Bank i Kredyt 57(5), 2026, 589-616
DOI: 10.5604/01.3001.0055.9229

BANK I KREDYT
BANK & CREDIT

Assessment of the accuracy of bank valuations prepared using selected valuation models in stock-market recommendations on the Polish capital market

Michał Radke*, Patryk Ciszewski#

Submitted: 10 December 2025. Accepted: 1 June 2026.

Suggested citation: Radke M., Ciszewski P. (2026), Assessment of the accuracy of bank valuations prepared using selected valuation models in stock-market recommendations on the Polish capital market, *Bank i Kredyt*, 57(5), 589–616, DOI: 10.5604/01.3001.0055.9229.

Abstract

The aim of the article is to assess the accuracy of bank valuations contained in stock-market recommendations on the Polish capital market and prepared using selected valuation models. The analysis covers 364 recommendations concerning 11 banks issued between 25 October 2005 and 13 August 2021. Valuation accuracy is defined as the percentage deviation of the target price from the closing share price observed on the recommendation's maturity date, consistent with the investment horizon specified in the report (3, 6, 9 or 12 months). A three-level classification is applied: 0–5% as accurate, 5–15% as acceptable, and above 15% as unacceptable. A one-sample t-test against the 15% benchmark is used to assess the average level of valuation error.

The study identifies four valuation models used in brokerage practice: the ROE–P/BV model, the residual income model (RIM), relative valuation based on multiples, and the dividend discount model (DDM). Relative valuation based on multiples is the most frequently used technique (71.15%), followed by the ROE–P/BV model (61.81%) and RIM (42.31%), whereas DDM appears only marginally (1.37%). The mean valuation errors for the ROE–P/BV model, RIM, and relative valuation are all close to 39–40%, while the highest mean error is recorded for DDM (50.68%). These findings imply that target prices published in brokerage recommendations should be interpreted cautiously, as valuation estimates subject to substantial uncertainty. At the same time, the interpretation of the results should consider the operational nature of the 15% benchmark and the very small number of DDM-based observations, which limits the strength of inference for that model.

Keywords: valuation, banks, stock-market recommendation, accuracy

JEL: G24, G14, G12

* Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Finansów Korporacji; e-mail: michal.radke@eksoc.uni.lodz.pl.
# Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny; e-mail: patryk.ciszewski@edu.uni.lodz.pl.

1. Introduction

The need for business valuation in Poland arose with the change in the socio-political system in 1990 and the introduction of the rules of a market economy (Zarzecki 2013, p. 534). The purchase and sale of enterprises, mergers and acquisitions, new share issues, and corporate break-ups are only the most important among the many situations in which a valuation is advisable (Zarzecki 2008, p. 105). Both theory and practice show that business valuation is a complex process for at least two reasons. First, due to the object of interest – the enterprise – which is a structure of varying degrees of complexity, a mechanism characterised by numerous internal interdependencies and multiple links with its environment. Second, business valuation draws on knowledge from economics, finance, and management, as well as, indirectly, other social sciences. Such a broad body of knowledge, with a diverse conceptual apparatus, hinders comprehension and generates a number of problems that translate into valuation practice (Szymański 2012, p. 435). Although, in both theory and practice, there is general agreement as to the approaches and specific models used in business valuation, many issues have not yet been conclusively resolved and are regarded as particularly difficult and controversial. Moreover, in Poland business valuation is not regulated by law. It is assumed that the market and generally accepted 'good practice' act as the regulator. There are, of course, advantages to such a solution (free competition), however, one may also point to certain risks, manifested above all in the low quality of valuations. This undermines the reliability of the results obtained, which in turn may lead to irrational and often simply erroneous decisions (Zarzecki 2011, p. 534).

The aim of this article is to examine the accuracy of bank valuations embedded in stock-market recommendations on the Polish capital market and prepared using selected valuation models. The following research hypotheses were formulated:

H1: The mean absolute error of bank valuations obtained using the RIM model exceeds the level regarded as acceptable (15%), which indicates low accuracy of those valuations.

H2: The mean absolute error of bank valuations obtained using the DDM model exceeds the level regarded as acceptable (15%), which indicates low accuracy of those valuations.

H3: The mean absolute error of bank valuations obtained using the relative valuation based on multiples exceeds the level regarded as acceptable (15%), which indicates low accuracy of those valuations.

H4: The mean absolute error of bank valuations obtained using the ROE–P/BV model exceeds the level regarded as acceptable (15%), which indicates low accuracy of those valuations.

2. Previous research

Studies on the valuation models used in brokerage recommendations for the entire Polish market were conducted, among others, by Głębocki et al. (2011), who found that in 100% of cases the discounted cash flow model was applied, in 98% the relative valuation based on multiples, 0.5% of the sample relied on the adjusted net asset model and the Wilcox-Gambler model, and in 0.89% of the sample the asset value model was used. Similar conclusions follow from the study by Adamczyk and Zbroszczyk (2017), who examined recommendations issued on the Warsaw Stock Exchange in 2016. Their research shows that two valuation models were most commonly applied (93.2%), and the most popular combination

was that of the market and income approaches (86.7%). Likewise, in 2017 Prusak demonstrated in his research that the DCF and relative valuation based on multiples were the most popular, being used in 91.25% and 89.82% of recommendations, respectively. Next in line was the P/B–ROE model, applied in 7.86% of the analysed sample, the net asset model in 1.79%, and the residual income model in 0.89% of recommendations. The liquidation approach was used in 0.71% of the recommendations examined, while in 0.18% of the recommendations subjected to analysis the sum-of-the-parts (SOTP) model and the Wilcox-Gambler model were applied. Among international studies, the results were similar. Imam, Barker and Clubb (2008) observed that the most popular model was DCF, applied in 50% of the recommendations analysed. Next, the most frequently used were the P/E and EV/EBITDA multiple models, employed in 45.92% and 25.51% of the sample, respectively. This was followed by P/BV, DDM and CFROI, applied in 11.22%, 10.20% and 9.18% of recommendations. The EV/Sales multiple was used in 7.14% of the sample. The PEG and EVA models were each employed in 3.06% of the sample. The Dividend Yield and P/CF models were used in 2.04% of recommendations, while the P/Sales model appeared in only 1.02% of recommendations. Similar conclusions were reached by Erkilet, Janke and Kasperzak (2022), who found that sell-side analysts more frequently used the market approach (39%; 340 observations) when formulating target prices, followed by the income approach (33%; 282 observations) and hybrid approaches (28%; 245 observations).

In the case of banks, Gianfrate and Vincenzi (2013) analysed 482 equity research reports on Italian listed banks published between 2005 and 2011, as well as the valuation models used by financial analysts to determine target prices, dividing them into two main categories: fundamental valuation (multi-period valuation) and relative valuation (valuation ratios). It was found that the dominant approach was relative valuation methodology, used in 70% of the data examined, using specific techniques such as price-to-book ratio, price-to-earnings ratio, and value map methodology. In contrast, fundamental valuation techniques such as discounted cash flow (DCF), dividend discount model (DDM), economic value added (EVA), and residual income model accounted for the remaining 30% of the data covered by the study.

The study by Ho et al. (2023) employed manual content analysis to examine a sample of 2,263 equity research reports on 23 large North American and European banks published between 2014 and 2017. Analysts faced the primary decision of valuing a bank either as a stand-alone entity (single entity approach) or using the sum-of-the-parts (SOTP) approach. Overall, the stand-alone entity approach was the most frequent choice, used in 1,764 reports (78%), while the SOTP approach was utilized in 499 reports (22%). This preference varied by region: analysts covering North American banks were significantly more likely to use the stand-alone entity approach (e.g. 92% of US reports and 100% of Canadian reports) compared to analysts covering European banks, who preferred the SOTP model in 63% of their reports. Regarding specific models, equity analysts predominantly use single-period multiple models to value banks, irrespective of whether they followed the SOTP or stand-alone entity approach, with the most popular multiples being price/earnings (P/E) and price/book-value (P/BV). The implementation of multi-period discounted cash flow models is mostly limited; these models include the dividend discount model (DDM) and residual income valuation model (RIV). Furthermore, when employing the SOTP framework, analysts sometimes use a hybrid approach, applying different valuation models for different segments of the bank, such as combining single-period multiples with discounted cash flow techniques.

