies, 52*(1), 100–105.

O'Brien, P. C. (1988). Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics, 10*(1), 53-83.

Ohlson, J. A. (1990). A synthesis of security valuation theory and the role of dividends, cash flows, and earnings. *Contemporary Accounting Research, 6*(2)*,* 648–676.

Ohlson, J. A. (1991). The theory of value and earnings, and an introduction to the Ball-Brown analysis. *Contemporary Accounting Research, 7*(1), 1-19.

Ohlson, J. A. (1995, Spring). Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research, 11*: 661-687.

Ohlson, J. A. (2000, March). Residual Income Valuation: The Problems by James A. Ohlson. Stern School of Business.

Ohlson, J. A. (2001, Spring). Earnings, Book Values, and Dividends in Equity Valuation: An Empirical Perspective. *Contemporary Accounting Research, 18*(1), 107–120.

Ohlson, J. A. and Gao, Z. (2006), 'Earnings, Earnings Growth and Value'. *Foundations and Trends in Accounting, 1*(1), 1–70.

Ohlson, J. A. and Juettner-Nauroth, B. E. (2005). Expected EPS and EPS Growth as Determinants of Value. *Review of Accounting Studies, 10:* 349–365.

Ohlson, J. A. and Lopes, A. B. (2007, May/August). Avaliação de Empresas com Base em Números Contábeis. *Brazilian Business Review, 4*(2), 96–103.

Penman, S. H. (1992, Fall) Return to Fundamentals. *Journal of Accounting, Auditing and Finance,* 7(4), 465-484.

Penman, S. H. (2005). Discussion of "On Accounting-based Valuation Formulae" and "Expected EPS and EPS Growth as Determinants of Value". *Review of Accounting Studies, 10*(2), 367–378.

Piotroski, J. D. (2000). Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research, 38*(3), 1–41.
Piotroski, J. D. and SO, E. C. (2012, April). Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach. *The Review of Financial Studies,25*(9),
Rosenberg, B., Reid, K., and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management, 11*(3), 9–16.
Rostagno, L., Soares, R. O., and Soares, K. T. C. (2006, September/October). Estratégias de valor e de crescimento em ações na Bovespa: uma análise de sete indicadores relacionados ao risco. *Revista Contabilidade & Finanças, 42*, 7–21.
Santos, L. R. and Montezano, R. M. S. (2011, May/August). Value and growth stocks in Brazil: risks and returns for one- and two-dimensional portfolios under different economic conditions. *Revista Contabilidade & Finanças, 22*(56), 189–202.
Sanvicente, A. Z. (2015, maio). Relevância de Prêmio por Risco País no Custo de Capital das Empresas. *Revista de Administração Contemporânea*, *19*, 38–52.
Sanvicente, A. Z., Sheng, H. H., and Guanais, L. F. P. (2017) Are Country and Size Risks Priced in the Brazilian Stock Market? *Brazilian Administration Review*, *14*(1), 1–17.
Sanvicente, A. Z. (2017, 23 de outubro). Estimativas do Equity Risk Premium para o Mercado Brasileiro de Capitais. *CEQEF/FGV*.
Sanvicente, A. Z. and Carvalho, M. R. A. (2020, March). Determinants of the implied equity risk premium in Brazil. *Brazilian Review of Finance, 18*(1), 69–90.
Spiess, K. and Affleck-Graves, J. (1995). Underperformance in long-run stock returns following seasoned equity offerings. *Journal of Financial Economics, 38*(3), 243–267.
Varga, G. and Brito, R. D. (2016, June). The Cross-Section of Expected Stock Returns in Brazil. *Revista Brasileira de Finanças, 14*(2), 151–187.
Villaschi, A. W., Galdi, F. C., and Nossa, S. N. (2011). Análise fundamentalista para seleção de uma carteira de investimento em ações com baixa razão book-to-market. *Revista de Administração e Contabilidade da Unisinos, 8*(4), 325–337.
Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach* (5th ed.). South-Western Cengage Learning.

