Bank Valuation Using Multiples in US and Europe: An Historical Perspective

Mario Massari* - Christopher Difonzo** - Gianfranco Gianfrate*** - Laura Zanetti****

We investigate the performance of relative valuation for US and European banks over the period 1990-2017. While the literature on the use of multiples is well developed, the relative valuation of financial institutions has received scant attention. We study the distribution and the main properties of each multiple's valuation errors, assessing which multiples work best and should be preferred when valuing banks. Our results show that on average high levels of accuracy are achieved by two years forward P/E. Moreover, diluted earnings not including extraordinary items should be preferred when computing trailing earnings multiples and, interestingly, P/BV consistently outperforms P/TBV. Dividends' multiples are not among the best performers, anyway, it is preferable to consider only common dividends when computing them. Most of the times P/Deposits and P/Revenues deliver poor performances, therefore it is advisable not to use them as main valuation approach.

1. Introduction

The performance of multiples with respect to equity valuation of non-financial companies has been extensively debated in financial and accounting literature (Liu et al., 2002; Liu et al., 2007; Nissim, 2013). However, research and evidence are limited concerning the equity valuation of banks. In fact, the relative valuation approach (also referred to as "market approach") may represent the simplest way to value a bank: the approach specifies the value of the bank as a function of selected fundamentals and the average price of peer banks (Forte et al., 2018; Nissim, 2013).

This work analyzes the accuracy of the market approach for US and European bank valuation. We first measure the performance 08 multiples based on value drivers such as the book value of equity, the tangible book value of equity, revenue, trailing earnings, forward earnings, common dividends, total dividends, bank deposits, and customer deposits. Following Liu et al. (2002), we measure the accuracy of multiples by comparing the "theoretical" valuation of banks obtained using multiples to the actual prices: multiples that produce the lowest errors – meaning the difference between theoretical prices and actual prices – are considered to be most accurate.

The results of our analysis show that the accuracy of multiples for US entities is significantly higher when European metrics are used, whereas small retail and investment banks present more of a valuation challenge than large retail banks. Forward Price/Equity (P/E) multiples outperform historical multiples, and multiples based on two-year-ahead forecasts (not just one-year-ahead) are more accurate. Despite the usual practitioner assumptions, Price/Tangible Book Value (P/TBV) is not found to be more meaningful and precise than Price/Book Value (P/BV). The P/BV is preferred. This study also reveals the weak relationship between value and the amount of preferred dividends: P/Common Dividends is a more precise tool than P/Total Dividends. Finally, P/Bank Deposits appears to be an accurate value driver when valuing investment banks, whereas P/Customer Deposits is preferred when addressing commercial banks.

The structure of this paper is as follows. A description of relative valuation (introducing all of the multiples analyzed) is presented in Section 2, while Section 3 summarizes the major contributions published in literature. Section 4 describes the data and the methodology adopted to assess the performance and accuracy of multiples and presents all of the results for each analysis subsample. The impact of the financial crisis and the introduction of the Euro on relative valuation precision are studied. Additionally, regression and correlation analyses investigate whether significant positive and negative errors, corresponding to undervalued and overvalued banks, reflect subsequent price reactions.

2. Relative valuation

The use of multiples to perform company valuation has been showing an increasingly positive trend, following the development of financial markets and corporate finance deals during the last decades. Moreover,

* Bocconi University.
** Bocconi University.
*** EDHEC Business School.
**** Bocconi University.

great supporters of complex valuation techniques frequently recall the use of multiples when estimating terminal values or checking for plausibility of their results (Bhojraj & Lee, 2002).

The logic behind relative valuation is grounded on the assumption that market prices are largely efficient, that, on average, fundamentals are correctly priced in, and that the law of one price holds. The basic principle governing relative valuation (i.e., similar assets, and so similar firms, should trade at similar levels) fully relies on this set of assumptions. Market prices thus need to be close to the true intrinsic value of firms.

The following list of multiples is the collection of the ones selected to run this empirical analysis, which correspond to the ones mostly used and considered by analysts and practitioners when running valuation for banks.

Price/Book Value of Equity (P/BV)

The ratio between the market capitalization of the firm and the book value of equity is widely used for capital-intensive businesses, whereas it is less appropriate for sectors where the main driver of price performances is future growth (e.g., technology). It is considered one of the most suitable multiples for financial institutions since it captures the regulatory attention on solvency, capital requirements and equity maintenance.

Price/Tangible Book Value of Equity (P/TBV)

This multiple is a variation of the previous one and deducts the value of all the intangible assets from the equity. Many practitioners prefer to use this multiple over the simple P/BV in order to obtain a more conservative figure, which uses a liquid representation of book value eliminating the potential bias deriving from the accounting of illiquid intangible assets. The intuition is that, in case of default, the value of intangible assets may easily collapse to zero, so it is advisable to use a multiple that eliminates their interferences. Moreover, it recalls the regulatory capital composition of the CET1 Capital according to Basel III, which indeed deducts goodwill and other intangibles.

Price/Revenues (P/Revenues)

Market capitalization divided by revenues is one of the less used and most criticized multiple. Firstly, because revenues should be compared with an asset-side measure (e.g., enterprise value). Secondly, comparing banks using only this multiple may lead to misjudgements because the cost structure and the riskiness of the underlying assets, which generated those revenues, are not considered.

Price/Deposits (P/Deposits)

Market capitalization is divided by the deposits, which is the core driver for the vast majority of commercial banks. This multiple used to be popular in the past but nowadays banks are more diversified and their revenues and profitability depend more and more on fees-generating activities rather than on the sole interest income. This explains why this multiple is now less used, but it is still helpful. In fact, given that deposits should be rather uniform among retail banks, they are a good candidate as explicative operating multiple.

When considering deposits, these could be measured in two different ways: the first one considers customer deposits only (i.e., demand, savings and time deposits held on account for individuals and corporations), the other one includes also bank deposits (i.e., the deposits held on account for other banks). The latter thus considers not only collections from customers, but also the involvement in the interbank market. This broad measure is the one used in this work.

Price/Dividends (P/Dividends)

Share price is divided by dividends per share, or, alternatively, market capitalization is divided by the entire amount of dividends. This is another operating multiple typically used for banks because of the importance of dividends for these institutions. In fact, dividends are the unique meaningful cash flow in the banking sector, as also highlighted when discussing about intrinsic valuation. However, this multiple could be applied to firms operating in any sector, but it would be meaningless in many circumstances.

Many companies, differently from banks, do not distribute dividends so frequently because they prefer to implement alternative shareholders' remuneration practices (e.g., shares buy-back programs) or they simply prefer to retain earnings to finance investment opportunities using internal sources.

When computing dividends multiples, it is important to consider that dividends distributions may occur more than once in a year and so all the relevant flows have to be summed up to obtain a yearly value. Moreover, this multiple can be built in two different ways, depending on the choice made for dividends. Considering that total dividends is the sum of common dividends, paid on common shares, and preferred dividends, paid on preferred shares, the multiple can be computed using total dividends (P/Total Dividends) or common dividends only (P/Common Dividends) as the denominator. The distinction wants to catch potential connections between preferred stocks and value. The use of common dividends only is generally preferred because they should better reflect value with respect to preferred dividends, which are more stable and less dependent on the actual level of profitability achieved in a given year. However, for the sake of this empirical study, both will be computed.

The last important consideration is related to outliers, which in this case can strongly affect the multiple. If dividends are particularly low because of a lack of financial resources or as a result of a strategic choice, the average multiple may reach extremely high values,

negatively affecting the quality of the valuation performed.

Price/Earnings (P/E)

The last multiple analysed is considered "the king" of relative valuation, in particular for banks. It is computed as the ratio between the share price and earnings per share (EPS) or alternatively as market capitalization over total earnings. The multiple can be built in different ways, depending on the methodology used to select earnings.

Firstly, the choice can be made distinguishing between trailing and forward earnings. If historical values, i.e., the earnings of the last twelve months (LTM earnings) are used, it is classified as a trailing – or LTM. Alternatively, if analysts' forecasts for earnings are used, it is classified as a forward multiple. Forecasts can be computed on a one year, two years or more years basis, but best practices generally use one or two years forecasts, to avoid a strong dependence on estimates based on unobservable and unpredictable figures. Nevertheless, Yee (2004) demonstrated that from a theoretical standpoint the use of more forward earnings represents an effective and important attribute in order to obtain more accurate results when performing valuation (i.e., the more forward, the more accurate).

The second element affecting EPS calculation is dilution. The resulting multiples are the Basic P/E, if EPS are computed considering only outstanding common shares or instead the Diluted P/E if diluted common shares are considered. Diluted common shares include the effects generated by the hypothetical exercise of all the outstanding convertible securities (e.g., convertible bonds, stock options, warrants), which causes an increase in the number of outstanding shares. This assumed increase pushes EPS down (diluted EPS are lower than basic EPS, if there are convertible security outstanding) and, consequently, the resulting multiple is higher.

The last point about earnings calculation is whether to include non-recurring items. The rationale behind the exclusion of these items is that unusual and extraordinary gains or losses should not affect valuation, since these will not constantly take place in the future. In this way, earnings excluding extraordinary items communicate better the actual profitability of a company without suffering from any interference directed by one-offs.

The combination of all these aspects and considerations gives rise to the identification of six different P/E multiples, which will be inspected in this work. They are:

– P / 1 Year Forward Earnings
– P / 2 Years Forward Earnings
– P / LTM Diluted Earnings, considering extraordinary items
– P / LTM Diluted Earnings, excluding extraordinary items
– P / LTM Basic Earnings, considering extraordinary items
– P / LTM Basic Earnings, excluding extraordinary items

Nonetheless, P/E has an important drawback that can limit its applicability. In case of negative earnings, the multiple becomes completely meaningless because of its negative value. In order to avoid any issue, the set of comparables must be built accordingly. Moreover, the presence of outliers should be accurately monitored in case of very low earnings that can generate an abnormal increase in the multiple. In particular, Dermine (2010) outlines that the use of the P/E is biased when banks report large provisions for credit losses (a problem that, recently, has been affecting the banking system of many countries, such as Italy) implying lower earnings. This causes large volatility in the multiple and drives bias.

3. Literature review

While the extensive use of multiples among both practitioners and academicians has progressively grown, theory and empirical research have also demonstrated some advancements, but still limited guidance is available to assess relative valuation metrics performance.

Essentially, some practitioners consider the use of multiples as an art form[1] rather than a science. Therefore, they suggest that the practice should be left only to industry professionals. Notwithstanding, the importance of multiples in valuation methods and their efficacy in supporting investment decisions have attracted many researchers to this field. Both standard literature and empirical studies on multiples have experienced notable advances over the past decades, becoming a debated topic among academicians.

Methodologies and findings from Nissim (2002 & 2011) and Cooper (2008) are particularly relevant for the development of this empirical study. Additionally, contributions to the literature coming from other authors provided an important theoretical support and many relevant intuitions.

Nissim (2011) analysed the accuracy of relative valuation for U.S. insurance companies. From March 1990 to January 2011, he monthly analysed a sample

[1] Bhojraj (2003) noted that the level of subjectivity required in the application of multiples, is inconsistent with a scientific standpoint. In particular, the selection process of comparable firms tends to rely strongly on individual analyst's expertise.

of 372 different firms, demonstrating that valuation performs better when using earnings forecasts (i.e., forward multiples) rather than reported earnings (i.e., trailing multiples). The same result will come from the analysis here performed. His study also proved that book value multiples perform robustly, in particular if the price-to-book ratio is conditioned to ROE. Moreover, Nissim observed other two relevant aspects, which are less marked in the results from the analysis here performed, but still evident. He compared the performance of Basic P/E and Diluted P/E, observing that the latter has higher predictive properties. He also showed that valuation accuracy substantially improves when using income before special items instead of reported income.