The accuracy of investment recommendations has been widely examined in both international and domestic research. One of the earliest studies in this field was conducted by Cowles (1933), who concluded, based on historical data, that analysts' forecasting performance was no better than chance. Bradshaw (2002), analysing 103 reports published in 1998–1999, found that analysts often used target prices merely to justify their recommendations rather than as independent forecasts of a security's value. A comprehensive example is the study conducted by Bradshaw, Brown and Huang (2012), which analysed US market data from 2000–2009. The authors found that only 38% of the recommendations reached the indicated target price within a 12-month investment horizon. The remaining 62% failed to achieve the forecast price level, suggesting significant deviations caused by excessive optimism or pessimism in analysts' forecasts. Asquith, Mikhail and Au (2005) examined 1,126 analytical reports from 1997–1999 in the US. For recommendations based on the multiplier model, the target price was reached in 48.4% to 55.1% of cases. For those using the discounted cash flow (DCF) model, accuracy was 52.3%, indicating moderate effectiveness of these valuation techniques.

In Europe, Bolliger (2004) analysed data from 14 capital markets and found that older forecasts exhibited significantly lower accuracy than more recent ones, suggesting the usefulness of applying time-based weights when assessing the predictive value of analyst reports. In Poland, Biedrzyński (2008) studied 516 recommendations issued between 1 January 2006 and 1 February 2008, assessing forecast accuracy over two-time frames: after nine months and at the end of the observation period. The average deviation between market and target prices was 44.77% and 60%, respectively. In the nine-month analysis, 271 cases were undervalued and 120 overvalued, indicating a predominance of pessimistic forecasts. By the end of the study period, only 36.5% of recommendations had reached the target price, while 63.5% were inaccurate, confirming significant forecast deviations. Prusak (2015), analysing 470 recommendations with target prices over 6-, 9- and 12-month horizons, used a 10% deviation threshold as the criterion for forecast accuracy. The accuracy rates were 30% (6 months), 68% (9 months) and 21.92% (12 months), confirming a negative relationship between forecast horizon and accuracy. The financial daily *Parkiet* analysed 71 reports concerning selected WSE-listed companies and found that forecasts issued during bull markets were significantly overestimated relative to actual outcomes. Kowalke (2015) focused on the usefulness of recommendations in terms of achieving target prices. In his study, only 30% of recommendations deviated by less than 10% from projected levels, indicating a relatively low predictive accuracy of equity analysis. Pastusiak (2016) examined recommendations for ten WSE-listed companies to determine whether excessive optimism was present among analysts. His study, covering 2000–2013 with a 250-session recommendation horizon, found that valuations were on average overestimated by approximately 30%, indicating a bias towards optimistic predictions that was not reflected in market prices. In a subsequent study, Pastusiak (2017) analysed over 10,500 recommendations between 2000 and 2014. He found that buy recommendations were, on average, overestimated by 29%, while sell recommendations were underestimated by 12%. Dąbrowski (2013) analysed recommendations for five blue-chip companies between 2007 and 2012. He found that analysts were overly optimistic, with 70% of recommendations being 'buy' and only 30% 'sell'. Only 44.63% of the recommendations reached the target price within the planned time frame. Radke (2023) conducted a study based on a sample of 3,245 stock-market recommendations issued between 2005 and 2018. His findings show that only 12.2% of the recommendations were accurate.

The accuracy of recommendations for the banking sector has also been examined, although such studies are considerably less numerous. The study conducted by Gross (2006) aimed to determine

which valuation models are applicable to banks and how standard models should be adapted to their institutional specificity. The study was conducted through a synthesis of the literature on bank valuation and value-based management. The results suggest that although the fundamental principles of valuation are applicable, it is necessary to use the equity approach in all models. The discounted cash flow (DCF) and residual income (RI) models are identified as the superior models for measuring shareholder value. Six key adjustments to these models were identified, including the explicit consideration of required capital retention and the adjustment of risk provisions. The residual income (RI) model is preferred in most cases, because it is easier to apply in a banking context and better supports value-based management.

The study by Mehran and Thakor (2011) demonstrates that bank capital has a positive impact on the total value of a bank in cross-sectional terms, as well as on individual components of this value. Their results strongly support the main theoretical predictions, indicating that the total value of the target bank, the purchase price, goodwill (in the acquisition) and net present value (NPV) for shareholders are all positively correlated with the target bank's equity capital. Gounder and Venkateshwarlu (2017) examined and compared the accuracy of the residual income valuation (RIV) model and a relative valuation model in order to determine the most appropriate model for valuing banks from a shareholder perspective. The research sample comprised 40 major commercial banks in India between 2001 and 2015. The empirical analysis showed that the residual income valuation (RIV) model is a more appropriate and accurate model for valuing banks, as the intrinsic value of a bank's shares determined using RIV showed a strong positive correlation with the market price of the shares, explaining approximately 70% of market price volatility. Consequently, the authors concluded that the RIV model is superior to relative valuation in measuring shareholder value and provides more reliable value estimates. Massari et al. (2018) assessed the historical development and performance of relative valuation of banks in the United States (US) and Europe between 1990 and 2017. The researchers measured the accuracy of various multiples, such as those based on book value, revenues, historical and projected earnings, and deposits, by comparing the 'theoretical' valuation of banks obtained from multiples with their actual market prices, treating those that generated the lowest valuation errors as the most accurate. The results showed that multiples based on forward earnings (forward P/E), especially those projected two years ahead, achieve the highest accuracy, and that valuation precision is significantly higher for US banks than for European banks. Leong et al. (2023) addressed the lack of empirical evidence regarding bank share pricing theory and the ability of theory-based models to forecast share prices accurately over future time horizons. In their study, they tested how the four most popular bank valuation models explained future prices using 140 observations from Malaysia. The study found that each model significantly explained bank share price volatility, with the best-performing models based on the P/E ratio, the dividend discount model, the residual income model, and free cash flow to equity (FCFE), respectively. Moyo and Obadire (2024) examined bank equity valuation models used in practice by equity analysts in South Africa, Nigeria and Kenya. The study sample included a total of 284 reports issued between 2018 and 2023. They found that the dominant valuation approaches were the warranted equity model (WEM) and the excess return model. WEM was widely used, often in conjunction with tangible net asset value (tNAV) and return on tangible equity (ROTE). More recent studies also broaden the discussion by linking bank valuation to specific drivers of market value. Martínez, Jiménez and Matas (2024), analysing 46 European banking groups between 2005 and 2020, show that profit-related equity components, especially yearly net

income and other comprehensive income, exert a significant positive influence on banks' price-to-book ratios, whereas investor contributions aimed mainly at strengthening solvency have a much weaker impact. A similar perspective is offered by Kwiatkowska (2024), whose literature review indicates that bank market value is shaped primarily by profitability, capital strength, loan portfolio quality, market power, and selected risk-related variables. These findings are important for bank valuation because they suggest that the explanatory power of accounting and market-based models depends not only on model choice, but also on how bank-specific value drivers are captured. Complementary evidence is provided by recent studies from non-European markets. Chen (2023), comparing discount-based and multiples-based valuation approaches for major Chinese banks, concludes that under relatively stable market conditions the discrepancies between model-based estimates and observed market prices may remain limited. Adipramana, Dirgantari and Masharyono (2025), in turn, demonstrate for Indonesian listed banks that earnings-related indicators and price-to-book value are significant determinants of stock-price variation. Although these studies do not directly evaluate target price accuracy in brokerage reports, they support the view that bank valuation outcomes are highly sensitive to profitability, book-value-related indicators, and the broader market environment.

3. Methodology

In order to achieve the research objective relating to the accuracy of bank valuations embedded in stock-market recommendations and prepared using selected valuation models, sell-side recommendations issued by Polish and foreign brokerage houses for companies in the banking sector were examined. Each recommendation included in the analysis was required to contain information on:

– the recommended company,
– the date of issue of the recommendation,
– the investment horizon for which it was issued,
– the target price together with an indication of the valuation model used.

The list of published stock exchange recommendations in the assumed research period was obtained via specialist industry websites:

– Bankier.pl (https://www.bankier.pl/gielda/rekomendacje),
– StockWatch.pl (https://www.stockwatch.pl/rekomendacje/).

Due to the need to verify the accuracy of valuations presented in recommendations – understood as the difference between the target price and the market price on the recommendation's maturity date – it was necessary to obtain share-price data for each of the analysed companies on the maturity date of the recommendation, corresponding to the period for which it had been issued. Quotations were obtained using the Stooq (www.stooq.pl) and Warsaw Stock Exchange (WSE) (www.gpw.pl) services. On this basis, the deviation of the target price from the market price observed on the recommendation's maturity date was determined, which was essential for verifying the hypotheses relating to the accuracy of bank valuations prepared using the examined models.

Given the nature of the analyses performed, it was also necessary to remove certain observations that could potentially distort the final conclusions. First and foremost, recommendations for which the investment horizon was not specified were excluded. This was necessary because, in such cases, it was not possible to determine the recommendation's maturity date, which is required to calculate the deviation of the target price from the market price at the end of the analysed period.

The research sample constructed according to the above criteria comprised 364 recommendations prepared for 11 banks.

