JOURNAL OF ECONOMIC SURVEYS

doi: 10.1111/joes.12133

A CLOSER LOOK AT VALUE PREMIUM: LITERATURE REVIEW AND SYNTHESIS

Eero Pätäri* and Timo Leivo

Lappeenranta University of Technology
School of Business and Management

Abstract. This paper provides a systematic review of value premium literature that examines the performance difference between value and growth stocks and the possible reasons for it. We compare and synthesize the results from the different regional stock markets and different sample periods. The literature is categorized according to stock selection criteria that are based on either individual valuation ratios, such as E/P, B/P, D/P, S/P, CF/P, and enterprise value-based multiples, or composite value criteria that aim to capture more than one dimension of relative value simultaneously or combine them with other classification criteria. We also compare the efficacy of various selection criteria to each other and synthesize the literature on the explanations for the value anomalies. The overall evidence shows that the best criterion varies over time and across the markets. The relative efficacy of different valuation criteria also seems to depend on numerous methodological choices. Recent studies have given mild evidence that combining traditional valuation ratios either with each other or with some financial statement variables could at least in some cases enhance the value premium, although very few studies have provided transparent comparisons between the results based on individual valuation ratios and those based on composite value criteria.

Keywords. Valuation multiples; Valuation ratios; Value anomalies; Value investing; Value premium; Value strategies

1. Introduction

The debate on the value anomalies is an excellent example of the fruitful interplay between scholars and investment practitioners. As early as the 1930s, after the stock market crash caused the Great Depression, academics started to develop theories of a fair value of common stocks. These pricing theories motivated investors to chase abnormal returns by using trading strategies that were based on the mispricing of the stocks. Soon after the introduction of the Capital Asset Pricing Model (CAPM), the first contrarian results according to which the relationship between risk and return is not linear were published: Lintner (1965), who is acknowledged as one of the developers of the CAPM, already documented that the security market line was too flat in comparison with the predictions. The follow-up anomaly studies began the new (and current) era of stock market research. During the recent decades, several investment strategies have been proven to generate abnormal returns. In almost every case, proponents of the CAPM have understated such results by invoking data mining, methodological flaws, or even a misinterpretation of the results. However, new evidence against stock market efficiency is continuously being published in academic studies. For example, numerous studies have identified the existence of price momentum on stock returns (e.g., see Jegadeesh and Titman, 1993, 2001; Chan *et al.*, 1996; Rouwenhorst, 1998; Chan *et al.*, 2000;

*Corresponding author contact email: eero.patari@lut.fi; Tel: +3584 0833 6907.

© 2015 John Wiley & Sons Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Korajczyk and Sadka, 2004; Gutierrez and Kelly, 2008; Chui *et al.*, 2010; Israel and Moskowitz, 2013), which refers to the tendency of recent winner stocks to generate abnormal returns in the near future. However, there is plenty of international evidence of a value premium in stock returns (e.g., see Chan and Lakonishok, 2004; Fama and French, 2006, 2012; Cakici *et al.*, 2013), which refers to the tendency of value stocks to outperform glamour stocks. Recently, new evidence of the added value of combining value and momentum strategies has also been documented (e.g., see Bird and Casavecchia, 2007a; Bettman *et al.*, 2009; Leivo and Pätäri, 2011; Guerard, Jr. *et al.*, 2012; Asness *et al.*, 2013). The empirical results of academic studies have formed the basis of many investment strategies commonly used in equity markets. However, dilemmas encountered by portfolio managers and consultants, such as the style identification of value and growth investments and the tailoring of style-specific benchmark indices, have contributed to the academic literature.

The reason value anomalies have piqued a broader interest among economists is that their existence would challenge the semi-strong-form market efficiency, which asserts that security prices fully reflect all publicly available information (Fama, 1970, 1991). Another reason for the increased attention to the value anomalies is their persistence. Though all pricing anomalies should disappear in efficient markets soon after they have been discovered, the value anomaly has not done so (e.g., see Fama and French, 2006; Israel and Moskowitz, 2013). Therefore, academic interest in the value premium has exploded since the seminal papers of Fama and French (1992) and Lakonishok *et al.* (1994), although the origins of value anomaly literature can be traced back to the 1960s (e.g., see Nicholson, 1960, 1968; McWilliams, 1966; Breen, 1968). The reason for the proliferation was Fama and French's (1992) results that strongly challenged the validity of the standard single-factor CAPM, particularly in the minds of the CAPM proponents. The debate stemming from that specific paper has made it one of the most influential articles in the current financial literature.