Appendix A: Methodology

A.1 Model and estimation: procedure for estimating future ROEs (FROEs)

For the purpose of estimating the stock fundamental value from equations 6 to 8, we start to estimate both the future book value and the firms' future ROEs (FROEs). By assuming the clean surplus relation (CSR), represented by equation 14, and applying it together with the firm's dividend payout ratio (*k*), the RIV-EBO model allows for the estimation of future book values:

$$B_t = B_{t-1} + NI_t - d_t \tag{14}$$

in which B_t= book value at the end of period (t), B_{t-1}= book value at the beginning of period (t), NI_t = net income at the end of period (t), d_t = dividends paid at the end of period (t).

Given that the dividend payout ratio (*k*) is represented by $k = \frac{d_{t-1}}{NI_{t-1}} = \frac{d_t}{NI_t}$, by including k and taking the variable B_{t-1} as a common factor in equation 14 we estimate the future book value as:

$$\begin{aligned} B_t &= B_{t-1} + NI_t - k * NI_t \\ &= \mathrm{B}_{t-1} + (1-k) * NI_t \\ &= \left[1 + (1-k)\frac{NI_t}{B_{t-1}}\right] * B_{t-1} \\ &= [1 + (1-k)\ ROE_t] * B_{t-1} \end{aligned} \tag{15}$$

Similarly, for the book value estimation in subsequent periods, we have:

$$\begin{aligned} B_{t+1} &= [1 + (1-k)\ ROE_{t+1}] * B_t \\ B_{t+1} &= [1 + (1-k)\ ROE_{t+1}] * [1 + (1-k)\ ROE_t] * B_{t-1} \end{aligned} \tag{16}$$

in which B_{t-1}= reported book value for period (t-1), B_t= estimated book value, based on CSR, for the end of period (t), B_{t+1} = estimated book value, based on CSR, for the end of period (t+1), k = dividend payout ratio or the net income percentage paid out as dividends in period (t-1), ROE_t = estimated ROE, based on CSR and the forecasted EPS by market analysts, for the end of year (t) according to equation 17, ROE_{t+1} = estimated ROE, based on CSR and the forecasted EPS by market analysts, for the end of year (t+1) according to equation 19.

Future ROE forecasts (FROE), in turn, are calculated based on earnings-per-share (EPS) estimated by market analysts (i) for the end of the current year (t) (one-year-ahead or *FY1*), (ii) for the end of the following year (t+1) (two-years-ahead or *FY2*), and (iii) for the end of the year after next (t+2) (three-years-ahead or *FY3*) and disclosed by I/B/E/S. Once $FROE_t$ is estimated based on equation 17 and rearranging equation 15 into the form of equation 18, it is possible to estimate the future book value for period (t) (B_t) and consequently the FROEs and future book values for the subsequent periods in the following terms:

$$FROE_t = FY1/B_{t-1} \tag{17}$$
$$B_t = B_{t-1}[1 + FROE_t(1-k)] \tag{18}$$
$$FROE_{t+1} = FY2/[(B_t + B_{t-1})/2] \tag{19}$$
$$B_{t+1} = B_t[1 + FROE_{t+1}(1-k)] \tag{20}$$
$$FROE_{t+2} = FY3/[(B_{t+1} + B_t)/2] \tag{21}$$

in which $FROE_t$ = estimated ROE based on FY1, $FY1$ = market analysts earnings per share (EPS) forecast for the end of year (t), B_{t-1} = reported book value per share for period (t-1), calculated by dividing book value by the number of shares outstanding, B_t = estimated book value per share for the end of period (t), $FROE_{t+1}$ = estimated ROE based on FY2, $FY2$ = market analysts earnings per share (EPS) forecast for the end of year (t+1), B_{t+1} = estimated book value per share at the end of period (t+1), $FROE_{t+2}$ = estimated ROE based on FY3, $FY3$ = market analysts earnings per share (EPS) forecast for the end of year (t+2).