In considering the presence of excessive analyst optimism in the valuation of enterprises based on stock exchange reports, the selection of the sample was guided by the aim of including the broadest possible dataset to construct a database for the purposes of this study, while maintaining reliability and internal consistency in relation to the phenomenon under investigation. The study of the Polish stock market is naturally constrained historically by the date of the establishment of the WSE, namely 16 April 1991. However, the first years of its operation differed significantly from the market structure observed today (from the perspective of the present study), both in terms of the number of listed companies, the number of institutions operating on the market, the frequency of valuations and recommendations, the number of investors and the extent of regulations governing trading in instruments on this market.

Another issue that had to be considered in the data collection process is that the legal regulations concerning stock exchange recommendations did not come into force in their entirety at the time the WSE was established in 1991. The document analysed here, namely the stock exchange recommendation, was defined in detail in the Regulation of 19 October 2005 on information constituting recommendations concerning financial instruments, their issuers or providers (Journal of Laws 2005 No. 206, item 1715). For this reason, the present study uses only recommendations published after the entry into force of the Regulation of 19 October 2005 on information constituting recommendations concerning financial instruments, their issuers or providers, that is, after 24 October 2005.

The final research sample covers the period from 25 October 2005 to 13 August 2021, i.e. almost 16 years. The time span of the study is one of the longest among all previous studies described in the literature dealing with stock exchange recommendations on the Polish capital market. Consequently, the conclusions drawn on this basis should provide a more complete picture of the capital market in the analysed area than in earlier studies.

It should also be noted that this period includes cyclical economic and stock market fluctuations, which means that the phenomenon and variables under study are tested both in periods of falling and rising share prices, as well as during phases of weaker and stronger overall economic conditions.

The period covered by the analysis between October 2005 and August 2021 comprises 190 months. Relating this to several basic theories of business and stock market cycles, it should be noted that, according to the research of Wośko (2005) conducted directly on the Polish market, the business cycle in Polish conditions lasts approximately 52.7 months, while a full stock market cycle lasts around 31.2 months. This implies that the adopted time span of the study covers almost four full business cycles and more than six stock market cycles. It should also be noted that the Kitchin cycle lasts around 40 months (Godłów-Legiędź 2013, p. 582), which means that the period examined encompasses almost five full Kitchin cycles. In relation to Juglar cycles, which last between 8 and 10 years (Szarzec 2019, p. 370), the research sample covers more than one and a half cycles.

To summarise the assumptions adopted with respect to the research period, it should be emphasised that it refers to a period subject to appropriate legal regulations, allows for the collection of adequate information in the context of the research topic, and covers the full range of economic and market phases. This makes it possible to depict the phenomenon in a comprehensive manner, without omitting specific periods.

The recommendations included in the sample were issued for one of four investment horizons: 3, 6, 9 or 12 months. Valuation accuracy was defined as the percentage deviation of the target price from the closing share price observed on the recommendation's maturity date, corresponding to the investment horizon specified in the given recommendation. The deviation is determined according to the following formula (Prusak 2015; Brycz, Dudycz, Włodarczyk 2021):

$$O_k = \left| \frac{CD - C_k}{C_k} \right| \cdot 100\%$$

where:

O_k – deviation of the target price from the market price on the recommendation's maturity date, corresponding to the investment horizon specified in the recommendation,

CD – target price specified in the report,

C_k – market closing price on the recommendation's maturity date, corresponding to the investment horizon specified in the recommendation.

It should be emphasised that bank valuation is not deterministic in nature. A valuation outcome is a conditional estimate of value that depends on the underlying modelling assumptions and input parameters, including expected profitability, the cost of risk, growth prospects, capital requirements, asset quality, and the regulatory environment. Consequently, the target price reported in a brokerage recommendation should be treated as a point estimate subject to uncertainty rather than as an exact prediction of the future market price of the share. For this reason, the present study focuses on an ex post assessment of deviations between target prices and observed market prices over specified investment horizons, rather than on a formal modelling of the stochastic process of bank value.

Subjectively determined levels of valuation accuracy assessment are illustrated in the table below.

The subjective valuation-accuracy scale was developed on the basis of earlier studies by Prusak (2015), Majewski and Majewska (2017), and Klepacki (2019) concerning the accuracy of target-price forecasts in stock exchange recommendations. Majewski and Majewska (2017) classifies recommendations as accurate when the deviation amounts to ±5%. Klepacki (2019) regards forecasts as very accurate when the deviation is less than or equal to 3%; accurate when the deviation lies between 3% and 5%; acceptable when it ranges from 5 to 10%; and unacceptable when it exceeds 10%. Prusak (2015), in turn, adopts the following scale: 0–5% – very good, 5–10% – good, 10–15% – average, 15–20% – poor, and above 20% – very poor.

Accordingly, the 15% threshold is treated in this study not as an objectively fixed universal boundary, but as an operational benchmark adopted for the purposes of hypothesis testing and comparative interpretation, grounded in prior studies on target-price accuracy in brokerage recommendations.

In order to achieve the main research objective and verify the research hypotheses, tests of single-mean hypotheses were conducted using the Student's t-test. The t-distribution is a theoretical distribution of the relative frequencies of all values of $\overline{X}$ transformed into t-values that would be obtained randomly if an infinite number of samples of a given size were drawn from the population. The formula defining t for testing a single-mean hypothesis is as follows (King, Minium 2020, p. 313):

$$t = \frac{\bar{X} - \mu_x}{s_{\bar{x}}}$$

where:

$\bar{X}$ – sample mean,

μ_x – hypothetical population mean,

$s_{\bar{x}}$ – estimate of the standard error of the mean.

The significance test for the mean is used to verify the hypothesis that the average value of a characteristic in the general population is equal to a hypothetical value (a specific number). Verification is carried out using a simple *n*-element sample. The null hypothesis takes the form H0: $\mu = \mu_x$ where μ_x is the hypothetical value. The alternative hypothesis may take one of three forms: H1: (I) $\mu \neq \mu_x$; (II) $\mu > \mu_x$; (III) $\mu < \mu_x$ (Malska, Koziorowska 2015, p. 324).

The Student's t-test was performed using quantitative data on valuation accuracy, defined as the deviation of the target price from the market price observed on the recommendation's maturity date, consistent with the investment horizon specified in the recommendation. In accordance with the operational benchmark adopted on the basis of the literature, a deviation of 15% was treated as the threshold above which valuations were classified as inaccurate.

A fundamental research issue is to determine how large the difference between the hypothetical value μ_x and the true value μ_x must be. In general, it is convenient to express the magnitude of the difference between μ_{hip} and μ_{true} in terms of the number of standard deviations of the measure used. Such an index is called the effect size and is denoted by the symbol *d* (King, Minium 2020, p. 338):

$$d = \frac{\mu_{hip} - \mu_p}{\sigma}$$

For practical purposes, this formula is transformed into the following form (King, Minium 2020, p. 338):

$$d = \frac{\bar{X} - \mu_{true}}{\sigma_x}$$

Cohen's *d* expresses the difference in relation to the standard deviation of the set of observations rather than to the standard error. If the value of σ_x is unknown, it can be estimated using *s*. The resulting estimate of *d* is called *g* (Hedges' correction) (King, Minium 2020, p. 338):

The effect size proposed by Cohen is defined as follows: 0.2 – small; 0.5 – medium; 0.8 – large (King, Minium 2020, p. 338).

The formal verification of hypotheses H1–H4 concerns the mean absolute valuation error only. Measures of dispersion and distributional shape, such as the standard deviation, coefficient of variation, skewness, and kurtosis, are reported as supplementary descriptive statistics intended to provide additional insight into the heterogeneity of valuation errors, rather than as separate hypothesis-tested dimensions.

4. Own research

In the sample under study, bank valuations were prepared using four models: the residual income model (RIM), the dividend discount model (DDM), the relative valuation based on multiples, and the ROE–P/BV model. The most frequently used model was the relative valuation based on multiples, which was applied in 259 recommendations, i.e. 71.15%. The second most commonly used model was ROE–P/BV, employed in 225 recommendations, i.e. 61.81%. The RIM model was used in 154 recommendations, representing 4% of all recommendations. The least frequently applied model was DDM, which was used in only five recommendations, accounting for 1.37% of all recommendations. These figures are illustrated Figure 1.

For valuations obtained using the ROE–P/BV model, only 10.67% of recommendations were classified as accurate, while a further 15.56% fell into the acceptable range. The vast majority, 73.78%, were assessed as unacceptable, which indicates that, in the analysed sample, valuations prepared using the ROE–P/BV model exhibited relatively low predictive accuracy in estimating banks' target prices. These figures are illustrated in Figure 2.