In a previous work Liu, Nissim & Thomas (2002) carried out a comprehensive analysis of multiples' precision in the U.S. between 1982 and 1999, drawing up a ranking of the better performing multiples, which holds true for almost every sector analysed. Multiples were ranked as follows: forward earnings measures as the best ones, then historical earnings measures as a valid second best option, cash flows measures and book value measures perform equally ranking as third, sales measures as the worst ones. These results are in line with the ones here obtained.

Cooper (2008) aimed at finding the optimal number of comparable firms to use when computing out-of-sample multiples. The results of his analysis highlighted that the use of about five comparables is optimal when some requirements are met (i.e., comparables operate in the same industry, their expected growth rates are close to the one of the target firm and their average growth rate stays within 1% of the target firm's growth rate). Cooper's work is extremely useful here for the statistical tools implemented, which will be introduced later in the empirical section, more than for the results achieved.

Cheng & McNamara (2000) inspected valuation accuracy when using historical P/E multiples, P/BV multiples and a combination of the two using equal weights. The analysis was performed for the U.S. equity market, firstly considered as an aggregate and then split depending on SIC codes[2]. They found that the equally weighted combination of P/E and P/BV performed better than both multiples alone, underlying that both earnings and book values are significant value drivers.

Alford (1992) tested the effects of the choice of comparable firms on the precision of valuation estimates when using earnings multiples. In particular, he focused on the use of industry membership and proxies for growth and risk for the selection of comparables. Results showed that valuation accuracy increases when the level of detail for the industry definition used to identify comparables is not too specific (i.e., three-digit SIC codes). Differently, Bhojraj & Lee (2002) implemented a matching mechanism to identify comparable firms based on the use of economic variables, rather than industry membership. The analysis here performed combines the different intuitions coming from these two studies: only banks will be considered, but the sample will be then subdivided depending on balance sheet figures determining size (large or small) and business model (commercial or investment bank).

Minjina (2009) implemented the same analysis done by Nissim, but he did not focus on the same market and on a unique sector. Indeed, his analysis embraced all the companies listed in the Bucharest Stock Exchange from January 2003 to June 2008, but excluded the financial sector. Results underlined that Price/Cash Flows (P/CF) and Enterprise Value/EBITDA (EV/EBITDA) are the first and second best multiples to use when valuing Romanian companies, whereas Price/Sales appeared to be the least reliable. As already mentioned, these multiples are not significant and somehow meaningless for banks, which also explains why the financial sector was excluded to perform this analysis. Another relevant outcome of Minjina's study was the observation of a lower performance accuracy for Romanian listed companies, if compared with companies from more developed countries. The lower efficiency of Romanian capital markets and the smaller size of Romanian companies are considered the main determinants of this finding. The same difference in accuracy will be evident later, when comparing results for multiples' accuracy between American and European banks.

Forte et al. (2018) investigates the role of relative valuation in the banking industry by evaluating the accuracy of a group of industry specific multiples. The results highlight that stock market multiples are best suited for US institutions, and that a two-year-forward P/E is the most precise metric. Contrary to practitioner beliefs, P/Tangible Book Value is less meaningful than P/BV. Multiples accuracy declines in case of small commercial banks relative to large commercial banks and investment bank relative to retail banks pointing out that for small retail bank and investment bank equity valuation using multiples becomes a more challenging exercise. Additionally, error distributions are exploited to assess whether large po-

[2] The Standard Industrial Classification (SIC) is a system for classifying industries by a four-digit code. SIC codes can be grouped into progressively broader industry classifications: industry group (the first three digits), major group (the first two digits) and division (encompassing a range of SIC codes).

sitive errors lead to systematic one-year positive price performances and whether negative errors lead to negative price changes.

4. Performance and accuracy of multiples

This section discusses about the main findings of the empirical research conducted to analyse the effectiveness of multiples and to understand if there is one multiple outperforming the others, which should be therefore considered the most reliable to perform banks valuation. Detailed analyses on valuation accuracy and on the distribution of valuation errors are presented. Furthermore, the effects on the performance of relative valuation driven by the introduction of the Euro and of the 2007/2008 financial turmoil are analysed. In the last part, a summary ranking of the best multiples to be used in each subgroup of banks is proposed.

4.1. Data

The timespan considered in our empirical analysis starts in January 1990 and terminates in April 2018. The dataset comprises all the banks currently listed, but also banks that have been listed during the period analysed. Delisting mostly derives from M&A activity or bankruptcy, more rarely it represents a strategic choice taken by the management. The final dataset is composed of 1,118 banks, of which 181 located in the Eurozone, while 937 are American. When building a database for comparables, there is always a "bias versus variability" trade-off to consider. In this case, variability is minimized, but the bias deriving from big differences among comparables may be relevant. The different techniques, which have been adopted to limit this effect and increase homogeneity, are showed later.

The analysis required a wide range of financial data, which have been collected from different data providers and then merged into a unique dataset. The list of these data providers and the corresponding data collected follows.

– Wharton Research Data Services – Compustat

This database provides all the historical Balance Sheet and Income Statement figures, along with other accounting measures (e.g., the number of shares outstanding). Data from 1990 until 2017 have been collected, using two different queries. The "Bank – Daily" query provides data for North America banks only, so data for banks in the United States were easily collected filtering only for the country. To collect data for the Eurozone, the "Compustat Global – Fundamental Annual" query was used. However, this dataset covers all the industries globally. Data were therefore filtered according to the GICS[3] codes, including only the entire Industry Group 4010 (Banks) and the Industry 402030 (Capital Markets), and according to the country, including only the ones within the Eurozone (Table 1). Furthermore, it is important to underline that these two queries do not provide the same information. In particular, the one used for the Eurozone does not provide data on diluted earnings, so that two multiples (i.e., P/LTM Diluted Earnings, considering extraordinary items and excluding extraordinary items) cannot be computed for these banks.

– Institutional Brokers' Estimate System (I/B/E/S)

This database, accessed via Thomson Reuters Datastream, provides all the analysts' forecasts, which are fundamental to compute forward multiples. Moreover, the measure of volatility of forecasts (i.e., the standard deviation of two years forward earnings) and the number of analysts covering each bank were collected from I/B/E/S.

– Bloomberg

Weekly prices have been collected from Bloomberg and then a monthly value has been obtained computing the median of the corresponding weekly observations. However, to compute multiples, only one price for each year is required and the April one has been selected. This choice is consistent with practice and follows the procedure implemented from Nissim in his study on relative valuation performance for the insurance sector. Market prices are selected four months after the fiscal year end to ensure that all year-end information are publicly available and reflected in prices (Schreiner, 2007). Moreover, in April, I/B/E/S updates and publishes summary forecasts, maximising also consistency between prices and future estimates.

Once that all data have been collected, in order to increase comparability and to reduce bias, banks have been divided in subgroups, maintaining the American and the European sample separated.

The first differentiation is related to the business model, distinguishing between Investment and Commercial Banks. Following the example of Beltratti and Stulz (2009), a summary ratio is computed for each bank as the median of the available ratios of Customer Loans over Total Assets, between 1990 and 2017. A threshold is set at 40%. Banks with a summary ratio exceeding the threshold are labelled as Commercial Banks, since the ratio signals that the business model is particularly focused on lending money to clients, more than on offering advisory services. Loans to banks are not included when computing the ratio to eliminate the effects of banks participating to the in-

[3] The Global Industry Classification Standard (CIGS) is an industry taxonomy developed by MSCI and Standard & Poor's used to categorise all major public companies. It consists of 11 sectors, 24 industry groups, 68 industries and 157 sub-industries.

terbank market, which may be the case also for pure Investment Banks.

The second distinction, which applies to Commercial Banks only, refers to size. Differentiating for size can make significant contributions since size has strong implications for the value of different banks (Alford, 1992). Large banks are generally less risky because their international scope gives them better access to customers and deposits, enhancing recurring revenues (Schreiner, 2007). Moreover, they can be perceived as "too big to fail", they can have more market power and enjoy economies of scale or scope and they can benefit from increased diversification, while small banks generally operate as niche players on a regional basis. Compared to small banks, large banks also enjoy greater financial flexibility having better access to capital market funds (Calomiris & Nissim, 2007). However, small banks can have higher strategic flexibility and growth potential, under the precondition of financial health and financing power. To fulfil a distinction based on size, for every bank, the median of Total Assets during the analysed years is computed and it is then compared with the median of the entire dataset. Banks exceeding this median are labelled as Large, the others as Small (Table 1).

Table 1 – Summary of Classification of Banks

		Investment Banks	Large Comm. Banks	Small Comm. Banks	Total
Number of Banks	U.S.	32	452	453	937
	Eurozone	31	75	75	181

4.2. Methodology

Relative valuation can be performed on the basis of out-of-sample multiples, so excluding the institution being valued from the group of banks considered for the computation of the multiple in each year. This methodology is considered the most reliable, since it minimises potential biases. Furthermore, multiples are computed using the harmonic mean: this way, the effects of outliers and of right asymmetry are strongly reduced (Nissim, 2011). A "theoretical" price is then computed multiplying the out-of-sample mean multiple by the corresponding value driver. If market prices are efficient, a theoretical price close to the actual market price suggests that a specific multiple performs well when running relative valuation. Therefore, to assess the performance of different multiples, for each bank in each year, the theoretical price is compared with the actual market price. It allows to calculate valuation errors (as percentage errors), as the difference between the theoretical price and the market price, divided by the actual price. According to Dittmann and Maug (2008), percentage errors, even though they are more basic than log errors, generate the least biased error when using the harmonic mean to aggregate the multiples of comparables. However, percentage errors penalise overvaluation more than undervaluation. This is why undervaluation in excess of −100% is impossible, while overvaluation is not limited and can easily go over +100%.

Setting x as the firm under analysis and t as the selected year, errors are computed as follows:

$$Error\ (x;t) = \\ = \frac{Multiple\ (all\ banks\ except\ x;t) * Value\ Driver\ (x;t) - Market\ Price\ (x;t)}{Market\ Price\ (x;t)}$$

Furthermore, in order to evaluate the performance of multiples, bias, mean absolute deviation (MAD) and mean-squared error (MSE) of the errors are computed, replicating Cooper and Cordeiro (2008) analysis. These measures are calculated according to the following formulas, where T is the total sum of observations (every bank for every year) and N is the total number of banks in each subsample:

$$Bias = \frac{1}{T}\sum_{t=1990}^{2017}\sum_{x=1}^{N} Error\ (x;t)$$

$$MAD = \frac{1}{T}\sum_{t=1990}^{2017}\sum_{x=1}^{N} |Error\ (x;t)|$$

$$MSE = \frac{1}{T}\sum_{t=1990}^{2017}\sum_{x=1}^{N} Error\ (x;t)^2$$

It is relevant to underline that MSE has been computed exploiting 95% Winsorization[4]. It is a common practice, which reduces the effects of large outliers that, once squared, can become too big and compromise results. Graphs are built using Normal Kernel Density[5] estimation, choosing a suitable bandwidth[6] and imposing the maximum level (1,000) of number of points at which evaluate the density function (or grid points) in order to do not lose the informative power of data. In order to lighten the chart, the bottom axis endpoint is set at 4.5 and not all multiples are included.