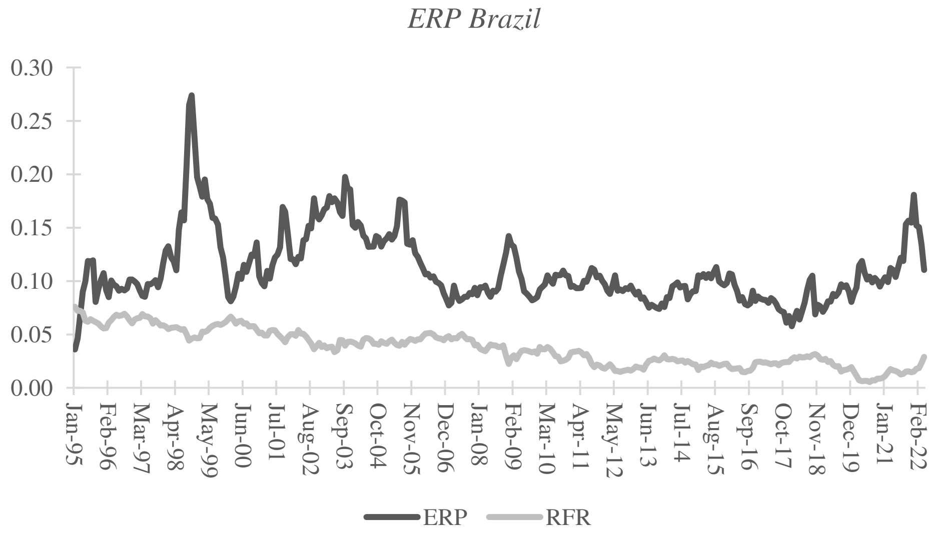
A.2 Data and sample description

A.2.1 Data description:

Eikon (Refinitiv) defines the accounting data used from its database in the following terms:
1) Common Shareholder Equity: this is the Total Shareholder's Equity as of the fiscal period Balance Sheet minus Preferred Stock and Redeemable Preferred Stock for the same period.
2) Income Available to Common Shareholders Excl Extra: represents net income available to common shareholders, not including the effect of extraordinary items, and is used to calculate Basic EPS Excluding Extraordinary Items.
3) Cash Dividend Paid – Common: represents cash dividends paid to common stockholders.
4) Total Assets Reported: represents the total assets of a company.
5) Total Common Shares Outstanding: represents the number of primary common shares equivalent outstanding.

A.2.2 Equity Risk Premium Brazil:

Monthly values of the equity risk premium (ERP) and the risk-free rate (RFR) represented by the current yield to maturity of the 10-year U.S. Treasury bond, from January 1995 to April 2022.



Source: Center of Quantitative Studies in Economics and Finance at São Paulo School of Economics – FGV https://ceqef.fgv.br/bancos-de-dados (CEQEF) and Economatica, respectively.

A.2.3 Closing Price:

The closing price obtained in Eikon (Refinitiv) database is adjusted for capital change events (e.g., rights issue, share split and stock dividend) and is defined as: "Price Close: the latest available closing price. If there are no trades for the most recent completed tradable day, the most recent prior tradable day with trading activity is used, provided the last tradable day for the instrument is within 378 completed calendar days (54 weeks)."

If no data is available from Eikon (Refinitiv) at the end of each month, the last available closing price at end of the relevant month, even if it exceeded 54 weeks, is considered. This measure (i) is justified by the need of uninterrupted availability of the closing prices for the estimation of post-ranking market betas over a period of 36 months for each stock and for the firms' market value of equity calculation at the end of June of each year (t) and (ii) does not seem to have a relevant or material impact on the sample, as there is no significant price variation (due to periods of illiquidity or mere unavailability of data in the database in period less than 54 weeks) and its use is isolated (given that the stocks in the sample were covered by market analysts).

A.3 Portfolio formation and statistical methods

A.3.1 Beta and Sharp ratio estimation:

The beta is estimated for each observation using monthly returns from the beginning of July in year (t) over a period of 36 months, considering the Ibovespa as the market return estimate. The beta of the portfolio, in turn, represents the equally weighted average of the observations (pooled data) that make up the portfolio, regardless of year (t) to which the observation refers.

The Sharpe ratio is calculated for each portfolio from (i) the annualized return derived from the equally weighted average of the 36-month cumulative returns of the observations that form the portfolio, (ii) the annualized return of Ibovespa (replacing the risk-free rate) derived from the equally weighted average of the 36-month cumulative return of each year between 1998 and 2018 and (iii) the standard deviation of equally weighted annualized returns derived from the 36-month cumulative return of the observations that form the portfolio. Despite the uneven data distribution by year in the portfolios, the use of an equally weighted average in Ibovespa annualized return estimation, item (ii), to all portfolios is justified to be consistent with the use of an equally weighted average in estimating the (pooled data) portfolio returns, item (i). The relatively high average of 13.78% obtained for Ibovespa annualized return explains the low ratios and negative signs for the Sharpe ratio in Table 3.