The mean valuation error using the ROE–P/BV model in the study sample (N = 225) was M = 39.33%, with high variability (SD = 40.01%, SE = 2.67%). Values ranged from 0.43% to 310.80% (range = 310.38%). The distribution was clearly positively skewed (skewness = 2.724; SE = 0.162) and strongly leptokurtic (kurtosis = 11.558; SE = 0.323), suggesting the presence of high extreme values. The coefficient of variation was CV = 101.73%, indicating very substantial dispersion. This interpretation is reinforced by the robust location measures: the median valuation error was 29.03%, and the 5% trimmed mean was 34.58%, both below the arithmetic mean, which indicates that the mean-based result was influenced by extreme observations. Descriptive statistics are presented in Table 2.

In a one-sample t-test, the mean valuation error for valuations prepared using the ROE–P/BV model was compared with a reference value of 15%. The results indicated that the mean was significantly higher than the test value, $t(224) = 9.112$, $p < 0.001$ (two-tailed). The mean difference was 24.33%, with a 95% confidence interval for the difference of [19.07%, 29.59%]. The observed difference was characterised by a moderate effect, Cohen's $d = 0.608$ (95% CI: [0.465, 0.749]). After correcting for small-sample bias, a very similar estimate was obtained: Hedges' $g = 0.606$ (95% CI: [0.463, 0.747]). The results are presented in Tables 3 and 4.

For valuations obtained using the RIM model, only 7.79% of recommendations were classified as accurate, while a further 14.94% fell into the acceptable range. The vast majority, 77.27%, were assessed as unacceptable, indicating that, in the analysed sample, valuations prepared using the RIM model exhibited very low predictive accuracy in estimating banks' target prices. These figures are illustrated in Figure 3.

The mean valuation error for the RIM model in the analysed subsample (N = 154) was M = 40.37%, with very high variability (SD = 53.80%, SE = 4.34%). Values ranged from 0.80% to 441.48% (range = 440.68%). The distribution was strongly positively skewed (skewness = 5.142; SE = 0.195) and extremely leptokurtic (kurtosis = 32.374; SE = 0.389), indicating the presence of very high extreme observations. The coefficient of variation was high (CV = 133.28%), confirming substantial dispersion in the results. This interpretation is reinforced by the robust location measures: the median valuation error was 30.39%, and the 5% trimmed mean was 32.71%, both below the arithmetic mean, which indicates a substantial influence of extreme observations on the mean-based result. Descriptive statistics are presented in Table 5.

In a one-sample t-test, the mean valuation error for valuations prepared using the RIM model was compared with a reference value of 15%. The results indicated that the mean was significantly higher than the test value, $t(153) = 5.851$, $p < 0.001$ (two-tailed). The mean difference was 25.366%, with a 95% confidence interval for the difference of [16.801%, 33.930%]. The observed difference was characterised by a moderate effect, Cohen's $d = 0.471$ (95% CI: [0.304, 0.637]). After correcting for small-sample bias, a very similar estimate was obtained: Hedges' $g = 0.469$ (95% CI: [0.303, 0.634]). The results are presented in Tables 6 and 7.

For valuations obtained using the relative valuation based on multiples, 11.29% of recommendations were classified as accurate and a further 11.29% as acceptable. The remaining 77.42% were assessed as unacceptable. This distribution indicates that, in the analysed sample, valuations prepared using the relative valuation based on multiples also displayed low predictive accuracy in estimating banks' target prices.

The mean valuation error for relative valuation based on multiples in the analysed sample ($N = 259$) was $M = 39.34\%$, with high variability ($SD = 38.56\%$, $SE = 2.40\%$). Values ranged from 0.02% to 246.78% (range = 246.77%). The distribution was clearly positively skewed (skewness = 2.287; SE = 0.151) and leptokurtic (kurtosis = 6.956; SE = 0.302), suggesting the presence of high extreme observations. The coefficient of variation was $CV = 98.03\%$, indicating very substantial dispersion in the results. This interpretation is reinforced by the robust location measures: the median valuation error was 29.39%, and the 5% trimmed mean was 34.45%, both below the arithmetic mean, which indicates that the mean-based result was influenced by extreme observations. Descriptive statistics are presented in Table 8.

In a one-sample t-test, the mean valuation error for valuations prepared using relative valuation based on multiples was compared with a reference value of 15%. The results indicated that the mean was significantly higher than the test value, $t(258) = 10.157$, $p < 0.001$ (two-tailed). The mean difference was 24.336%, with a 95% confidence interval for the difference of [19.618%, 29.055%]. The observed difference was characterised by a moderate effect, Cohen's $d = 0.631$ (95% CI: [0.497, 0.764]). After correcting for small-sample bias, a very similar estimate was obtained: Hedges' $g = 0.629$ (95% CI: [0.496, 0.762]). The results are presented in Tables 9 and 10.

For valuations obtained using the DDM, none of the recommendations were classified as accurate or even acceptable; all of them (100%) fell into the unacceptable category. This indicates that, in the analysed sample, valuations prepared using the DDM exhibited the lowest predictive accuracy of all the examined valuation models in estimating banks' target prices.

The mean valuation error for the DDM model in the analysed very small subsample ($N = 5$) was $M = 50.68\%$, with moderately high variability ($SD = 29.39\%$, $SE = 13.14\%$). Values ranged from 21.43% to 89.42% (range = 67.99%). The distributional shape indices suggest a slight positive skew (skewness = 0.330; SE = 0.913) and a platykurtic (flattened) distribution (kurtosis = –1.872; SE = 2.000); however, given $N = 5$, these estimates should be interpreted with caution due to substantial uncertainty. This interpretation is consistent with the robust location measures: the median valuation error was 49.30%, and the 5% trimmed mean was 50.15%, both very close to the arithmetic mean. This indicates that, despite the small sample size, the mean-based result does not appear to be strongly distorted by extreme observations. Descriptive statistics are presented in Table 11.

Due to the small sample size, the t-test was not conducted.

5. Conclusions

This study examines the accuracy of bank valuations embedded in stock-market recommendations issued on the Polish capital market. The empirical evidence indicates that these valuations exhibit limited predictive accuracy. Across the three formally tested valuation approaches, the average target-price error exceeds the 15% benchmark adopted in the study as the threshold of acceptable accuracy. For the DDM model, the mean error was also above this benchmark; however, due to the very small number of observations, this result should be treated as descriptive rather than as formally hypothesis-tested evidence. Accordingly, target prices reported in brokerage recommendations should be interpreted with caution: rather than representing precise forecasts of the future market price of the share, they should be viewed as conditional valuation estimates subject to substantial uncertainty.

A further important finding is that the results do not support the superiority of any single valuation model. The ROE–P/BV model, the residual income model (RIM), and relative valuation based on multiples generate broadly similar average errors, while the dividend discount model (DDM) records the highest mean error; however, the very limited number of DDM-based observations prevents strong comparative conclusions in this respect. Taken together, the results suggest that the key challenge lies less in the selection of one universally superior model and more in the broader difficulty of valuing banks in an environment shaped by substantial uncertainty, regulatory constraints, and sector-specific risk.

The distributional properties of the valuation errors provide additional insight into this problem. The error distributions are characterised by substantial dispersion, positive skewness, and the presence of extreme observations, indicating that valuation outcomes are uneven and occasionally subject to large deviations from realised market prices. These distributional characteristics should be interpreted as descriptive features of the empirical error structure, complementing the hypothesis--based assessment focused on average valuation error. This pattern is consistent with the notion that bank valuation is particularly sensitive to shifts in macroeconomic conditions, the risk environment, supervisory expectations, and capital-market sentiment. In this sense, the limited accuracy documented in the study should not be interpreted solely as evidence of weak model performance, but also as a reflection of the intrinsic complexity of valuing banking institutions.

These findings have several implications for valuation practice. First, improving the quality of bank valuation is unlikely to be achieved through the mechanical preference for one model over another. Greater gains may instead come from refining model implementation. In particular, analysts may benefit from replacing point target prices with valuation ranges and explicit scenario analysis, including baseline, optimistic, and downside cases. Second, valuation frameworks should more systematically incorporate bank-specific drivers of value, such as capital adequacy, the cost of risk, asset quality, profitability, and changes in the regulatory environment. Third, the results support the use of regular recalibration of valuation assumptions to both business-cycle and market-cycle conditions, combined with more explicit sensitivity analysis for key parameters. Such adjustments may improve not only analytical discipline, but also the transparency and interpretability of valuation outputs.

The study also carries practical implications for market participants. For investors, the results suggest that brokerage target prices should be treated primarily as one input into the investment decision-making process rather than as a sufficient standalone signal. For analysts, the findings point to the importance of communicating assumptions, uncertainty, and parameter sensitivity more explicitly, especially in the context of bank valuation, where relatively small changes in inputs may materially alter valuation outcomes.