In order to evaluate multiples' performances, valuation accuracy is inspected following Nissim's procedure. The percentage of observations with estimated error in absolute value within 10%, 25%, 50%, 75% and 90% of price are computed. These measures are useful to understand which multiple is more accurate and reduces the size of errors.

4.3. Results

After having explained all the methodologies implemented to perform the empirical analysis, this section will focus on the main findings deriving from the analysis of data. A general overview presenting common elements among all the subsamples is firstly proposed. Next, banks in each subgroup are considered, mostly focusing on the distribution of errors. Percentage errors that are at most 25% of the price are then examined on a yearly basis, to observe the evolution of multiples' performance through the entire period under scrutiny. In addition, the effects on relative valuation and on multiples' efficiency caused by the introduction of the Euro in 2001 and by the 2007/2008 financial turmoil will be analysed. Finally, all the results will be summarised providing a ranking, which suggests the multiples to prefer and the ones to avoid for each subgroup of banks.

4.3.1. General Overview

There are some results that are common among all the subgroups analysed, so they are summarised here in order to avoid redundancy.

– The use of multiples is much more precise for American banks than for European ones, as highlighted by valuation accuracy and measures of performance. This can be easily explained by the negative effects deriving from wide heterogeneity among European peers. In fact, the Eurozone includes countries with strong differences, namely in culture, financial education, regulation and stock markets. Moreover, the higher performance of multiples can be related to the fact that market-oriented financial system, like the American one, show a stronger demand for value relevant accounting information and to the higher capital markets efficiency, which distinguish the U.S. from

[4] Winsorization is a statistical technique that substitutes values exceeding a certain threshold (in this case, the 95th percentile) with the threshold itself. It is preferred to simple trimming because thanks to Winsorization no observation is lost and the original size of the sample is always maintained (Kokic & Bell, 1994).

[5] Kernel Density estimation is a non-parametric way to estimate the probability density of a random variable. Heuristically, it is an adjusted histogram in which "boxes" are replaced by smooth "bumps" (Silverman, 1986). Smoothing is done using a Kernel weighting function that puts less weight on observations that are further from the point being evaluated. The Normal Kernel weighting function is computed according to the following formula:

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right),$$

where u is the argument of the Kernel function.

[6] The bandwidth controls the smoothness of the density estimate: the larger the bandwidth the smoother the estimate. Although there is no general rule for the appropriate choice of the bandwidth, Silverman (1986) makes a case for undersmoothing by choosing a somewhat small bandwidth, since it is easier for the eye to smooth than it is to unsmooth. The same approach has been here used in order to give a clear representation.

Continental Europe, as also highlighted by Herrmann and Richter (2003).

– Large Commercial Banks show the stronger multiples' precision in the U.S., with marked differences with the other subsamples. In Europe, these differences in predictability are less evident between Large Commercial Banks and Investment Banks, apart from P/BV and P/TBV multiples that clearly show higher accuracy for Investment Banks. Conversely, multiples of Small Commercial Banks show the lowest level of accuracy in Europe, while the worst performers in the U.S. are Investment Banks, suggesting that these institutions should be valued with more caution, in particular when selecting comparables.

– In every subsample, forward P/Es are markedly better indicators of any other multiple. They consistently show the highest level of accuracy and performance. This result was actually expected, since prices should reflect future expectations. For instance, compared to reported earnings, analysts' earnings forecasts provide a more direct estimate of future profitability and, since they reflect a larger information set, they are likely to be more accurate (Nissim, 2011). Moreover, I/B/E/S forecasts obviously exclude impacts of extraordinary events, providing a sustainable proxy for permanent core earnings that should therefore persist in the future.

– In line with what theoretically hypothesized by Yee (2004), multiples based on two years forward forecasts of earnings are generally more precise than the ones using one year forecasts. The only exception is the European Small Commercial Banks subsample, where the latter delivers slightly better results. Considering historical P/Es, in Europe the one excluding extraordinary items appears to perform slightly better, but there are no marked differences to take a strong position. Conversely, in the U.S., this difference is more evident, suggesting the use of diluted earnings excluding extraordinary items, in fact this choice should reduce the volatility of book value and mitigate potential accounting distortions.

– Among practitioners, it is a common practice to prefer P/TBV to P/BV, since the tangible book value, which is a more liquid representation of book value, is considered less biased and more accurate for the banking sector. Interestingly, the analysis here performed evidences opposite results with P/BV always showing smaller valuation errors than P/TBV. American Investment Banks are the unique exception, where at 10% accurateness P/TBV gets the 7.0% of banks while P/BV the 6.2%. However, if the precision bound is relaxed to higher value, P/BV always outperforms P/TBV. Moreover, the accuracy of book value multiples is particularly low for Large and Small Commercial Banks in Europe, indeed at 10% accurateness they get approximately 3.0% of banks. Looking at these errors more in-depth, they are particularly high during crises and may have been driven by a large number of outliers, for instance banks consistently trading below book value in countries such as Cyprus, Greece, Italy and Spain. High levels of heterogeneity in Europe, boosted by different governments' responses to the financial crisis, suggest to attentively selecting comparables when using these multiples that can anyway deliver sufficient accuracy, as results for American banks demonstrate.

– Considering the two alternative ways of computing multiples based on dividends, results show that the use of common dividends should be preferred, in particular in Europe. While, in the U.S., differences in performances are less evident. For these reasons, there is not a real connection between preferred dividends and value, in fact, in some extend they can be compared to extraordinary items and, therefore, they should be excluded when valuing a company. Moreover, the analysis of multiples' performances through time shows that P/Common Dividends always follows an individual path, delivering poor accuracy but being enough stable. It suggests that dividends are not the best fundamental to use and that they can potentially be misleading.

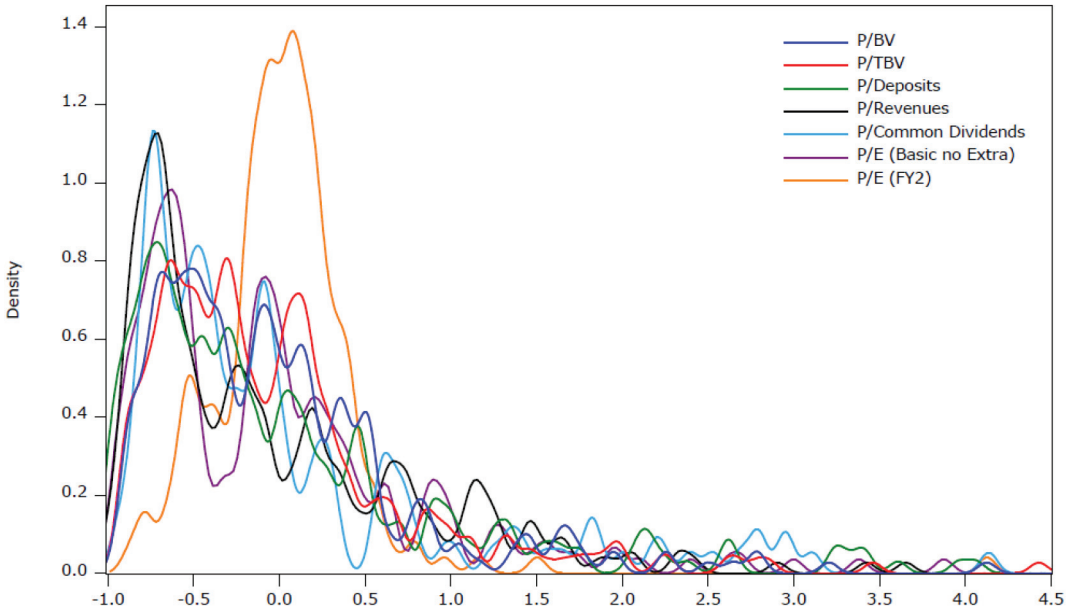
– Both multiples based on deposits and revenues do not show interesting levels of accuracy, in particular in Europe, where they are characterised by high asymmetry. However, performances of these multiples when valuing American Large Commercial Banks is quite satisfactory, in particular in recent years. Nevertheless, it has to be considered that they are consistently overperformed by multiples based on other value drivers. On the one side, the role of deposits within banks has become less crucial in recent years since their business model is shifting towards the offering of many different services disentangled from deposits collection. On the other side, revenues can be strongly misleading since they should be compared with asset-side measures and the level of risk underlying the activities generating these revenues in not considered.

4.3.1. European Investment Banks

Compared to the other European subgroups and considering the lower number of observations, performance of relative valuation for European Investment Banks is quite satisfactory. In general, the distribution of errors (Graph 1) appears noisy and asymmetric, apart from P/E (FY2). In fact, forward multiples are the ones better performing in this subsample and should always be preferred to trailing P/Es, which are anyway an acceptable second best option. Valuation accuracy for P/E (FY2) reaches 27.9% at a 10% level, while bias, MAD and MSE are very limited. Moreover, P/BV and P/TBV perform particularly well if compared

with the other European banks and show a low MSE. Conversely, P/Revenues and P/Deposits are not among the best performers, but still they work better than for the other European Banks (in particular Small Commercial Banks). Multiples based on dividends show the higher levels of bias, MAD and MSE, signalling high volatility and the presence of many outliers, as it is also evidenced by the distribution of errors.

Graph 1 – Distribution of errors for European Investment Banks



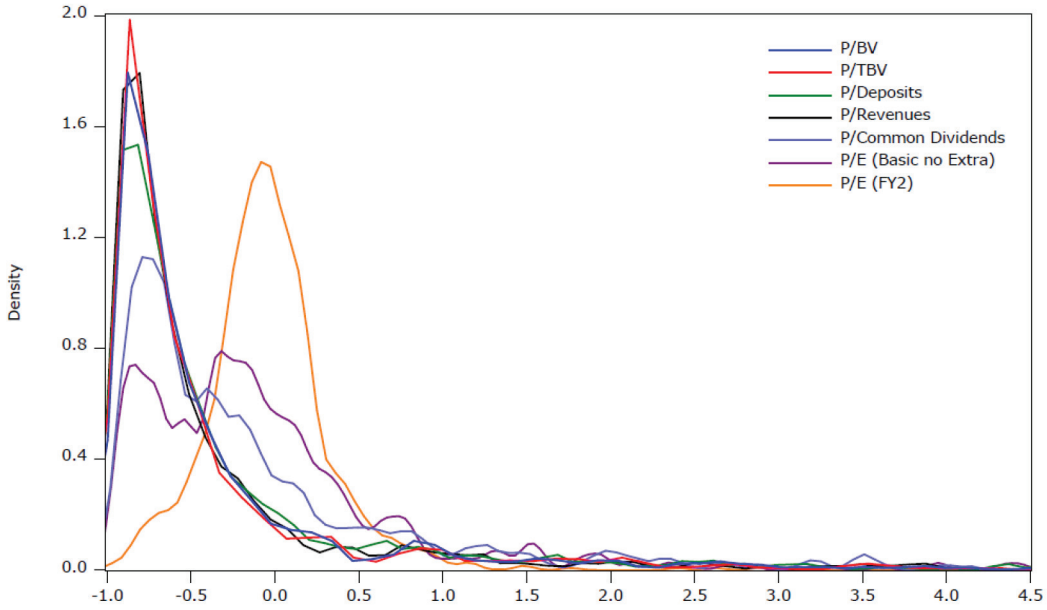
Note: Normal Kernel Density (bandwidth=0.05).

The high level of heterogeneity in Europe has been highlighted as a potential driver of inaccuracy. However, this case is characterised by sufficient homogeneity (due to the restrictive selection process resulting into a low number of banks included in this group) that plays a positive role: errors are overall better distributed than in the other European subsamples.