A.3.2 Monte Carlo simulation:

Using random allocation (without replacement) of the entire sample into tertiles (T1, T2, and T3), simulated portfolios are formed and the T3-T1 differences are estimated. The procedure is repeated 10,000 times to estimate an average of the T3-T1 difference results, enabling the construction of a probability curve and the calculation of p-values to determine whether the observed T3-T1 differences are significantly different from those estimated from the simulated portfolios (reference portfolios). In the uni-dimensional analysis, the simulated portfolios formed have the same number of observations as the original portfolios. In the bi-dimensional analyses, while we hold one of the tertiles membership of one variable constant, the tertiles membership of the other variable, which is of interest to analyze, is randomized. That is, simulated portfolios are formed from random allocation (without replacement) of the entire sample and combined with the original portfolio of the variable that remained constant.

Appendix B: Results and discussion

B.1 Original sample used for estimating predictive power for stock returns

The percentage participation of firms and sample observations by industries and the specific betas for each of the 6 industries in which the sample is sorted for the calculation of the cost of equity (*Ke*) for each firm are shown below. The distribution of the observations in each industry is reported in order to not only evaluate the sample distribution among economic industries, but also to assess to what extent the observed distribution (derived from RIV-EBO model-imposed conditions such as analysts' coverage of sample firms) differs from a distribution obtained from studies carried out by research centers, such as Nefin, that use all available data in the market. In this sense, the equally weighted average of the annual distribution of Nefin's observations from 2007 to 2018 (the period with the highest concentration of data in Nefin's analysis) is calculated for each industry (Nefin Avg Distr.)[1] and it is compared to the percentage participation of the observations of the sample (% Obs.) in the same industries.

Table 11: Data distribution by industry.

	Distribution of data by industry			
Nefin Industries	% Firms	% Obs.	Nefin Avg Distr.	Beta Nefin
Basic Products	7.4%	5.3%	8.4%	0.72
Manufacturing	12.5%	11.0%	10.2%	1.24
Construction	11.0%	11.2%	9.0%	1.59
Energy	15.4%	13.5%	17.5%	0.94
Consumer	17.6%	21.6%	31.4%	0.98
Other	36.0%	37.4%	23.6%	1.09
Total	136	902		

This Table presents the industries used by Nefin to estimate industry-based betas, the respective betas for each industry and the sample firms and observations distribution percentages by industry. Nefin Industries represent the 6 (six) industries of the economy used by Nefin to estimate industry-based betas. % Firms represent the percentage obtained from the distribution of all sample firms in each of these industries. % Obs. the percentage obtained from the distribution of all sample observations in each of these industries. Nefin Avg Distr. represents the equally weighted average of the annual distribution of all observations collected by Nefin in each of the industries between 2007 and 2018. Beta Nefin represents the specific betas estimated by Nefin for each of the 6 industries.

In general, except for the Consumer and Other industries, there are no significant differences between the distribution by industry percentage mean obtained from Nefin and from the sample. In the sample, the higher participation of the "Other" industry compared to the "Consumer" industry, as opposed to the Nefin distribution (Others 37.4% vs. Consumer 21.6% for the sample and Others 23.6% vs. Consumer 31.4% for Nefin), is justified by the fact that firms that perform in more than one activity or service (e.g., manufacturing, distribution and commerce) are classified as "Other" industry in order to incorporate a beta that reflects multiple activities (1.09), which turned out to be equivalent to the simple average of all industries (1.09). In practice, the greater distribution of observations in the "Other" industry (compared to the "Consumer" industry) observed in the sample ended up making it slightly more conservative by assigning a higher beta (Other: 1.09 vs. Consumer: 0.98) to a higher percentage of observations in the sample. Furthermore, there are no significant differences between the percentage obtained from the distribution of all firms in the sample in each industry (% Firms) and the percentage obtained from the distribution of all observations in the sample by industry (% Obs.), meaning that in general there is no excessive concentration of firms relative to the number of observations per industry. Overall, the sample appears to represent well, as in studies conducted by research centers, the distribution of data available in the market by industry.