Several limitations should be acknowledged. Most importantly, the small number of DDM-based observations constrains the scope for robust comparison across all valuation approaches. Moreover, the analysis relies primarily on a mean-based measure of valuation accuracy, defined as the deviation between the target price and the subsequent market price over a specified horizon. An additional limitation is that the study does not verify ex ante whether the specific assumptions underlying each valuation model were fully satisfied in the individual brokerage reports; instead, it evaluates the ex post accuracy of the target prices generated using those models. The results should also be interpreted with caution because the distributions of valuation errors are positively skewed and leptokurtic, which implies that conventional summary measures such as the mean and standard deviation may be sensitive to extreme observations. For this reason, the reported results should be treated primarily as evidence on average valuation error within the adopted research design. Future research could therefore usefully complement the present framework with more robust descriptive and inferential measures, including median-based statistics, quantile analysis, trimmed means, and other outlier-resistant estimation procedures. This would allow for a more reliable assessment of valuation accuracy in the presence of skewed and heavy-tailed error distributions. Further extensions could also explicitly account for differences across forecast horizons, market regimes, and macro-financial conditions. It would also be valuable to examine whether the patterns identified in this study are specific to the banking sector or reflect a broader feature of equity valuation in analysts' recommendations across sectors and markets.

References

Adamczyk P., Zbroszczyk A. (2017), Business valuation – practice of domestic WSE members in 2016, *Finanse, Rynki Finansowe, Ubezpieczenia*, 5(89), 233–249.

Adipramana D., Dirgantari P.D., Masharyono (2025), *The influence of profitability and price on book value on stock prices (a study on banking companies in the Indonesia Stock Exchange in the LQ45 Index for the years 2019–2023)*, Proceedings of the 9th Global Conference on Business, Management and Entrepreneurship (GCBME 2024).

Asquith P., Mikhail M.B., Au A.S. (2005), Information content of equity analyst reports, *Journal of Financial Economics*, 75(2), 245–282.

Biedrzyński R. (2008), *Wycena w rekomendacjach sporządzanych przez biura maklerskie a cena rynkowa spółek notowanych na giełdzie papierów wartościowych w Warszawie*, unpublished, Uniwersytet Ekonomiczny w Szczecinie.

Bolliger G. (2004), The characteristics of individual analysts' forecasts in Europe, *Journal of Banking & Finance*, 28(9), 2283–2309.

Bradshaw M.T. (2002), The use of target prices to justify sell-side analysts' stock recommendations, *Accounting Horizons*, 16(1), 27–41.

Bradshaw M.T., Brown L.D., Huang K. (2012), Do sell-side analysts exhibit differential target price forecasting ability?, *SSRN Electronic Journal*, 1–49.

Brycz B., Dudycz T., Włodarczyk K. (2021), Are analysts really optimistic in their stock recommendations? The case of the Polish capital market, *Emerging Markets Finance and Trade*, 57(13), 3649–3676.

Chen N. (2023), Valuation analysis of the banking industry based on DCF, P/E and EBITDA valuation, *BCP Business & Management*, 47, 154–160.

Cowles A. (1933), Can stock market forecasters forecast?, *Econometrica*, 1(3), 309–324.

Dąbrowski P. (2013), Huraoptymizm rekomendacji maklerskich w okresie giełdowej dekoniunktury, *Studia Ekonomiczne*, 174, 88–104.

Erkilet G., Janke G., Kasperzak R. (2022), How valuation approach choice affects financial analysts' target price accuracy, *Journal of Business Economics*, 92(5), 741–779.

Gianfrate G., Vincenzi R. (2013), How do analysts value banks? An empirical analysis, *Bancaria,* 2, 32–46.

Głębocki M., Grudziński M., Kundera M., Sylwestrzak M. (2011), Studium metodyki wyceny przedsiębiorstw stosowanej w rekomendacjach giełdowych, *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia*, 37, 575–588.

Godłów-Legiędź J. (2013), Ekonomia behawioralna: od koncepcji racjonalności do wizji ustroju ekonomicznego, *Ekonomia*, 4(25), 24–41.

Gounder C.G., Venkateshwarlu M. (2017), Bank valuation models – a comparative analysis, *Accounting and Finance Research*, 6(3), 116–134.

Gross S. (2006), *Banks and Shareholder Value. An Overview of Bank Valuation and Empirical Evidence on Shareholder Value for Banks*, Deutscher Universitäts-Verlag.

Ho T., Nguyen T., Nguyen Y., Brownen-Trinh R. (2023), *How do equity research analysts value banks? Evidence from North American and European banks*, unpublished, Working Paper, University of Bristol, St. Francis Xavier University, https://www.bayes.citystgeorges.ac.uk/__data/assets/pdf_file/0010/729694/Brownen-Trinh-et-al-2023-How-equity-research-analysts-value-banks.pdf.

Imam S., Barker R., Clubb C. (2008), The use of valuation models by UK investment analysts, *European Accounting Review*, 17(3), 503–535.

King B.M., Minium E.W. (2020), *Statystyka dla psychologów i pedagogów,* Wyd. Naukowe PWN.

Klepacki J. (2019), *Indeks behawioralny w prognozowaniu i decyzjach finansowych*, Difin.

Kowalke K. (2015), Prognozy finansowe biur maklerskich a rzeczywiste wyniki finansowe spółek giełdowych, *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia*, 75, 269–281.

Kwiatkowska K. (2024), Czynniki finansowe a wartość rynkowa banków: przegląd literatury, *Zeszyty Naukowe Akademii Górnośląskiej*, 26, 24–33.

Leong K.Y., Ariff M., Alireza Z., Bhatti M.I. (2023), Bank stock valuation theories: do they explain prices based on theories?, *International Journal of Managerial Finance*, 19(2), 331–350.

Majewski S., Majewska A. (2017), Wartość wewnętrzna jako determinanta procesu decyzyjnego w ujęciu behawioralnym, *Finanse, Rynki Finansowe, Ubezpieczenia*, 6, 37–50.

Malska W., Koziorowska A. (2015), Wykorzystanie testu *t* dla pojedynczej próby we wnioskowaniu statystycznym, *Edukacja – Technika – Informatyka*, 6(3), 323–327.

Martínez F.G., Jiménez J.D., Matas R.Q.S.d.l. (2024), Reevaluating bank price-to-book ratios: an in-depth analysis of equity components across economic cycles, *Journal of Risk and Financial Management*, 17(8), 363.

Massari M., Difonzo C., Gianfrate G., Zanetti L. (2018), Bank valuation using multiples in US and Europe: an historical perspective, *Business Valuation OIV Journal*, 53–76.

Mehran H., Thakor A. (2011), Bank capital and value in the cross-section, *The Review of Financial Studies*, 24(4), 1019–1067.

Moyo V., Obadire A.M. (2024), A case study of bank equity valuation methods employed by South African, Nigerian and Kenyan equity researchers, *Risks*, 12(6), 89.

Pastusiak R. (2016), Nadmierny optymizm w wycenach przedsiębiorstw, *Annales Universitatis Mariae Curie-Skłodowska. Sectio H Oeconomia*, 50(4), 383–392.

Pastusiak R. (2017), Heurystyka nadmiernego optymizmu w wycenach przedsiębiorstw, *Annales Universitatis Mariae Curie-Skłodowska. Sectio H Oeconomia*, 51(6), 283–291.

Prusak B. (2015), *Wielokryterialna analiza rekomendacji giełdowych*, CeDeWu.

Prusak B. (2017), The accuracy of alternative stock valuation methods – the case of the Warsaw Stock Exchange, *Economic Research – Ekonomska Istraživanja*, 30(1), 416–438.

Radke M. (2023), *Nadmierny optymizm analityków sporządzających rekomendacje giełdowe w wycenach przedsiębiorstw*, Wydawnictwo Uniwersytetu Łódzkiego.

Szarzec K. (2019), *Koniunktura gospodarcza*, in: A.Z. Nowak, T. Zalega (eds.), *Makroekonomia*, Polskie Wydawnictwo Ekonomiczne.

Szymański P. (2012), Jakie problemy napotykają eksperci przy wycenie przedsiębiorstw? Wyniki badań, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 261, 435–448.

Wośko Z. (2005), *Wpływ zmian na rynkach finansowych na przebieg wahań koniunkturalnych w Polsce,* unpublished, Uniwersytet Łódzki.

Zarzecki D. (2008), Indeks ryzyka w wycenie przedsiębiorstw, *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania/Uniwersytet Szczeciński*, (1), 105–114.

Zarzecki D. (2011), O kluczowych wyzwaniach wyceny przedsiębiorstw, wycenie zobowiązań warunkowych i wycenie aktywów nieoperacyjnych, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 158, 514–524.