4.3.2. European Large and Small Commercial Banks

Earnings multiples are the most important value driver for European Large Commercial Banks, in particular, at a 10% level, P/E (FY2) can predict the 27.9% of banks' prices, while P/E (Basic no Extra) the 11.1% only. The apparent bell shaped distribution of errors (Graph 2) for P/E (FY2) highlights the presence of few outliers, being quite gratifying.

Graph 2 – Distribution of errors for European Large Commercial Banks

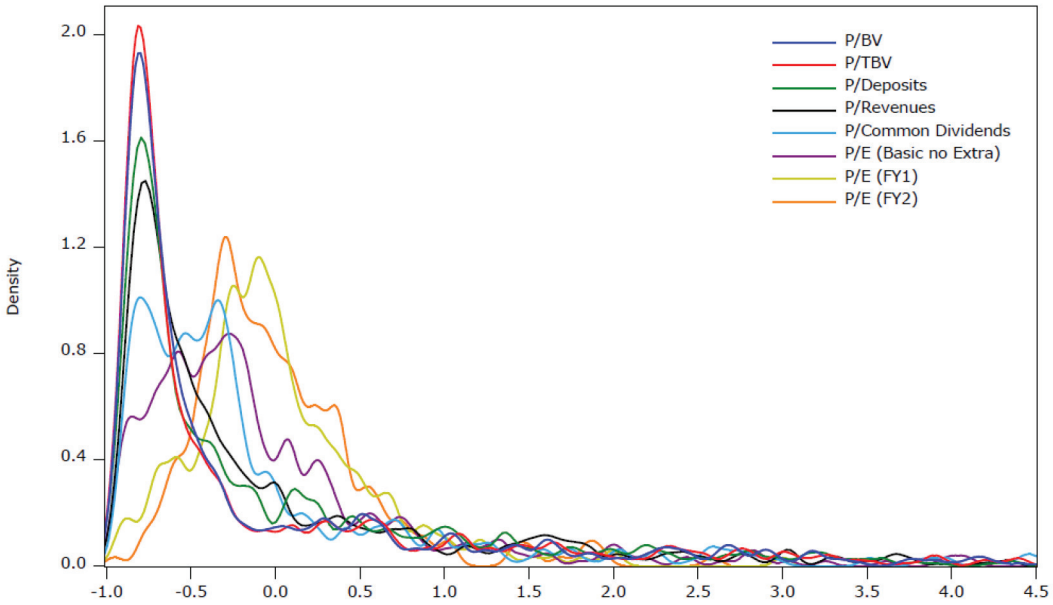


Note: Normal Kernel Density (bandwidth=0.05).

However, the distribution of errors for the other multiples brings to opposite considerations. Apart from P/Common Dividends, which shows discrete accuracy and performance indicators, multiples based on BV, TBV, Deposits and Revenues are affected by a strong left asymmetry as highlighted by their distribution peaked at very negative values. Indeed, the 75th percentile is negative for all these multiples, indicating that more than the 75% of the observations are below zero. Additionally, the fact that bias of these multiples registers the highest positive values signals that there are few, but very big, positive outliers. Indeed, MAD is three times bigger than bias and MSE registers the highest values. Overall, apart from earnings and, in particular, forward ones, the other value drivers do not deliver positive results and they work better for European Investment Banks.

Focusing on European Small Commercial Banks, considerations are even worse, as a first look at the distribution of errors (Graph 3) communicates.

Graph 3 – Distribution of errors for European Small Commercial Banks



Note: Normal Kernel Density (bandwidth=0.05).

This is the unique subsample where P/E (FY1) delivers the highest accuracy: at a 10% level it predicts 20.2% of prices against 17.1% of P/E (FY2), but looking at bias, MAD and MSE, values are lower for the latter. The distribution of errors for both forward earnings multiples is quite similar. The one of P/E (FY1), on the one side, seems to show higher density for values closer to zero, on the other side, big negative errors appear more frequent. Furthermore, the less reliable multiples are the ones based on BV and TBV, registering the highest MSE and very bad accuracy, which is lower than 3.0% at a 10% level. Moreover, the median multiples stand respectively at 0.67x and 0.70x, suggesting high risk of undervaluation. In fact, left asymmetry is extremely evident: the distribution of errors is peaked at very negative values. However, it appears to be lower than the case for large banks, in fact the 75th percentile takes on positive values because of the presence of a higher number of positive observations. Moreover, also large outliers are more frequent, as confirmed by the bumps in the right tails. Errors of P/Common Dividends are better distributed, but still performance and accuracy are quite low.

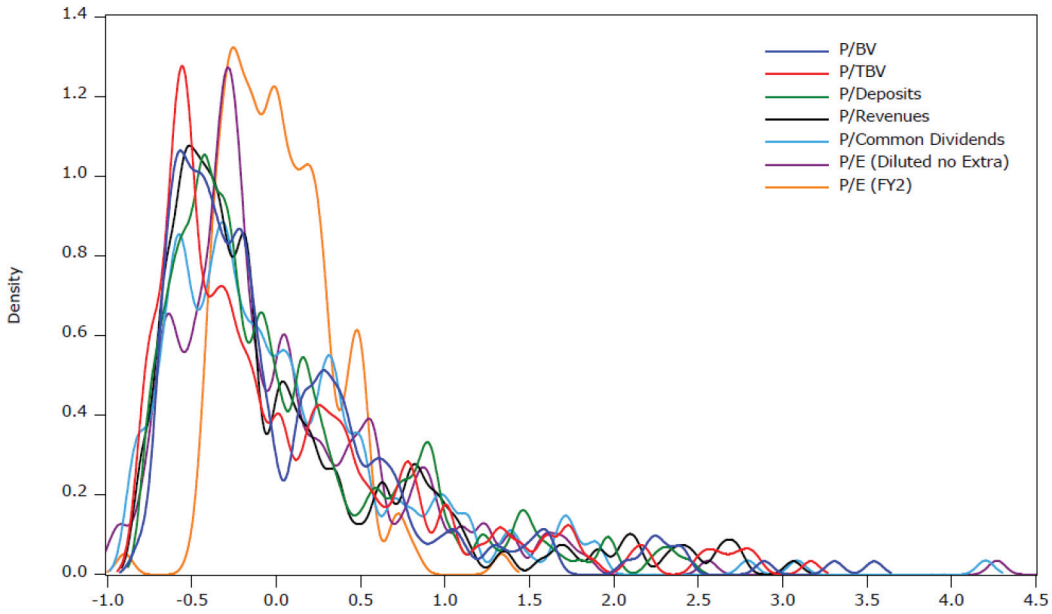
Overall, these results suggest that multiples should be used more as a confirmatory tool than as primary valuation methodology for European Commercial Banks. Unique exception is the use of forward earnings multiples for Large Commercial Banks. While results for Investment Banks are quite acceptable. However, it is important to recall that the selection of comparables can have significant impacts in this case and can potentially deliver stronger results.

To sum up, it is advisable to prefer forward P/E multiples since they deliver more precise values than any other multiple. However, trailing multiples could be considered a second best option when forecasts are not available.

4.3.3. American Investment Banks

Accuracy of forward P/Es is quite satisfactory for American Investment Banks, showing substantially similar results when using FY1 and FY2 earnings (at a 10% level, both predict 23.2% of prices), while MAD and MSE are lower for the latter measure. Apart from P/TBV, which shows low performances and a strong left asymmetry, with high values for MAD, MSE and median (in absolute value), the other multiples are characterised by a relative homogeneous distribution of errors (Graph 4) and similar performances.

Graph 4 – Distribution of errors for American Investment Banks



Note: Normal Kernel Density (bandwidth=0.05).

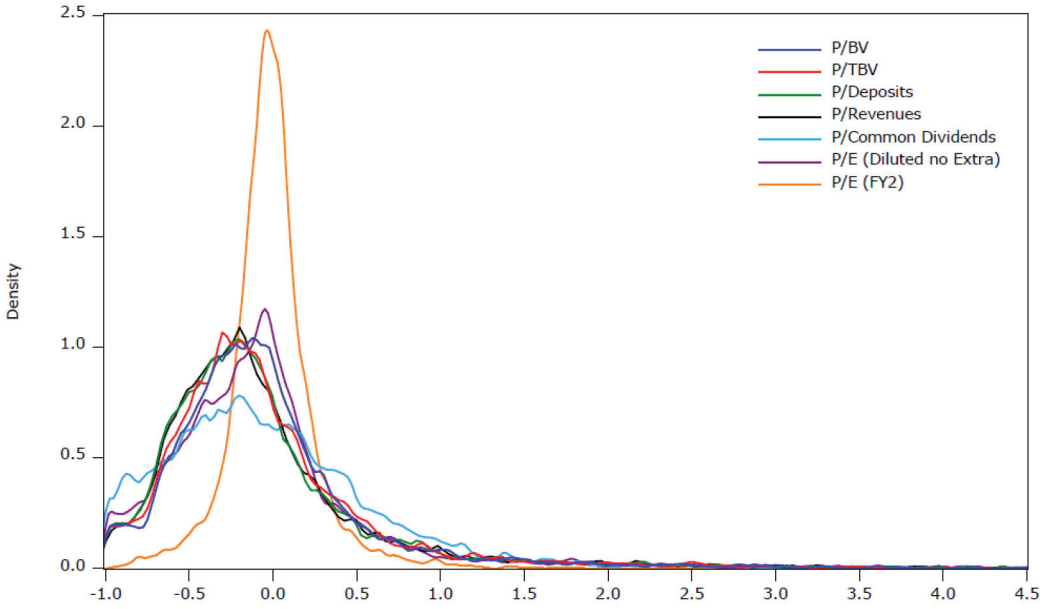
P/Deposits and P/Common Dividends work well registering respectively a 10.3% and 11.5% precision at a 10% level, doing even better than the 10.1% of P/E (Diluted no Extra). These results are anyway worse than the ones registered in the other American subsamples, but they are still better than Europeans' ones. P/TBV and P/Revenues are among the most volatile measures, suggesting their poor reliability and demystifying again the widespread preference of TBV over BV.

Overall, the distribution of errors may result messy and perfromances not convincing if compared to the other American subsamples. However, the analysis of these banks brings to stronger considerations with respect to European ones, enlighting higher suitability of multiples in the United States.

4.3.4. American Large and Small Commercial Banks

The high number of observations (ranging between a maximum of 6,337 for P/BV and a minimum of 5,197 for P/E (FY2)) collected for American Large Commercial Banks, allows to make considerable comments. Multiples for this group of banks show impressive results and, above all, the outstanding performance of P/E (FY2) deserves particular attention. Valuation accuracy stands at 44.6% at a 10% level and it reaches 78.5% and 93.6%, if the accuracy level is relaxed respectively to 25% and 50%. These numbers underline the strong power and the limited size of errors deriving from the use of this multiple. Errors (Graph 5) are overall well distributed and the distribution of P/E (FY2) appears to be bell shaped, really peaked to zero and with relative thin tails. Bias is practically zero, while MAD and MSE are extremely low. These numbers confirm the small magnitude and dispersion of errors when using forward earnings (also results of P/E (FY1) are very similar to these).