Now we report the percentage distribution of all sample observations for each year (t) between 1998 and 2018. We observe that, due to a limitation in the availability of data in the Brazilian stock market, the distribution of data by year gradually increases as time passes. The use of observations from years prior to 2007 becomes relevant as we attempt to overcome this limitation of data availability by forming portfolios and estimating returns from pooled data independently of year (t) to which the observation refers.

[1] The estimation excluded financial firms, as it is done in the sample.

Table 12: Data distribution by year.

Distribution of data by year					
Year	%	Year	%	Year	%
1998	0.1%	2005	1.4%	2012	9.5%
1999	0.2%	2006	1.3%	2013	9.3%
2000	0.2%	2007	3.4%	2014	10.2%
2001	0.3%	2008	6.1%	2015	9.0%
2002	0.6%	2009	6.4%	2016	7.8%
2003	0.4%	2010	7.0%	2017	7.6%
2004	1.1%	2011	9.0%	2018	8.9%

This Table displays the percentage distribution of all sample observations for each year (t) between 1998 and 2018.

B.2 Modified sample used for estimating analyst forecast errors in the second part of our analysis.

Table 13: Descriptive Statistics.

Descriptive Statistics						
	Avg.	*Min.*	*Max.*	*Median*	*Stdv.*	*Total*
B(t-1)	44.13	0.10	4827.72	7.32	273.45	-
ROE(t-1)	0.1682	0.0015	0.9245	0.1443	0.1212	-
k	0.3198	0.0000	0.9998	0.2788	0.2593	-
B/P	1.13	0.01	24.51	0.76	1.43	-
ROA	0.0707	0.0002	0.3619	0.0593	0.0511	-
ME	9165.89	86.94	321204.32	2359.62	28464.80	-
Ke	0.1304	0.0764	0.2576	0.1273	0.0281	-
Beta Nefin	1.10	0.72	1.59	1.09	0.21	-
Obs.	-	-	-	-	-	889
N. Firms	-	-	-	-	-	136

Table values represent equally weighted average of all sample observations for all years (t) between 1998 and 2018. *B(t-1) Avg* represents the average book value per share of the firms in the sample at the end of December in year (t-1). *ROE Avg* is the average return on equity of the sample, calculated as net income divided by book value, both at the end of December in year (t-1). *k Avg* is the average dividend payout ratio of the firms in the sample, calculated as dividends paid to ordinary shareholders divided by net income, both at the end of December in year (t-1). As for ROE and *k*, for firms that reported negative net income in December of the year (t-1), a proxy was used for net income, equivalent to the percentage average of ROA (return on assets) of the entire sample (equivalent to 7%) multiplied by the total assets reported by the respective firm at the end of December in year (t-1). *B/P Avg* represents the average book-to-market ratio of the sample at the end of December of the year (t-1), where B represents book value per share and P represents the price of the firm's share, both at the end of December in year (t-1). ROA Avg is the average return on assets of the sample, calculated as net income divided by total assets, both at the end of December in year (t-1). *ME Avg* is the average market value of equity of the firms in the sample in millions of Brazilian reais, calculated as shares outstanding multiplied by the share price, both determined at the end of June in year (t). *Ke Avg* is the average cost of equity of the firms in the sample. *Beta Nefin Avg* is the average of the industry-based betas of the sample firms estimated by Nefin. *Min.* and *Max.* represent the minimum and maximum values observed for each parameter in the sample. *Median* and *Stdv.* represent the median and the standard deviation values for each parameter in the sample. N. Firms is the total number of firms that make up the sample. Obs. is the total number of observations in the sample.

Table 14: Data distribution by industry.

Data distribution by industry				
Nefin Industries	% Firms	% Obs.	Nefin Avg Distr.	Beta Nefin
Basic Products	7.4%	4.9%	8.4%	0.72
Manufacturing	12.5%	11.1%	10.2%	1.24
Construction	11.0%	11.4%	9.0%	1.59
Energy	15.4%	13.4%	17.5%	0.94
Consumer	17.6%	21.6%	31.4%	0.98
Other	36.0%	37.6%	23.6%	1.09
Total	136	889		

This Table presents the industries used by Nefin to estimate industry-based betas, the respective betas for each industry and the sample firms and observations distribution percentages by industry. Nefin Industries represent the 6 (six) industries of the economy used by Nefin to estimate industry-based betas. % Firms represent the percentage obtained from the distribution of all sample firms in each of these industries. % Obs. the percentage obtained from the distribution of all sample observations in each of these industries. Nefin Avg Distr. represents the equally weighted average of the annual distribution of all observations collected by Nefin in each of the industries between 2007 and 2018. Beta Nefin represents the specific betas estimated by Nefin for each of the 6 industries.