Zarzecki D. (2013), Kluczowe wyzwania wyceny przedsiębiorstw, wycena zobowiązań warunkowych i aktywów nieoperacyjnych, *Finanse, Rynki Finansowe, Ubezpieczenia*, 63, 533–547.

Information

In preparing this paper, artificial intelligence (AI) tools were used solely for language editing and proofreading. AI support was limited to improving grammar, spelling, punctuation, clarity, and overall linguistic quality. No AI tools were used to generate scientific content, interpret results, draw conclusions, create figures, or for data collection, analysis, or research-related activities.

Appendix

Table 1
Assessment scale

Assessment	Range of deviation
Accurate	0–5%
Acceptable	> 5% to 15%
Unacceptable	> 15%

Source: own elaboration.

Table 2
Descriptive statistics for ROE–P/BV model

N	Statistic	225
Range	Statistic	310.38%
Minimum	Statistic	0.43%
Maximum	Statistic	310.80%
Mean	Statistic	39.33%
	Std. error	2.67%
	5% trimmed mean	34.58%
Median	Statistic	29.03%
Std. deviation	Statistic	40.01%
Variance	Statistic	1600.882
Skewness	Statistic	2.724
	Std. error	0.162
Kurtosis	Statistic	11.558
	Std. error	0.323
CV	Statistic	101.73%

Source: own elaboration.

Table 3
One-sample test for ROE–P/BV model (test value = 15%)

	t	df	Significance		Mean difference	95% confidence interval of the difference	
			one-sided p	two-sided p		lower	upper
Valuation error	9.112	224	< 0.001	< 0.001	24.33%	19.08%	29.59%

Source: own elaboration.

Table 4
One-sample effect sizes for ROE–P/BV model

		Standardizer[a]	Point estimate	95% confidence interval	
				lower	upper
Valuation error	Cohen's d	40.01103%	0.608	0.465	0.750
	Hedges' correction	40.14562%	0.606	0.464	0.747

[a] The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

Source: own elaboration.

Table 5
Descriptive statistics for RIM model

N	Statistic	154
Range	Statistic	440.68%
Minimum	Statistic	0.80%
Maximum	Statistic	441.48%
Mean	Statistic	40.37%
	Std. error	4.34%
	5% trimmed mean	32.71%
Median	Statistic	30.39%
Std. deviation	Statistic	53.80%
Variance	Statistic	2894.401
Skewness	Statistic	5.142
	Std. error	0.195
Kurtosis	Statistic	32.374
	Std. error	0.389
CV	Statistic	133.28%

Source: own elaboration.

Table 6
One-sample test for RIM model (test value = 15%)

	t	df	Significance		Mean difference	95% confidence interval of the difference	
			one-sided p	two-sided p		lower	upper
Valuation error	5.851	153	< 0.001	< 0.001	25.37%	16.80%	33.93%

Source: own elaboration.

Table 7
One-sample effect sizes for RIM model

		Standardizer[a]	Point estimate	95% confidence interval	
				lower	upper
Valuation error	Cohen's d	53.79963%	0.471	0.304	0.637
	Hedges' correction	54.06517%	0.469	0.303	0.634

[a] The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

Source: own elaboration.

Table 8
Descriptive statistics for relative valuation based on multiples

N	Statistic	259
Range	Statistic	246.77%
Minimum	Statistic	0.02%
Maximum	Statistic	246.78%
Mean	Statistic	39.34%
	Std. error	2.40%
	5% trimmed mean	34.45%
Median	Statistic	29.39%
Std. deviation	Statistic	38.56%
Variance	Statistic	1486.988
Skewness	Statistic	2.287
	Std. error	0.151
Kurtosis	Statistic	6.956
	Std. error	0.302
CV	Statistic	98.03%

Source: own elaboration.

Table 9

One-sample test for relative valuation based on multiples (test value = 15%)

	t	df	Significance		Mean difference	95% confidence interval of the difference	
			one-sided p	two-sided p		lower	upper
Valuation error	10.157	258	< 0.001	< 0.001	24.34%	19.62%	29.05%

Source: own elaboration.

Table 10

One-sample effect sizes for relative valuation based on multiples

		Standardizer[a]	Point estimate	95% confidence interval	
				lower	upper
Valuation error	Cohen's d	38.56148%	0.631	0.497	0.764
	Hedges' correction	38.67403%	0.629	0.496	0.762

[a] The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

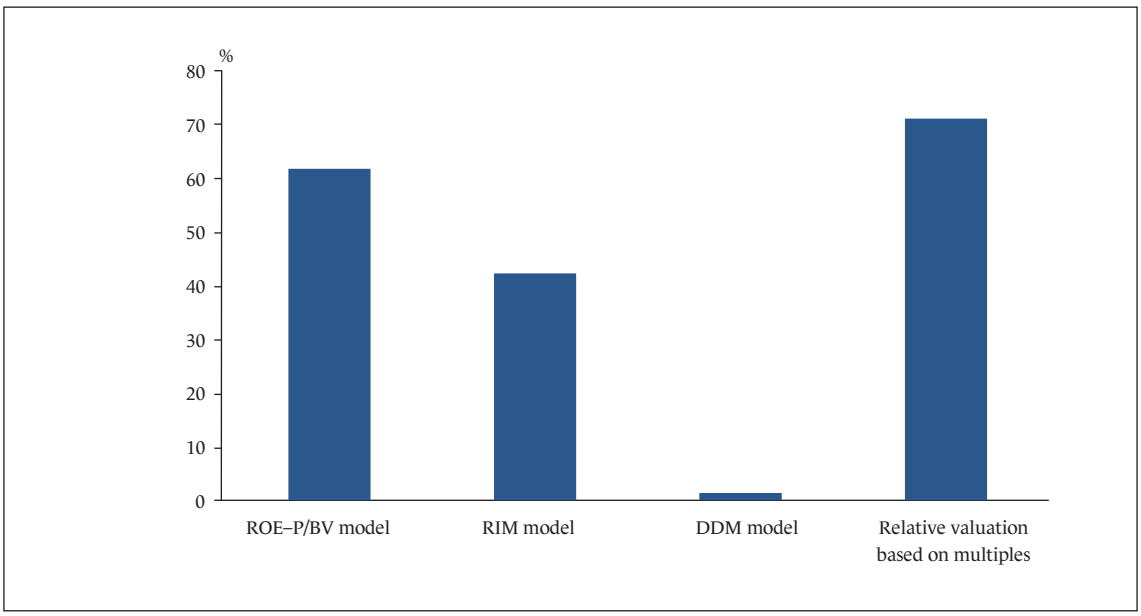
Source: own elaboration.

Table 11
Descriptive statistics for DDM model

N	Statistic	5
Range	Statistic	67.99%
Minimum	Statistic	21.43%
Maximum	Statistic	89.42%
Mean	Statistic	50.68%
	Std. error	13.14%
	5% trimmed mean	50.15%
Median	Statistic	49.30%
Std. deviation	Statistic	29.39%
Variance	Statistic	863.533
Skewness	Statistic	0.33
	Std. error	0.913
Kurtosis	Statistic	-1.872
	Std. error	2

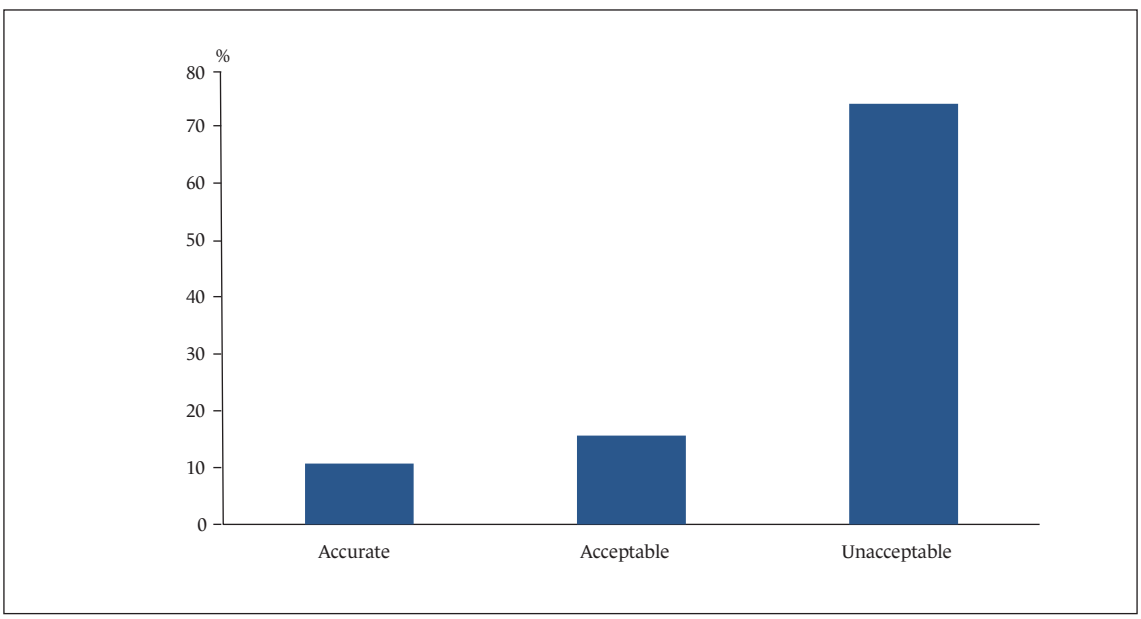
Source: own elaboration.