In this article, we review and update the empirical literature on value anomalies and value premium. Our survey covers studies that examine the portfolio performance of value investment strategies as well as those that examine the cross-sectional explanatory power of various factor combinations on expected or realized returns. Due to space limitations, this literature review focuses on those articles that use variables from individual companies as input data and therefore omits many articles that are based on aggregate stock market data. Because several preceding articles have already provided extensive reviews on the theoretical issues related to the debate over value versus growth investing (e.g., see Fama, 1998; Campbell, 2000), we focus on empirical papers.

The remainder of the paper is structured as follows. Section 2 gives an overview of the evidence of the earnings yield (E/P) anomaly and the related value premium. Section 3 reviews the literature on the book-to-price (B/P) anomaly, whereas Section 4 discusses the dividend yield (D/P) anomaly. Section 5 summarizes and synthesizes the evidence of the sales-to-price (S/P) anomaly, whereas Section 6 focuses on the cash flow-to-price (CF/P) anomaly. Section 7 overviews the enterprise value-based anomalies,[1] and Section 8 introduces evidence of the added value of using composite value criteria. Section 9 discusses the explanations for the value anomalies and the related value premium. Concluding remarks and suggestions of directions for future research are given in Section 10.

2. Earnings Yield (E/P) Anomaly

By definition, the value premium refers to the performance difference, calculated either in terms of absolute or risk-adjusted returns, between value and glamour (or growth) stocks. Correspondingly, the demarcation between value and glamour stocks is based on their relative valuation that is determined by valuation ratios or composite value criteria. Individual valuation ratios can be thought of as output/input ratios, where the output is represented, e.g. by earnings or other measure of profitability, dividends, book value or sales, and the input by, e.g. market value of equity or enterprise value. The more output one

© 2015 John Wiley & Sons Ltd

can get against one unit of input, the better in terms of relative value. Therefore, high valuation ratios, given that they have been denoted as output/input ratios, are characteristics of value stocks, although corresponding ratios of glamour stocks are low.[2]

Although the principles of value investing can be traced back to the 1930s (e.g., see Graham and Dodd, 1934), the first scientific evidence of the E/P anomaly was documented by Nicholson (1960). However, he did not report any risk measures or risk-adjusted performance measures for the portfolios being compared. In the second half of the 1960s, McWilliams (1966), Breen (1968), and Nicholson (1968) released similar types of studies. To the best of our knowledge, Basu (1975, 1977) was the first to document the outperformance of high E/P portfolios also on a risk-adjusted basis. For the large sample of U.S. industrial firms, he reported a monotonically declining performance of quintile portfolios when moving from the high E/P to low E/P portfolios. Basu's seminal work was challenged by Banz (1981) and Reinganum (1981), who both concluded that the E/P anomaly is explained by the small-cap anomaly, and furthermore, that the latter subsumes the former. However, Basu (1983) showed in his response study that the E/P anomaly still exists after exercising experimental control over differences in firm size. He proved further that the size effect virtually disappears when returns are controlled for differences in risk and E/P ratios.[3] In contrast, Cook and Rozeff (1984) attached approximately equal significance to both the E/P and the size factors. However, Banz and Breen (1986) reported a size effect but no independent E/P effect across all months, which is consistent with Reinganum (1981), whose results were criticized by Basu (1983). Chan *et al.* (1993) also did not find the E/P anomaly, but documented significant CF/P and B/P anomalies in the Japanese stock markets. These seemingly paradoxical results can be for the most part explained by differences in the samples, sample periods, and methodologies employed.

After correcting several methodological flaws made in previous studies, Jaffe *et al.* (1989) found significant E/P and size effects when estimated across all months, consistent with Cook and Rozeff (1984). Moreover, Jaffe *et al.* (1989) reported that the E/P effect was significant in all months, whereas the size effect was significant only in January. Interestingly, the authors found evidence of consistently high returns for firms of all sizes with negative earnings. In contrast, Fuller *et al.* (1993) reported low risk-adjusted returns (i.e., the Jensen alphas) for low E/P quintile portfolios that also include stocks with negative earnings. However, their results support the existence of the E/P anomaly despite that they test the robustness of the results to numerous potential explanatory factors.

Fama and French (1992) found that differential returns to E/P strategies are captured by a combination of size and B/P, and therefore, ended up excluding the earnings yield from their famous three-factor model.[4] In contrast, Roll's (1995) performance comparison of three portfolio formation criteria (i.e., size, E/P, and B/P) based on monthly frequency in the reformation of portfolios indicated the superiority of E/P over B/P and size. In their later study, Fama and French (1998) reported that in two out of 13 major national stock markets, the use of E/P as value portfolio formation criteria would have resulted in the highest value premium when comparing four different portfolio formation criteria (the three other were B/P, CF/P, and D/P). However, the E/P criterion did not generate the highest value portfolio return in any of the 13 markets, although it generated a significant value premium in three national markets.