Table 15: Data distribution by year.

Data distribution by year					
Year	%	Year	%	Year	%
1998	0.1%	2005	1.5%	2012	9.7%
1999	0.2%	2006	1.3%	2013	9.2%
2000	0.2%	2007	3.5%	2014	10.0%
2001	0.3%	2008	6.2%	2015	9.0%
2002	0.6%	2009	6.4%	2016	7.9%
2003	0.4%	2010	7.0%	2017	7.5%
2004	1.1%	2011	9.0%	2018	8.8%

This Table displays the percentage distribution of all sample observations for each year (t) between 1998 and 2018.

Table 16: Spearman correlation.

Spearman correlation							
	B(t-1)	*V1h*	*V2h*	*V3h*	*V1f*	*V2f*	*V3f*
$\boldsymbol{P_{(t-1)}}$							
Corr	0.57	0.63	0.63	0.62	0.74	0.77	0.78
R^2	0.324	0.398	0.392	0.386	0.541	0.598	0.609
$\boldsymbol{P_{(t)}}$							
Corr	0.55	0.63	0.63	0.63	0.73	0.76	0.77
R^2	0.305	0.401	0.397	0.392	0.529	0.584	0.593
Obs.	889	889	889	889	889	889	889

This Table presents the correlation coefficients and respective R^2 (i) between the stock price at the end of December of year (t-1) ($P_{(t-1)}$) and at the end of June of year (t) ($P_{(t)}$) and B*(t-1)*, (ii) between $P_{(t-1)}$ and $P_{(t)}$ and each of the 3 (three) fundamental values estimated based on analyst earnings forecasts (*Vf*) and (iii) between $P_{(t-1)}$ and $P_{(t)}$ and each of the 3 (three) fundamental values estimated based on historical values (*Vh*), specifically ROE_t, for the sample comprising all years (t) from 1998 to 2018. *B(t-1)* represents the book value per share of the firms in the sample at the end of December of year (t-1). *V1f, V2f* and *V3f* and *V1h, V2h* and *V3h* represent fundamental values based on analyst forecasts and historical earnings, respectively. Obs. represents the total number of observations in the sample.

B.3 Analyzing the predictive power for stock returns: bi-dimensional portfolios

It becomes relevant to evaluate to what extent the performance of *V2f/P* is influenced by *B/P* or *ME* given the robust relation between B/P and the market value of equity of the firm (ME) and the average stock return (Fama and French, 1992). To this end, bi-dimensional analyses are performed. In Table 17, we show the results of the cumulative return of portfolios formed by the simultaneous application of two ex-ante firm criteria or variables (i) *V2f/P – B/P* and (ii) *V2f/P – ME* for 36 months, adopting a buy-and-hold strategy from the sample composed of all years (t) from 1998 to 2018. A total of 18 distinct bi-dimensional portfolios are formed by combining each of the tertiles of the *V2f/P* model with each of the tertiles of the *B/P* and *ME* models, respectively. When combining the tertiles of each variable, duplicate observations are excluded. To assess the statistical significance of the differences in the results of the combined portfolios T3 (*High ratio*) and T1 (*Low ratio*) (T3-T1) the Monte Carlo technique is used to simulate portfolios.

Table 17: Bi-dimensional Portfolios.

Table 17.a: V2f/P – B/P

			V2f/P		
	B/P	T1 *(Low Vf/P)*	T2	T3 *(High Vf/P)*	T3-T1
	T1 *(Low B/P)*	t1t1	t1t2	t1t3	
Ret36m		**0.357**	**0.455**	**0.528**	0.171**
Obs.		428	490	586	
	T2	t2t1	t2t2	t2t3	
Ret36m		**0.471**	**0.553**	**0.621**	0.150**
Obs.		514	481	507	
	T3 *(High B/P)*	t3t1	t3t2	t3t3	
Ret36m		**0.523**	**0.591**	**0.635**	*0.111*
Obs.		562	531	411	
	T3-T1	0.167**	0.136*	*0.107*	