Figure 1
Frequency of use of bank valuation models



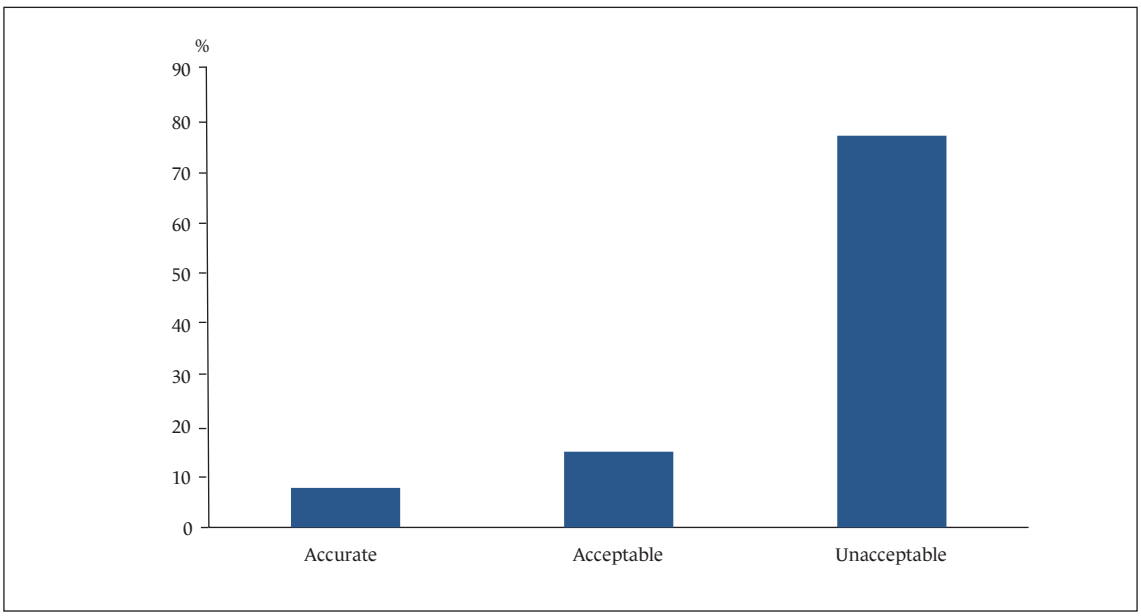
Source: own elaboration.

Figure 2
Accuracy of valuations using the ROE–P/BV model



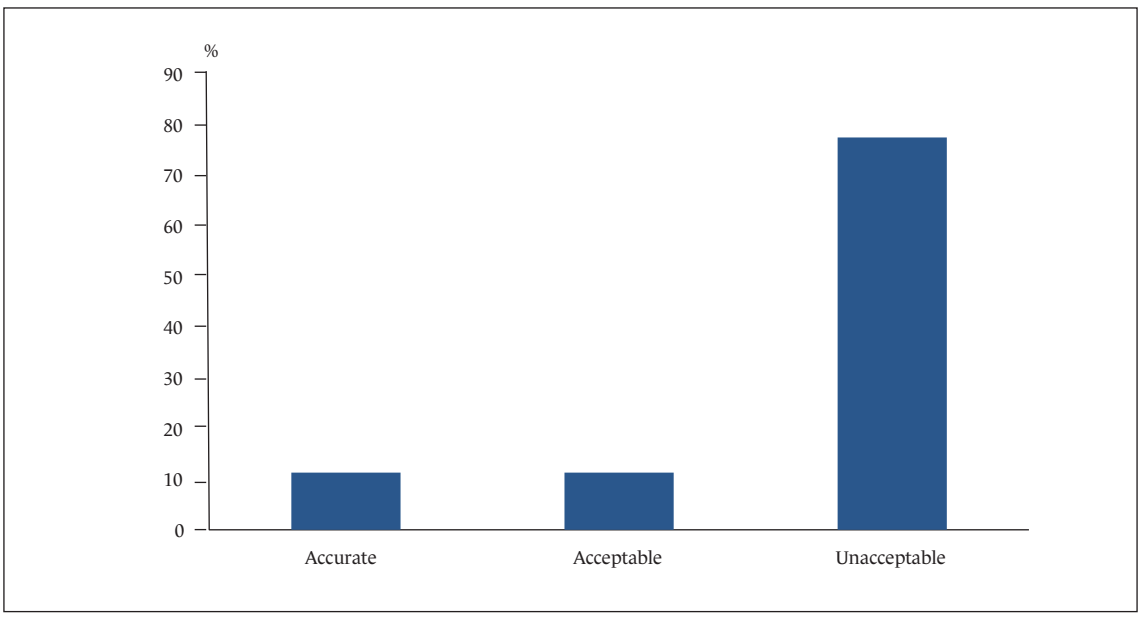
Source: own elaboration.

Figure 3
Accuracy of valuations using the RIM model



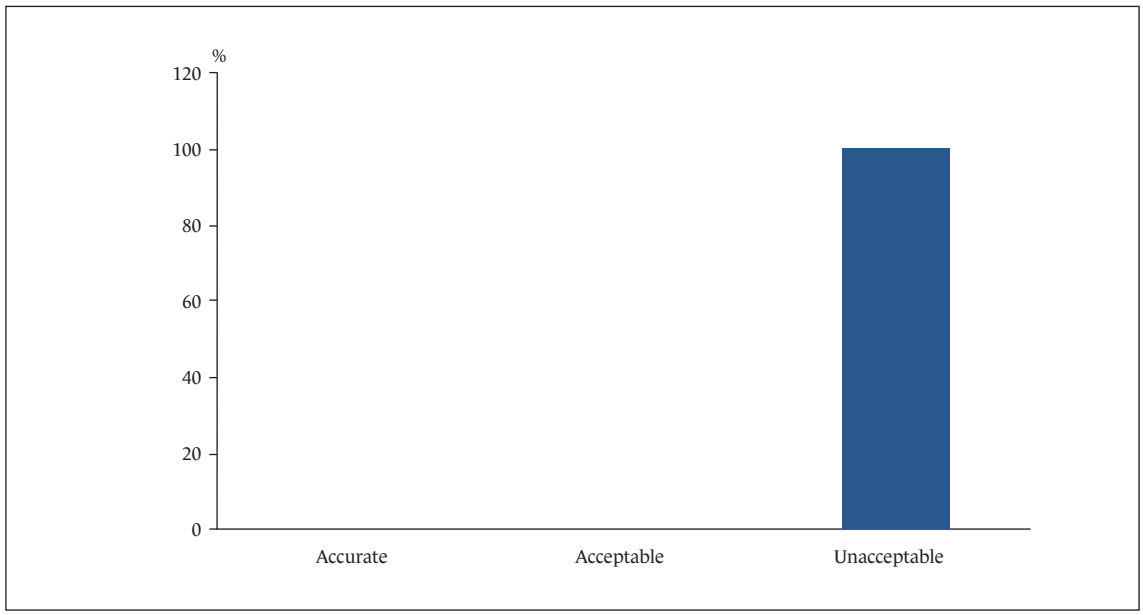
Source: own elaboration.

Figure 4
Accuracy of valuations using the relative valuation based on multiples



Source: own elaboration.

Figure 5
Accuracy of valuations using the DDM model



Source: own elaboration.