Graph 5 – Distribution of errors for American Large Commercial Banks



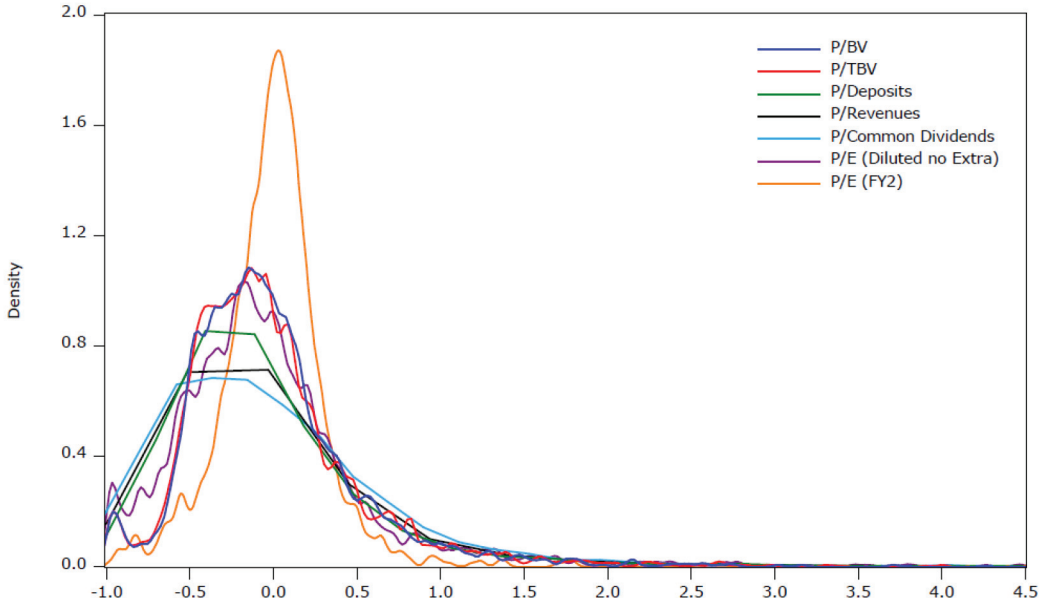
Note: Normal Kernel Density (bandwidth=0.025).

The performance of the other multiples is quite satisfactory too. Bias is extremely low for every multiple, but it benefits from the high number of observations. Trailing P/Es and P/BV can be considered the best alternatives, in fact valuation accuracy at a 10% level stands respectively at 20.6% (diluted earnings excluding extraordinary items) and 18.1%. The remaining multiples, despite the very acceptable levels of accuracy, show higher volatility, as confirmed by MAD and MSE, in particular for multiples built using dividends and deposits. Results confirm, once again, that P/BV should be prefered over P/TBV, that dividends can be easily manipulated, generating distorsions in value, and that it is advisable not to use revenues and deposits as first choice while selecting value drivers.

Despite the lower number of observations for American Small Commercial Banks (in this case, ranging between a maximum of 4,510 for P/BV and a minimum of 1,415 for P/E (FY2)), results are still remarkable. P/E (FY2) produces errors that lie within 10% of price in 33.9% of the cases (which increases to 69.0% and 90.2% relaxing the pricision bound to 25% and 50% respectively). Bias, MAD and MSE are greater than the ones registered in the previous group, but still very limited. Moreover, in this case, the performance of P/E (FY1) is substantially lower than the one of P/E (FY2): bias, MAD and MSE are more than two times bigger and valuation accuracy loses more than 10 percentange points when considering the strictier precision bounds. However, the most impressive results come from P/BV and P/TBV, which deliver the best results among all the subsamples analysed. Indeed, valuation accuracy at a 10% level for these multiples reaches respectively 19.5% and 18.8%, overperforming all the other multiples, including trailing P/Es. Bias is very small and close to the one of P/E (FY2), while MAD and MSE are greater, but among the lowest. Data confirm also the quite stronger performance of P/BV over P/TBV.

The distribution of errors (Graph 6) confirms these findings, with a nice distribution peaked to zero for P/E (FY2) errors. Moreover, P/BV and P/TBV are confirmed as a second best option. Once again dividend multiples are affected by the highest value of MSE, showing high variability and the presence of many outliers. Also accuracy is pretty low, ranking them as the less reliable multiples. Multiples based on revenues and deposits, show acceptable levels of accuracy, but their MSE rank among the highest.

Graph 6 – Distribution of errors for American Small Commercial Banks



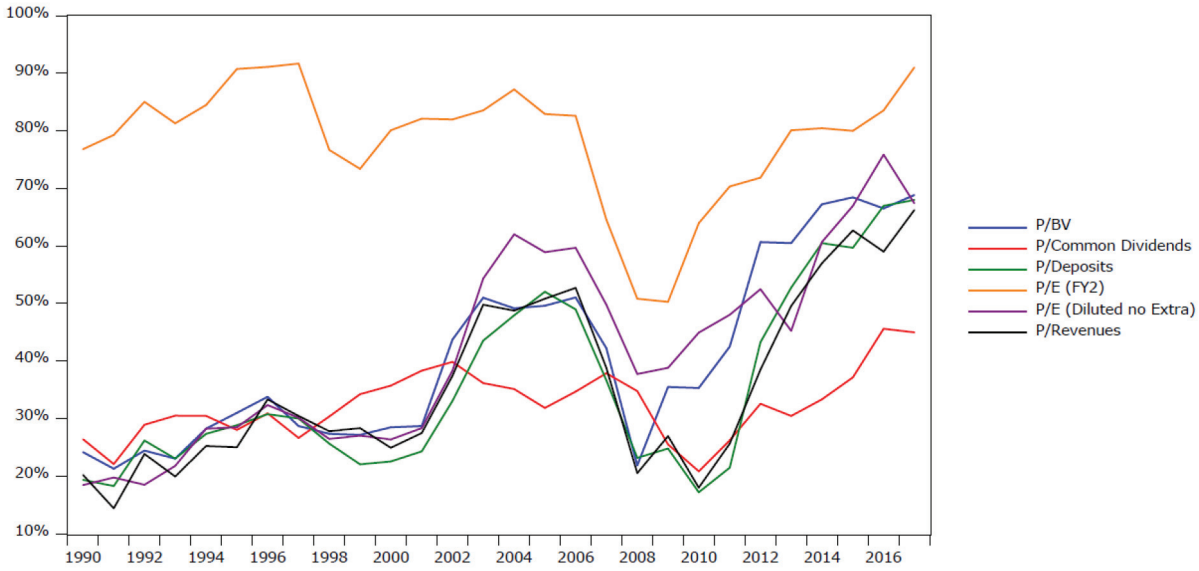
Note: Normal Kernel Density (bandwidth=0.025).

4.3.5. Historical Yearly Performance

This section analyses the evolution of multiples' accuracy on a yearly basis. Errors that lie within 25% of market prices are collected for each subsample in order to observe their yearly evolution. It is important to notice that, because of the split on a yearly basis, the smaller samples, in particular American and European Investment Banks, may show missing data or not reliable figures since very few measures are available in some years (this is why the number of banks under scrutiny is not constant over years because of delisting, new listing o simply availability of data). For these reasons, graphs and results will be commented only for the more relevant subsamples. Moreover, to lighten the chart, not all multiples are included in the graphical respresentation.

The most communicative representation is the one for American Large Commercial Banks (Graph 7). The strong performance of multiples and the high number of observations allow to get important intuitions.

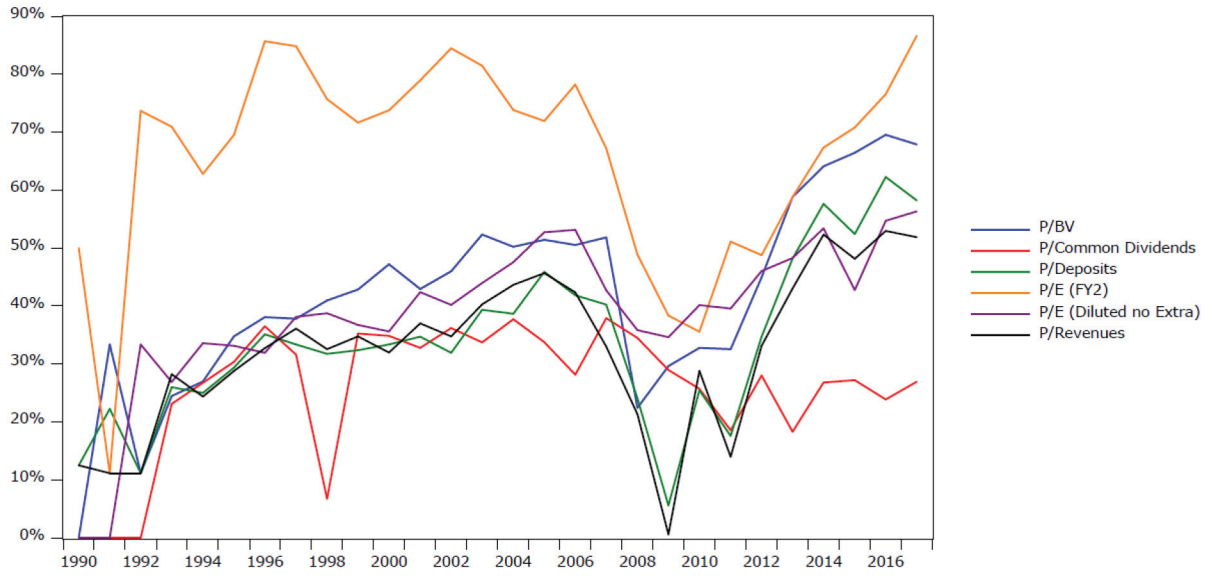
Graph 7 – Yearly multiples' performances for American Large Commercial Banks



Firstly, it is not so surprising to appreciate the excellent accuracy delivered by P/E (FY2) that every year overperforms all the other multiples. More interesting is the level of correlation between all the multiples, excluding P/Common Dividends that is quite stable and follows an individual path, which suggests similar reactions of multiples' performance to the same events. Moreover, it is evident that performances of multiples are negatively impacted around 2000, because of the explosion of the "dot-com bubble". Whereas, the effects of the 2007/2008 financial crisis are definitely more evident, signalling a huge decrease of accuracy. This means that the reaction of prices to the financial crisis was not homogeneous among banks in this subsample. After the crisis, it can be easily observed a recovery of performances, with multiples reaching interesting levels of accuracy in the more recent years. Indeed, it is remarkable to observe the low level of accuracy characterising all the multiples apart from P/E (FY2) during the Nineties, which instead, nowadays, is reaching very high levels. The combination of these elements suggests that restricting the analysis only to more recent years would definitely deliver stronger results than the one already achieved for American Large Commercial Banks. Suitability of multiples for these institutions is again confrmed.

Results for American Small Commercial Banks are messier, in particular during the first years, because very limited observations and a lack of data (Graph 8). However, general considerations are mostly similar.