Table 17.b: V2f/P – ME

			V2f/P		
	ME	T1 *(Low Vf/P)*	T2	T3 *(High Vf/P)*	T3-T1
	T1 *(Low ME)*	t1t1	t1t2	t1t3	
Ret36m		**0.559**	**0.590**	**0.672**	0.113
Obs.		543	495	466	
	T2	t2t1	t2t2	t2t3	
Ret36m		**0.360**	**0.491**	**0.511**	0.152**
Obs.		502	494	506	
	T3 *(High ME)*	t3t1	t3t2	t3t3	
Ret36m		**0.446**	**0.522**	**0.588**	*0.143***
Obs.		459	513	532	
	T3-T1	*-0.113*	-0.068	-0.084	

The Tables show the cumulative return of 18 different portfolios formed by the simultaneous application of 2 ex-ante criteria (i) *V2f/P – B/P* and (ii) *V2f/P – ME* using the entire sample, comprising all years (t) from 1998 to 2018. The returns of each portfolio represent equally weighted averages of all observations (pooled data) that make up each portfolio, regardless of the year (t) to which the observation refers, excluding duplicate observations. The return of each observation is cumulative and estimated from the beginning of July of year (t) in which the observation is collected to compose the sample, following a buy-and-hold strategy for the 36-month period. *V2f/P*, *B/P*, B, P, and *ME* have the same definitions as in Table 3. t1, t2, and t3 represent each of the tertiles into which the entire sample is independently sorted according to the *V2f/P*, *B/P* and *ME* criteria. t(i)t(i) is the reference of the tertiles used in the simultaneous application of the 2 criteria. T1, T2 and T3 are the combined portfolios formed from t(i)t(i). T3-T1 is the difference between the results obtained from the combined portfolios when applying the highest (High) and lowest (Low) tertiles of one criterion while the other is controlled. *Ret36m* is the return of each combined portfolio based on a buy-and-hold strategy for the 36-month period. Obs. represents the total number of observations in the sample that make up each combined portfolio. The statistical significance of the differences is based on Monte Carlo simulation tests. Simulated portfolios are formed by randomly allocating (without replacement) the entire population of the sample into tertiles (t1, t2 and t3) of one variable and combined with tertiles of the other variable whose membership remained constant. The notation ***, **, * indicates that the observed difference between T3-T1 is significantly different from the differences of the simulated portfolios at the 1%, 5% and 10% levels, respectively (two-tailed).

To make it easier visualizing the intended comparison, the results are identified in pairs with underlined, italicized, and normal font. As an example, in Table 17.a, the underlined font represents pairs of results in which the lowest ratio tertile (t1) of one criterion is controlled while the other varies across t1 and t3. The same occurs in Table 17.b, considering the presence of a small firm effect in the ME results represented by negative signs. By observing the results of the combined portfolios *V2f/P-B/P* (Table 17.a) we can conclude that there is a robust relation between the *V2f/P* and *B/P* models in predicting returns, since, while controlling both the highest and lowest tertiles, the T3-T1 differences are similar: 11.1% when controlling the *B/P* criterion and 10.7% when controlling the *V2f/P* criterion (significance close to 10%, but results not statistically significant), and 17.1% when controlling the *B/P* criterion vs. 16.7% when controlling *V2f/P* (both statistically significant at 5%). In other words, *V2f/P* shows only a slight superiority by allowing its tertiles to vary across the combined portfolios, corroborating the existence of a strong relation between the models in the sample and similar performances in a buy-and-hold strategy for 36 months, evidenced in Section 3.3.2 (Tables 3.a and 3.c). As to the results of the combined portfolios using both *V2f/P* and *ME* criteria (Table 17.b), we find that, when controlling for either the highest or the lowest tertile of one criterion while the other varies, *V2f/P* generally shows higher T3-T1 differences compared to *ME* (the T3-T1 differences 0.113 and -0.113, although not statistically significant, obtained p-values equivalent to 0.1083 and 0.1032, respectively). Nonetheless, these results do not allow us to disregard the influence of the market equity of firms on *V2f/P* in predicting returns. Our results of the bi-dimensional analyses of the models thus support the existence of a strong and resilient relation between *B/P* and *V2f/P* in predicting returns for 36 months, and the influence of the firm's market equity on the explanatory power of *V2f/P*.