Ocena trafności wycen banków sporządzanych z wykorzystaniem wybranych modeli wyceny w rekomendacjach giełdowych na polskim rynku kapitałowym

Streszczenie

Wycena przedsiębiorstw stanowi istotny element procesu podejmowania decyzji na rynku kapitałowym, szczególnie w sytuacjach związanych z kupnem i sprzedażą przedsiębiorstw, fuzjami i przejęciami, emisją nowych akcji oraz restrukturyzacją przedsiębiorstw. W Polsce znaczenie wyceny przedsiębiorstw istotnie wzrosło wraz z transformacją ustrojową i przejściem do gospodarki rynkowej. Jednocześnie wycena pozostaje procesem złożonym, ponieważ wartość przedsiębiorstwa zależy od licznych czynników wewnętrznych i zewnętrznych oraz wymaga wykorzystania wiedzy z zakresu ekonomii, finansów i zarządzania. Problemy te są szczególnie istotne w sektorze bankowym, w którym na wartość wpływają m.in.: rentowność, wymogi kapitałowe, jakość aktywów, ryzyko, otoczenie regulacyjne oraz nastroje rynkowe. Dotychczasowe badania wskazują, że analitycy wykorzystują różnorodne podejścia do wyceny, wśród których istotną rolę odgrywają wyceny porównawcze, metody dochodowe oraz modele dostosowane do specyfiki banków. Jednocześnie wcześniejsze badania wskazują na znaczne rozbieżności pomiędzy cenami docelowymi zawartymi w rekomendacjach a późniejszymi cenami rynkowymi. Nadal jednak stosunkowo niewiele badań dotyczy dokładności poszczególnych modeli wyceny banków notowanych na polskim rynku kapitałowym. Celem niniejszego badania jest zatem ocena dokładności wycen banków zawartych w rekomendacjach giełdowych na polskim rynku kapitałowym, przygotowanych z wykorzystaniem wybranych modeli wyceny.

Podstawowe pytanie badawcze dotyczy tego, czy wyceny polskich banków zawarte w rekomendacjach biur maklerskich zapewniają wystarczająco dokładne oszacowanie przyszłych cen rynkowych akcji. Analizie poddano cztery modele wyceny: model dochodu rezydualnego (RIM), model zdyskontowanych dywidend (DDM), wycenę porównawczą opartą na mnożnikach oraz model ROE–P/BV. W badaniu sprawdzono, czy średni bezwzględny błąd wyceny uzyskiwany przy zastosowaniu każdego z modeli przekracza poziom 15%, przyjęty jako operacyjny próg akceptowalnej dokładności. Sformułowano cztery hipotezy badawcze: H1 zakłada, że średni błąd wycen przygotowanych z wykorzystaniem RIM przekracza 15%; H2 odnosi to samo założenie do DDM; H3 do wyceny porównawczej opartej na mnożnikach; natomiast H4 do modelu ROE–P/BV.

Badanie empiryczne obejmuje 364 rekomendacje giełdowe dotyczące 11 banków. Rekomendacje zostały przygotowane przez polskie i zagraniczne biura maklerskie i musiały zawierać informacje dotyczące spółki, daty wydania rekomendacji, horyzontu inwestycyjnego, ceny docelowej oraz zastosowanego modelu wyceny. Z analizy wyłączono rekomendacje, dla których nie określono horyzontu inwestycyjnego, ponieważ uniemożliwiało to ustalenie daty zapadalności rekomendacji. Ostateczny okres badawczy obejmuje czas od 25 października 2005 r. do 13 sierpnia 2021 r., czyli niemal 16 lat, obejmujących różne fazy cyklu gospodarczego i giełdowego. Analizowane rekomendacje charakteryzowały się czterema horyzontami inwestycyjnymi: 3, 6, 9 oraz 12 miesięcy.

Dokładność wyceny określono jako procentowe odchylenie ceny docelowej zawartej w rekomendacji od kursu zamknięcia akcji obserwowanego na koniec właściwego horyzontu inwestycyjnego. Poziom

15% przyjęto na podstawie wcześniejszych badań dotyczących dokładności prognoz cen docelowych, traktując go jako próg operacyjny, a nie uniwersalną granicę dokładności. Do weryfikacji hipotez zastosowano jednopróbkowy test t-Studenta dla średniego błędu wyceny. Dodatkowo wykorzystano statystyki opisowe, m.in. odchylenie standardowe, współczynnik zmienności, skośność i kurtozę, aby określić zróżnicowanie i charakter rozkładów błędów wyceny.

Wyniki wskazują na ogólnie niską dokładność predykcyjną analizowanych modeli wyceny banków. Najczęściej stosowana była wycena porównawcza oparta na mnożnikach – wykorzystano ją w 259 rekomendacjach (71,15%). Drugim najczęściej stosowanym modelem był ROE–P/BV, wykorzystany w 225 rekomendacjach (61,81%). Model RIM zastosowano w 154 rekomendacjach, natomiast DDM jedynie w pięciu.

Dla modelu ROE–P/BV średni błąd wyceny wyniósł 39,33% (N = 225), a więc zdecydowanie przekroczył przyjęty próg 15%. Jednopróbkowy test t-Studenta potwierdził statystyczną istotność tej różnicy: $t(224) = 9{,}112$, $p < 0{,}001$. Wielkość efektu była umiarkowana (d Cohena = 0,608). Jedynie 10,67% rekomendacji zaklasyfikowano jako dokładne, 15,56% jako akceptowalne, natomiast aż 73,78% jako niedokładne.

W przypadku modelu RIM średni błąd wyceny wyniósł 40,37% (N = 154) i był istotnie wyższy od poziomu 15%, $t(153) = 5{,}851$, $p < 0{,}001$. Wielkość efektu również była umiarkowana (d Cohena = 0,471). Tylko 7,79% rekomendacji było dokładnych, 14,94% akceptowalnych, a 77,27% niedokładnych.

Wycena porównawcza oparta na mnożnikach charakteryzowała się średnim błędem wynoszącym 39,34% (N = 259). Różnica względem poziomu 15% była statystycznie istotna, $t(258) = 10{,}157$, $p < 0{,}001$, a wielkość efektu była umiarkowana (d Cohena = 0,631). Za dokładne uznano 11,29% rekomendacji, kolejne 11,29% za akceptowalne, natomiast 77,42% za niedokładne.

Najwyższy średni błąd odnotowano dla modelu DDM – 50,68%. Wynik ten opiera się jednak wyłącznie na pięciu obserwacjach. Wszystkie rekomendacje wykorzystujące DDM zostały sklasyfikowane jako niedokładne. Ze względu na bardzo małą liczebność próby nie przeprowadzono testu t.

Wyniki badania wskazują, że wyceny banków zawarte w rekomendacjach giełdowych na polskim rynku kapitałowym charakteryzują się ograniczoną dokładnością predykcyjną. W przypadku trzech modeli poddanych formalnej weryfikacji statystycznej średni błąd wyceny istotnie przekraczał przyjęty poziom 15%. Jednocześnie wyniki nie wskazują na istnienie jednego modelu, który można uznać za zdecydowanie najlepszy. Modele RIM, ROE–P/BV oraz wycena porównawcza oparta na mnożnikach charakteryzowały się zbliżonymi średnimi błędami, natomiast DDM uzyskał najwyższy średni błąd, przy czym jego bardzo mała próba nie pozwala na wyciąganie silnych wniosków porównawczych.

Rezultaty sugerują, że problem nie sprowadza się do mechanicznego wyboru jednego „najlepszego" modelu wyceny, lecz wynika w większym stopniu z trudności z wyceną banków w warunkach niepewności regulacyjnej, makroekonomicznej i rynkowej. Z punktu widzenia praktyki wyceny zasadne wydaje się zastąpienie pojedynczych cen docelowych przedziałami wartości oraz stosowanie jawnej analizy scenariuszowej, obejmującej wariant bazowy, optymistyczny i pesymistyczny. Modele powinny również w większym stopniu uwzględniać specyficzne czynniki determinujące wartość banków, takie jak adekwatność kapitałowa, koszt ryzyka, jakość aktywów, rentowność oraz zmiany otoczenia regulacyjnego. Dla inwestorów oznacza to, że cena docelowa zawarta w rekomendacji powinna być traktowana jako jeden z elementów procesu decyzyjnego, a nie jako samodzielny i wystarczająco wiarygodny sygnał inwestycyjny.

Najważniejszym ograniczeniem jest bardzo mała liczba obserwacji dotyczących modelu DDM, co uniemożliwia przeprowadzenie wiarygodnego porównania statystycznego tego modelu z pozostałymi podejściami. Ponadto dokładność wyceny oceniano przede wszystkim za pomocą średniego błędu, a badanie nie weryfikowało *ex ante*, czy założenia poszczególnych modeli wyceny były w pełni spełnione w przypadku każdej analizowanej rekomendacji. Dodatkowo rozkłady błędów charakteryzowały się dodatnią skośnością i wysoką kurtozą, co oznacza, że średnia i odchylenie standardowe mogą być wrażliwe na obserwacje ekstremalne. W przyszłych badaniach zasadne byłoby zatem wykorzystanie bardziej odpornych miar, takich jak mediana, analiza kwantylowa czy średnie ucinane, a także uwzględnienie różnic pomiędzy horyzontami inwestycyjnymi, fazami cyklu rynkowego i warunkami makroekonomicznymi.

Słowa kluczowe: wycena, banki, rekomendacje giełdowe, trafność