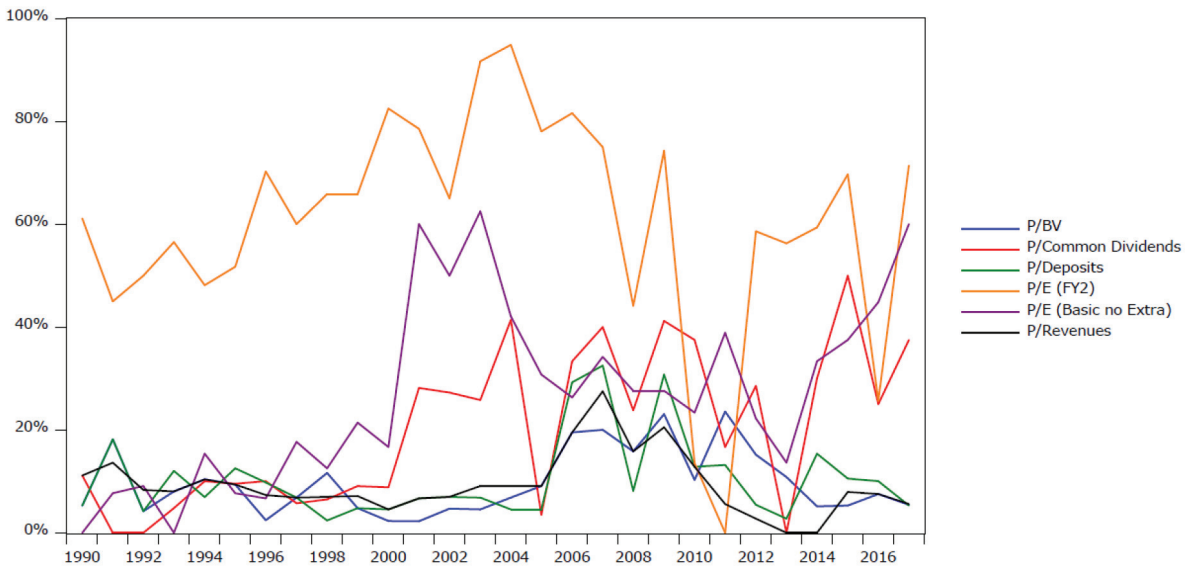
Graph 8 – Yearly multiples' performances for American Small Commercial Banks



The effects of the "dot-com" bubble are less evident here, but interestingly the performance of P/Common Dividends drops significantly in that period. It may be the consequence of different responses among banks in terms of dividend policy, (either in negative terms, because of accumulated losses and no availability of funds or in positive terms, to engage investors), generating high errors when basing valuation on dividends' multiples only. Conversely, the effects of the 2007/2008 financial breakdown are more visible, with the performance of deposits and revenues multiples dropping almost to zero. This is again a sign of heterogeneity among prices' reactions. However, performances registered in the last years, apart from P/Common Dividends suggests good applicability of multiples, in particular for P/E (FY2) and P/BV. In the end, correlation between movements is less evident in this subsample.

Finally, results for European Large Commercial Banks, the sole European subsample with a reasonable number of observations, are presented (Graph 9). The graph shows significant randomness and volatility among years, which is related to the already discussed low suitability of multiples for this group of banks, mainly due to high heterogeneity.

Graph 9 – Yearly multiples' performances for European Large Commercial Banks



Accuracy is not particularly brilliant, apart from earnings multiples which have already been identified as the most reliable for this group of banks. The positive effects of the introduction of the Euro, in particular for earnings multiples, can be here appreciated. However, they will be better scrutinised in the next section. Conversely, the effects of the "dot-com" bubble are quite negligible in this case, while the ones of the 2007/2008 are remarkable. Moreover, the graph shows relevant randomness during the period following this crisis. It may be mostly related to the sovreign-debt crisis, which strongly affected Portugal, Italy, Ireland, Spain and Greece. It is important to recall that during this period banks suffered extreme losses and many of them were bailed-out. Interestingly, the performance of P/E (FY2) drops to zero in 2011, low realiability of forecasts, due to the uncertainty of the economic envirnoment, and poor comparibility among banks, driven by the different economic conditions between Southern and Nothern Europe countries, are probably responsible of this negative impact. Overall, variability of multiples' performances across years is not negligible and their reaction during distressed periods may be significant.

4.3.6. The Effects of the Introduction of the Euro

The Euro was physically introduced as a common currency on the 1st of January 2002, while it was first created on the 1st of January 1999. From that moment on, the European Central Bank started operating to unify monetary policies across the member states.

The investigation here performed aims at understanding whether the implementation of these changes generated significant impacts over relative valuation performance. To implement this analysis, the period 1997-2006 has been selected to avoid distortions related to periods of crisis. After, it has been split into two 5-years subperiods: 1997-2001 (pre-Euro introduction) and 2002-2006 (post-Euro introduction). Therefore, the separation point coincides with the effective date in which the Euro started to circulate. Moreover, not all the banks included in the European dataset have been considered, since for many of them the introduction of Euro came later. Therefore, data were filtered considering only the 12 countries[7] that in the first place implemented together the project of having a common currency. Performances of multiples in the two different periods have been then computed, to observe the sign and the size of potential differences.

Looking at the results, it is clear how multiples' accuracy benefited from the implementation of a common currency for Large Commercial Banks (showing no worsening) and Investment Banks. Given their size and business model, these institutions are more likely to operate in different countries, which are characterised by a level playing field after the introduction of a common currency. Moreover, the introduction of

[7] Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain.

a common currency limited currency risks for banks, making also cheaper the access to capital markets. In addition, a unique monetary policy, setting a stable inflation target, drove interest rates down and increased their stability. Being value inversely proportional to the level of interest rates, these effects made value estimation easier and less discretionary, increasing multiples' performances. The overall level of homogeneity among comparables therefore increased, and this was particularly true for Large Commercial Banks and Investment Banks, which generally do not compete at a national level, being more geographically diversified. Finally, it is important to underline that the rank among multiples for these institutions does not show significant changes between the two periods analysed, apart from P/BV, which was the worst performer for Investment Banks before 2002. While forward P/E multiples are always the best performers and, interestingly, are the ones showing better improvements.

Conversely, the effects on Small Commercial Banks are less clear. This can be explained considering that these institutions are generally more focused on regional markets and compete on a national level. P/BV, P/TBV and dividends multiples show a decrease in accuracy, while earnings multiples register an increase.

4.3.7. The Effects of the 2007/2008 Financial Crisis

The worldwide effects of the 2007/2008 crisis were impressive: many banks went bankrupt and governemnt bail-outs were often necessary forcing them to run budget deficits, stock prices plummeted and unemployment surged, affecting every industry. It was the beginning of a global economic recession and of a long lasting sovreign-debt crisis in Europe: nowadays some Southern Europe countries still have to fully recover from the crisis. Taking Italy as an example: GDP growth still shows weak positive signs and the implementation of policies to boost growth and reduce government debt appears difficult to be achieved as long as political instability remains there.

This section focuses on the analysis of the effects of the financial crisis on valuation accuracy. The period 2003-2012 has been selected and it has been split into two 5-years subperiods: 2003-2007 (pre-financial crisis) and 2008-2012 (post-financial crisis). Therefore, the separation point coincides with Lehman Brothers' collapse. Performances of multiples in the two different periods are then computed to observe the sign and the size of potential differences.

The financial crisis had different effects on European Small Commercial Banks, depending on their level of international exposure and on their dependence on mortgages. In general, accuracy was badly hit, apart from trailing multiples that did slightly better. This can be related to market operators relying more on realized results than on expected results, which were surrounded by huge uncertainty.

European Large Commercial Banks, show a strong decrease in forward earnings multiples' accuracy, losing more than 20 and 45 percentage points at a 10% and 25% level respectively. Notwithstanding these multiples are still the best performers. Trailing multiples show a slightly better accuracy, but only at a 10% level. Interestingly, solvency-based multiples (P/BV and P/TBV) show greater accuracy, underlying analysts' attention on the level of capitalization of banks during the crisis. Indeed, large banks were largely affected by deteriorated and non-performing exposures, forcing them to account for massive write-downs. In addition, P/Common Dividends performed slightly better during the crisis, highlighting a focus on the capability of distributing dividends more than on the uncertain earnings achievable in the future.

Results for European Investment Banks, which, by definition, are strongly dependent on financial markets performances, show mostly a worsening of multiples' accuracy. P/E (FY2) is the multiple suffering the most in performances: it registers a decrease in accuracy at a 10% level higher than 40 percentage points. Small improvements are achieved when using P/Common Dividends and P/TBV and can be justified by market operators shifting their focus on cash flows and levels of capitalization.

Moving to American Commercial Banks, results show a strong worsening in multiples accuracy, highlighting how impressive the magnitude of the subprime market was in the United States. Both large and small commercial banks show negative performances for all the multiples, with the ones based on forward earnings suffering the most. Also American Investment Banks were affected by a decrease of the performance of almost every multiple and, in particular, of the ones based on earnings. Conversely, P/Revenues and P/Deposits show interestingly higher levels of accuracy during the crisis. An increase in the relevance of revenues may be related to analysts being more focused on the capabilities of these institutions in generating fees from a frozen M&A market and gains from extremely volatile stock markets. While the performance of P/Deposits is quite unexpected, since deposits are not a relevant measure for investment banks. This may imply that investment banks' business model was shifting to the one of retail banks. As a matter of fact, few days after the collapse of Lehman Brothers, the two major pure American Investment Banks, i.e., Goldman Sachs and Morgan Stanley, confirmed to become traditional bank holding companies, bringing an end to the era of pure investment banking on Wall Street.

4.3.8. Multiples' Ranking

We analyse multiples' performance during the last 28 years, mainly considering multiples' accuracy and the distribution of errors across different subsamples. Moreover, multiples' performance has been attentively scrutinised on an yearly basis, with a particular focus on specific periods (i.e., the introduction of the Euro and the 2007/2008 financial crisis) to catch performance's reactions. Combining the main findings deriving from these analyses here performed, it is possible to stipulate a ranking to show which multiple represents the best choice for each subsample. This can be considered as a useful summary tool for analysts when choosing the best multiples to perform relative valuation.

Supremacy of earnings multiples, in particular forward ones, is evident from the summary table (Table 2). Therefore, forward earnings should always be the first choice, while trailing earnings can be an acceptable alternative, only when forecasts are not available. Earnings excluding extraordinary items and including dilutive effects, when available, should always be preferred among the different measures of earnings. American Small Commercial Banks and Investment Banks are the unique exceptions, with P/BV and P/TBV, for the former group, and P/Common Dividends, for the latter, ranking above trailing earnings.

Table 2 – Multiples' Summary Ranking

		EU Investment Banks	EU Large Commercial Banks	EU Small Commercial Banks	US Investment Banks	US Large Commercial Banks	US Small Commercial Banks
Best choice	1	P/E (FY2)	P/E (FY2)	P/E (FY1)	P/E (FY2)	P/E (FY2)	P/E (FY2)
	2	P/E (FY1)	P/E (FY1)	P/E (FY2)	P/E (FY1)	P/E (FY1)	P/E (FY1)
	3	P/E (LTM Basic no Extra)	P/E (LTM Basic no Extra)	P/E (LTM Basic no Extra)	P/Common Dividends	P/E (LTM Diluted no Extra)	P/BV
	4	P/BV	P/Common Dividends	P/Common Dividends	P/E (LTM Diluted no Extra)	P/BV	P/TBV
	5	P/TBV	P/Deposits	P/Revenues	P/Deposits	P/TBV	P/E (LTM Diluted no Extra)
	6	P/Common Dividends	P/Revenues	P/Deposits	P/Revenues	P/Deposits	P/Revenues
	7	P/Deposits	P/BV	P/BV	P/BV	P/Revenues	P/Deposits
Worst choice	8	P/Revenues	P/TBV	P/TBV	P/TBV	P/Common Dividends	P/Common Dividends

Book value multiples work well also for American Large Commercial Banks and European Investment Banks, while they rank last in the other subsamples. However, it is important to recall that this low performance was magnified by the high levels of heterogeneity of European Commercial Banks. Moreover, P/TBV always ranks below P/BV, which highlights that the common practice of using a tangible measure of book value, eliminating intangibles, has no practical relevance. The remaining multiples, P/Common Dividends, P/Deposits and P/Revenues, never rank among the top performers, however their use could be still acceptable in some subsamples (i.e., European Commercial Banks and American Investment Banks), should other measures be not available. P/Common Dividends, as unique exception, ranks third for American Investment Banks, but at the same time it ranks last for American Commercial Banks, confirming distortions that can derive from the misuse of dividends.

5. Conclusions

We study the relative valuation accuracy of 1,118 listed and delisted banks across the United States and the Eurozone from 1990 to 2017. Multiples deliver strong valuation accuracy in the U.S. (in particular for Large Commercial Banks), while in Europe results are less univocal. Multiples based on forward earnings are the best performers and the one based on two years forecasts are the most accurate. However, the use of trailing earnings is quite often a valid second best option and diluted earnings not including extraordinary items should always be preferred among earnings measures. Despite practitioners consider P/TBV more reliable than P/BV, results show the opposite with the latter consistently overperforming the former. Moreover, the performance of these solvency-based multiples is very low in Europe, while they work quite well for American Commercial Banks. A very weak relationship between value and the amount of preferred dividends is also revealed with P/Common Dividends being a more precise tool than P/Total Dividends. Multiples based on Revenues and Deposits do not show particularly interesting performances.

We also investigate the historical performance of multiples. The effects of the financial crisis appear strongly negative in every subsample, while performances registered in recent years are at the highest levels. On the one hand, American Large Commercial Banks confirm the strong accuracy of P/E (FY2) that every year overperformed the other multiples. On the other hand, they register an increasing performance in the last 5 years for the multiples based on book value, trailing earnings, deposits and revenues. Finally, precision of different multiples appears to move in a correlate way, while P/Common Dividends tends to follow a proper path.

APPENDIX A – The 2007/08 Financial Crisis: Performance Statistics

	Pre-Financial Crisis (2003-2007)					Post-Financial Crisis (2008-20012)					Difference				
	Valuations within					Valuations within					Valuations within				
EU Investment Banks	10% of price	25% of price	50% of price	75% of price	90% of price	10% of price	25% of price	50% of price	75% of price	90% of price	10% of price	25% of price	50% of price	75% of price	90% of price
P/BV	20,8%	37,7%	73,6%	96,2%	96,2%	16,7%	29,2%	52,1%	87,5%	95,8%	-4,1%	-8,6%	-21,5%	-8,7%	-0,4%
P/TBV	13,2%	34,0%	73,6%	94,3%	94,3%	20,8%	37,5%	58,3%	87,5%	91,7%	7,6%	3,5%	-15,3%	-6,8%	-2,7%
P/Revenues	9,4%	20,8%	43,4%	77,4%	83,0%	2,1%	17,0%	44,7%	72,3%	85,1%	-7,3%	-3,7%	1,3%	-5,0%	2,1%
P/Deposits	10,0%	28,0%	56,0%	72,0%	78,0%	8,5%	31,9%	46,8%	68,1%	87,2%	-1,5%	3,9%	-9,2%	-3,9%	9,2%
P/Common Dividends	12,5%	35,0%	50,0%	85,0%	85,0%	15,4%	23,1%	53,8%	92,3%	92,3%	2,9%	-11,9%	3,8%	7,3%	7,3%
P/Total Dividends	12,5%	30,0%	52,5%	85,0%	85,0%	11,1%	33,3%	50,0%	72,2%	83,3%	-1,4%	3,3%	-2,5%	-12,8%	-1,7%
P/E (LTM Basic no Extra)	20,0%	37,8%	60,0%	84,4%	88,9%	12,1%	18,2%	36,4%	60,6%	81,8%	-7,9%	-19,6%	-23,6%	-23,8%	-7,1%
P/E (LTM Basic with Extra)	15,6%	35,6%	60,0%	84,4%	88,9%	9,1%	18,2%	33,3%	60,6%	81,8%	-6,5%	-17,4%	-26,7%	-23,8%	-7,1%
P/E (FY1)	44,1%	79,4%	88,2%	100,0%	100,0%	32,3%	61,3%	87,1%	93,5%	100,0%	-11,9%	-18,1%	-1,1%	-6,5%	0,0%
P/E (FY2)	54,5%	81,8%	87,9%	100,0%	100,0%	12,9%	54,8%	77,4%	90,3%	100,0%	-41,6%	-27,0%	-10,5%	-9,7%	0,0%
EU Large Commercial Banks															
P/BV	2,8%	11,7%	24,9%	49,3%	84,5%	5,5%	17,5%	47,0%	74,3%	82,5%	2,6%	5,7%	22,1%	25,0%	-2,0%
P/TBV	3,8%	9,9%	23,0%	42,7%	82,6%	4,4%	12,0%	40,4%	70,5%	83,1%	0,6%	2,2%	17,4%	27,8%	0,4%
P/Revenues	4,7%	14,6%	30,0%	53,5%	82,6%	4,8%	11,6%	28,0%	54,5%	83,1%	0,1%	-2,9%	-2,0%	1,0%	0,4%
P/Deposits	6,1%	15,0%	29,1%	51,2%	84,5%	5,8%	14,2%	37,9%	69,5%	80,0%	-0,3%	-0,8%	8,8%	18,3%	-4,5%
P/Common Dividends	10,3%	28,8%	52,1%	70,5%	85,6%	15,3%	30,5%	52,5%	76,3%	83,1%	5,0%	1,7%	0,5%	5,7%	-2,6%
P/Total Dividends	10,0%	28,7%	50,7%	70,0%	84,7%	7,1%	22,4%	41,2%	60,0%	76,5%	-2,9%	-6,3%	-9,5%	-10,0%	-8,2%
P/E (LTM Basic no Extra)	10,7%	35,3%	67,3%	82,0%	93,3%	16,9%	27,4%	52,4%	78,2%	88,7%	6,3%	-7,9%	-14,9%	-3,8%	-4,6%
P/E (LTM Basic with Extra)	12,7%	38,0%	67,3%	83,3%	93,3%	14,3%	26,2%	55,6%	78,6%	88,9%	1,6%	-11,8%	-11,8%	-4,8%	-4,4%
P/E (FY1)	39,4%	79,8%	95,9%	99,5%	100,0%	15,2%	33,8%	64,2%	86,1%	94,0%	-24,1%	-46,0%	-31,6%	-13,4%	-6,0%
P/E (FY2)	42,6%	84,2%	96,8%	100,0%	100,0%	19,0%	37,5%	62,5%	86,3%	95,8%	-23,6%	-46,7%	-34,3%	-13,7%	-4,2%
EU Small Commercial Banks															
P/BV	1,8%	4,9%	18,4%	48,4%	80,7%	0,9%	3,1%	12,1%	42,2%	69,5%	-0,9%	-1,8%	-6,3%	-6,3%	-11,2%
P/TBV	1,3%	4,5%	17,9%	45,7%	80,3%	0,9%	2,7%	12,1%	38,6%	69,5%	-0,4%	-1,8%	-5,8%	-7,2%	-10,8%
P/Revenues	5,5%	10,5%	29,5%	64,1%	81,8%	3,3%	10,5%	28,2%	53,6%	80,0%	-2,2%	0,0%	-1,4%	-10,5%	-1,8%
P/Deposits	4,6%	17,4%	36,1%	60,7%	82,2%	2,8%	9,6%	21,9%	47,0%	75,8%	-1,7%	-7,8%	-14,2%	-13,7%	-6,4%
P/Common Dividends	6,6%	11,9%	44,4%	68,2%	81,5%	3,3%	7,9%	19,2%	33,1%	46,4%	-3,3%	-4,0%	-25,2%	-35,1%	-35,1%
P/Total Dividends	5,7%	11,4%	42,4%	65,2%	81,6%	5,2%	7,6%	19,0%	34,2%	48,1%	-0,5%	-3,8%	-23,4%	-31,0%	-33,5%
P/E (LTM Basic no Extra)	7,4%	23,0%	51,1%	77,0%	83,0%	13,4%	31,1%	60,0%	74,1%	86,7%	6,0%	8,1%	8,9%	-3,0%	3,7%
P/E (LTM Basic with Extra)	7,4%	23,0%	50,4%	77,0%	83,0%	11,1%	28,9%	60,0%	74,1%	86,7%	3,7%	5,9%	9,6%	-3,0%	3,7%
P/E (FY1)	20,0%	46,7%	75,6%	95,6%	95,6%	23,9%	46,7%	66,7%	84,4%	88,9%	3,9%	0,0%	-8,9%	-11,1%	-6,7%
P/E (FY2)	19,6%	37,0%	80,4%	93,5%	97,8%	15,6%	39,1%	76,1%	89,1%	89,1%	-4,0%	2,2%	-4,3%	-4,3%	-8,7%

	Pre-Financial Crisis (2003-2007)					Post-Financial Crisis (2008-20012)					Difference				
	Valuations within					Valuations within					Valuations within				
US Investment Banks	10% of price	25% of price	50% of price	75% of price	90% of price	10% of price	25% of price	50% of price	75% of price	90% of price	10% of price	25% of price	50% of price	75% of price	90% of price
P/BV	16,1%	35,5%	71,0%	87,1%	96,8%	4,0%	32,0%	80,0%	88,0%	88,0%	-12,1%	-3,5%	9,0%	0,9%	-8,8%
P/TBV	12,9%	25,8%	45,2%	74,2%	90,3%	8,0%	20,0%	52,0%	76,0%	84,0%	-4,9%	-5,8%	6,8%	1,8%	-6,3%
P/Revenues	9,7%	22,6%	48,4%	77,4%	90,3%	16,0%	24,0%	52,0%	80,0%	88,0%	6,3%	1,4%	3,6%	2,6%	-2,3%
P/Deposits	12,9%	25,8%	77,4%	87,1%	90,3%	16,0%	44,0%	72,0%	88,0%	88,0%	3,1%	18,2%	-5,4%	0,9%	-2,3%
P/Common Dividends	9,7%	19,4%	51,6%	83,9%	87,1%	9,1%	18,2%	54,5%	72,7%	86,4%	-0,6%	-1,2%	2,9%	-11,1%	-0,7%
P/Total Dividends	9,7%	19,4%	51,6%	83,9%	87,1%	12,0%	16,0%	40,0%	80,0%	88,0%	2,3%	-3,4%	-11,6%	-3,9%	0,9%
P/E (LTM Diluted no Extra)	16,1%	38,7%	83,9%	90,3%	90,3%	9,1%	31,8%	59,1%	77,3%	86,4%	-7,0%	-6,9%	-24,8%	-13,0%	-4,0%
P/E (LTM Diluted with Extra)	16,1%	38,7%	83,9%	90,3%	90,3%	9,1%	27,3%	63,6%	77,3%	86,4%	-7,0%	-11,4%	-20,2%	-13,0%	-4,0%
P/E (LTM Basic no Extra)	16,1%	41,9%	80,6%	90,3%	90,3%	9,1%	31,8%	59,1%	81,8%	86,4%	-7,0%	-10,1%	-21,6%	-8,5%	-4,0%
P/E (LTM Basic with Extra)	16,1%	38,7%	80,6%	90,3%	90,3%	9,1%	27,3%	63,6%	81,8%	86,4%	-7,0%	-11,4%	-17,0%	-8,5%	-4,0%
P/E (FY1)	26,7%	40,0%	80,0%	100,0%	100,0%	12,5%	37,5%	75,0%	93,8%	100,0%	-14,2%	-2,5%	-5,0%	-6,3%	0,0%
P/E (FY2)	26,7%	33,3%	80,0%	100,0%	100,0%	6,3%	31,3%	81,3%	93,8%	100,0%	-20,4%	-2,1%	1,3%	-6,3%	0,0%
US Large Commercial Banks															
P/BV	20,3%	48,7%	77,0%	85,6%	91,4%	16,2%	39,3%	70,2%	84,9%	90,0%	-4,1%	-9,4%	-6,9%	-0,7%	-1,3%
P/TBV	16,6%	41,8%	73,2%	84,2%	90,4%	13,4%	34,4%	67,0%	84,4%	89,0%	-3,1%	-7,4%	-6,2%	0,3%	-1,5%
P/Revenues	20,3%	48,3%	74,5%	84,4%	90,6%	8,2%	26,0%	59,3%	84,2%	89,5%	-12,2%	-22,3%	-15,2%	-0,3%	-1,1%
P/Deposits	18,5%	45,9%	74,0%	83,5%	90,5%	9,5%	26,0%	58,9%	83,7%	90,0%	-9,0%	-19,9%	-15,1%	0,2%	-0,5%
P/Common Dividends	12,6%	35,1%	64,4%	79,5%	87,9%	11,3%	28,2%	52,4%	71,0%	84,3%	-1,3%	-6,9%	-12,0%	-8,4%	-3,6%
P/Total Dividends	13,2%	34,8%	63,9%	78,7%	87,7%	13,5%	33,6%	59,1%	78,7%	88,5%	0,3%	-1,3%	-4,8%	0,0%	0,8%
P/E (LTM Diluted no Extra)	28,7%	57,0%	75,2%	83,7%	90,9%	22,1%	45,2%	69,6%	84,8%	90,7%	-6,6%	-11,8%	-5,6%	1,1%	-0,2%
P/E (LTM Diluted with Extra)	27,9%	56,3%	74,9%	83,7%	90,8%	19,8%	43,7%	69,6%	84,3%	90,8%	-8,0%	-12,6%	-5,3%	0,6%	0,0%
P/E (LTM Basic no Extra)	27,9%	57,2%	75,6%	83,5%	90,8%	20,7%	44,7%	69,2%	84,7%	90,7%	-7,2%	-12,5%	-6,3%	1,2%	-0,2%
P/E (LTM Basic with Extra)	27,5%	56,6%	75,1%	83,6%	90,7%	19,1%	43,5%	69,3%	84,2%	90,7%	-8,5%	-13,1%	-5,8%	0,6%	0,0%
P/E (FY1)	45,5%	77,9%	94,7%	98,3%	99,1%	26,2%	54,5%	78,2%	91,2%	95,4%	-19,3%	-23,3%	-16,4%	-7,1%	-3,6%
P/E (FY2)	47,3%	80,5%	95,2%	98,7%	99,1%	32,3%	61,7%	84,2%	94,6%	97,5%	-15,0%	-18,8%	-11,0%	-4,1%	-1,6%
US Small Commercial Banks															
P/BV	20,3%	51,3%	80,8%	89,7%	92,7%	12,6%	32,3%	68,3%	85,8%	90,2%	-7,6%	-19,0%	-12,5%	-3,9%	-2,4%
P/TBV	20,7%	50,5%	80,5%	89,0%	93,0%	13,2%	33,2%	67,8%	85,7%	89,5%	-7,6%	-17,3%	-12,7%	-3,4%	-3,4%
P/Revenues	15,6%	41,1%	76,3%	87,4%	91,1%	6,7%	19,4%	41,6%	72,5%	88,5%	-8,9%	-21,7%	-34,7%	-14,9%	-2,6%
P/Deposits	15,3%	41,2%	74,6%	86,9%	91,1%	8,2%	21,3%	45,1%	76,5%	90,5%	-7,1%	-19,8%	-29,5%	-10,4%	-0,6%
P/Common Dividends	15,6%	34,2%	61,3%	80,4%	88,0%	13,0%	27,8%	54,8%	79,0%	88,3%	-2,6%	-6,4%	-6,5%	-1,4%	0,3%
P/Total Dividends	15,2%	34,9%	61,2%	79,8%	87,6%	14,0%	33,3%	57,0%	78,4%	89,5%	-1,2%	-1,5%	-4,2%	-1,4%	1,9%
P/E (LTM Diluted no Extra)	20,6%	48,1%	72,7%	83,7%	89,0%	17,5%	39,6%	64,3%	83,2%	90,5%	-3,1%	-8,4%	-8,3%	-0,4%	1,6%
P/E (LTM Diluted with Extra)	20,3%	47,6%	72,1%	83,3%	88,9%	17,2%	39,9%	64,3%	83,2%	90,5%	-3,1%	-7,7%	-7,7%	-0,1%	1,7%
P/E (LTM Basic no Extra)	21,4%	48,7%	72,1%	83,6%	88,9%	17,7%	39,6%	64,5%	82,9%	90,4%	-3,7%	-9,0%	-7,7%	-0,6%	1,5%
P/E (LTM Basic with Extra)	21,5%	47,9%	72,2%	83,3%	88,9%	17,2%	39,6%	64,3%	82,9%	90,5%	-4,3%	-8,3%	-7,8%	-0,4%	1,7%
P/E (FY1)	30,8%	70,8%	93,7%	98,6%	99,2%	18,4%	41,0%	70,9%	88,1%	94,3%	-12,4%	-29,8%	-22,9%	-10,5%	-4,9%
P/E (FY2)	36,3%	74,4%	93,0%	97,9%	99,1%	22,1%	43,8%	75,1%	90,8%	95,6%	-14,2%	-30,6%	-17,9%	-7,1%	-3,5%

Bibliography

Alford, A. W. (1992). The effect of the set of comparable firms on the accuracy of the price-earnings valuation method. *Journal of Accounting Research 30*(1): 94-108.

Allayannis, G., Rountree, B. & Weston, J. P. (2005). Earnings volatility, cash flow volatility and firm value. Working Paper, SSRN.

Barker, R. G. (1999). Survey and market-based evidence of industry-dependence in analysts' preferences between the dividend yield and price-earnings ratio valuation models. *Journal of Business and Accounting 26*(3): 393-418.

Baker, M. & Ruback, R. (1999). *Estimating industry multiples*. Working Paper, Harvard University, USA.

Beltratti, A & Stulz, R. M. (2009). *Why did some banks perform better during the credit crisis? A cross-country study of the impact of governance and regulation*. Working Paper 254, ECGI.

Bhojraj, S. & Lee, C. M. C. (2002). Who is my peer? A valuation based approach to the selection of comparable firms. *Journal of Accounting Research 40*(2): 407-439.

Brennan, M. & Hughes, P. (1991). Stock Prices and the Supply of Information. *Journal of Finance* 46: 1665-1691.

Brooks, C. (2014). *Introductory Econometrics for Finance*. 3rd ed. Cambridge University Press.

Calomiris, C. & Nissim, D. (2007). *Activity-based valuation for bank holding companies*. Working Paper 12918, NBER.

Cheng, C. S. A. & McNamara, R. (2000). The valuation accuracy of the price-earnings and price-book benchmark valuation methods. *Review of Quantitative Finance and Accounting 15*(4): 349-370.

Cooper, I. & Cordeiro, L. (2008). *Optimal equity valuation using multiples: The number of comparable firms*. Working Paper, SSRN.

Damodaran, A. (2002). *Investment Valuation. Tools and Techniques for Determining the Value of Any Asset*. 2nd ed. John Wiley & Sons.

Damodaran, A. (2006). *Damodaran on Valuation. Security Analysis for Investment and Corporate Finance*. 2nd ed. John Wiley & Sons.

Deev, O. (2011). Methods of bank valuation: A critical overview. *Financial Assets and Investing 2011*(3): 33-44.

Dermine, J. (2010). *Bank valuation: With an application to the implicit duration of non-maturing deposits*. Working Paper, INSEAD Business School, France.

Dittmann, I & Maug, E. G. (2008). *Biases and error measures: How to compare valuation methods*. Working Paper, SSRN.

Guatri, L. & Bini, M. (2002). *I moltiplicatori nella valutazione delle aziende*. Volume 1: *Principi e linee guida professionali*, Volume 2: *Metodi e Strumenti*. Università Bocconi Editore.

Forte, G., Gianfrate, G., & Rossi, E. (2018). Does relative valuation work for banks? *Global Finance Journal*, forthcoming.

Forte, G. & Rossi, E. (2016). *Assessing Relative Valuation in Equity Markets: Bridging Research and Practice*. Palgrave Macmillan.

Herrmann, V. & Richter, F. (2003). Pricing with performance-controlled multiples. *Schmalenbach Business Review, 55*: 194-219.

Kokic, P. N. & Bell, P. A. (1994). Optimal winsorizing cutoffs for a stratified finite population estimator. *Journal of Official Statistics, 10*(4): 419-435.

Koller, T.; Goedhart, M. & Wessels, D. (2005). *Valuation, Measuring and Managing the Value of Companies*. 4th ed. John Wiley & Sons.

Lie, E. & Lie, H. (2002). Multiple used to estimate corporate value. *Financial Analysts Journal, 58*(2): 44-54.

Lin, C. (2017). Impact of International Cross-Listing and Delisting on Return Volatility. *Universal Journal of Accounting and Finance* 5(3): 60-71.

Liu, J.; Nissim, D. & Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research, 40*(1): 135-172.

Lokey, O. K.; Braun, R. S. & Cefali, S. L. (1990). Business Valuation. *Litigation Services Handbook: The Role of Accountant as Expert Witness*.

Massari, M.; Gianfrate, G. & Zanetti, L. (2014). *The Valuation of Financial Companies*. John Wiley & Sons.

Minjina, D. I. (2009). Relative performance of valuation using multiples. Empirical evidence on Bucharest Stock Exchange. *The Review of Finance and Banking, 1*(1): 35-53.

Nissim, D. & Penman, S. (2001). Ratio analysis and equity valuation: from research to practice. *Review of Accounting Studies, 6*(1): 109-154.

Nissim, D. (2010). *Analysis and valuation of insurance companies*. Industry Study #2, Center for Excellence in Accounting and Security Analysis, Columbia Business School, USA.

Nissim, D. (2011). Relative valuation of U.S. insurance companies. *Review of Accounting Studies, 18*(2): 324-359.

Ou, J. A. & Penan, S. H. (1995). *Financial statement analysis and the evaluation of market-to-book ratios*. Working Paper, University of California at Berkeley, USA.

Park, Y. S. & Lee, J. (2003). An empirical study on the relevance of applying relative valuation models to investment strategies in the Japanese stock market. *Japan and the World Economy, 15*(3): 331-339.

Pratt, S. P. (1989). *Valuing a Business: The Analysis and Appraisal of Closely-Held Companies*. New York: Dow Jones-Irwin.

Resti, A. & Sironi, A. (2007). *Risk Management and Shareholders' Value in Banking*. John Wiley & Sons.

Schipper, K. (1991). Commentary on Analysts' Forecasts, *Accounting Horizons 5*, 105-121.

Schreiner, A. (2007). *Equity valuation using multiples: an empirical investigation*. Academic Network, Roland Berger Strategy Consultants, Wiesbaden: DUV Gabler Edition Wissenschaft.

Siems, M. (2009). The foundations of securities law. *European Business Law Review, 20:* 141-171.

Stockenstrand, A. & Nilsson, F. (2017). *Bank Regulation. Effects on Strategy, Financial Accounting and Management Control*. Routledge.

Tasker, S. C. (1998). *Industry-preferred multiples in acquisition valuation*. Working Paper, Cornell University, USA.

Yee, K. K. (2004). Forward versus trailing earnings in equity valuation. *Review of Accounting Studies* 9(2): 301-329.

Wu, Y. (2017). *What's Behind the Smooth Dividends? Evidence from Structural Estimation*. Working Paper, University of Illinois at Urbana-Champaign, USA.