In contrast, van der Hart *et al.* (2003) documented the highest and significant value premium as well as the highest value portfolio return for the E/P criterion when comparing the equally weighted returns of value and growth portfolios of emerging market stocks classified on the basis of B/P, D/P, and E/P rankings. However, following the same methodology for the slightly different sample period, van der Hart *et al.* (2005) documented the superiority of B/P over E/P. Instead, Hou *et al.* (2011) reported the highest global value premium based on E/P ratios for the sample data of over 27,000 stocks from 49 countries when analyzing factors that drive global returns (the other valuation-related factors beside E/P included in their study were B/P, CF/P, and D/P). Hence, the overall international evidence of the relative efficacy of E/P to identify undervalued stocks from their overvalued counterparts is contradictory.

© 2015 John Wiley & Sons Ltd

Table 1 summarizing the studies on the E/P anomaly made with the U.S. sample data, introduces 12 such papers that have compared at least three alternative portfolio formation criteria to each other and also included E/P as one of the criteria.[5] In two out of the 12 aforementioned studies, the E/P criterion has generated the highest value premium. Furthermore, in another of these two (i.e., Davis, 1994), the sample period is from 1940 to 1963 and covers the pre-Compustat era. Based on value portfolio returns, the U.S. evidence of relative efficacy of E/P is also parallel, because in three out of 11 studies, the E/P value portfolio generated the highest return.[6] At the other extreme, in four out of 12 cases, the lowest value premium has been generated by E/P, whereas the use of the same criterion has resulted in the lowest value portfolio returns in three out of 11 studies.

Table 2 includes 19 comparable papers that have employed non-U.S. sample data. In three out of 18 of them,[7] the greatest value premium has been generated by E/P, whereas the highest value portfolio return is based on E/P in only one out 17 comparable studies.[8] At the other end, the use of E/P has resulted in the lowest value premium in six out of 18 cases[9] and in the lowest value portfolio return in six out of 17 cases. Based on overall evidence, E/P has not been among the most efficient selection criteria, because in only 5 out of 30 cases, it has generated the highest value premium, whereas it has generated the lowest in 10 cases. In only three out of 28 cases has the value portfolio return been the highest based on E/P, whereas it has been the lowest in 10 cases. However, it must be noted that this type of comparison includes some selection bias, because only B/P beside E/P has been included among portfolio formation criteria in all of such 31 cases in which at least three different valuation ratios have been compared. The relative efficacy of E/P also varies across the samples because of different treatment practices of negative earnings' stocks. Moreover, the impact of inclusion or exclusion of those stocks varies across the samples and sample periods.[10]

In spite of the mixed evidence of the relative efficacy of E/P, many recent papers have shown its applicability for a stock selection. Chen and Zhang (2007), for example, concluded that beside the Fama–French factors, E/P ratios were useful in explaining price movements of the U.S. stocks. Parallel results from the same stock markets were also reported by Penman and Reggiani (2013) and by Artmann *et al.* (2012) in the German stock markets. According to the latter authors, the explanatory power of the standard Fama–French 3-factor model on the cross-section of average stock returns in Germany has not been strong. An alternative three-factor model in which the size factor was replaced with E/P factor explained the returns better, and adding the momentum factor further increased the explanatory power. The explanatory power of different portfolio formation criteria on subsequent stock returns seems to vary across both the stock markets and the sample periods (see also Barbee *et al.*, 2008 for recent U.S. evidence and Hou *et al.*, 2011 for global evidence). Beside Penman and Reggiani (2013), recent evidence of the E/P anomaly in the U.S. stock market is also documented by Li *et al.* (2009), Athanassakos (2011), and Israel and Moskowitz (2013), and in the U.K. stock market by Anderson and Brooks (2006).

3. Book-to-Price (B/P) Anomaly

The book value of equity provides a relatively stable, intuitive measure of value that can be compared to the market value of the equity that reflects the market's expectations of the firm's earning power and cash flows. High B/P ratios are sometimes considered to provide a "margin of safety," because book value is deemed to be a "floor" supporting the market price (Bodie *et al.*, 2010). However, the relationship between book value and price is much more complex because book values are not necessarily reliable indicators of the assets' current fair value or liquidation value. In spite of that, B/P is the most frequently used valuation ratio in the value premium literature, although recent evidence have also put forth several other criteria as a competitive alternative for that criterion (e.g., see Loughran and Wellman, 2011; Gray and Vogel, 2012; Gharghori *et al.*, 2013; Pätäri *et al.*, 2015).

© 2015 John Wiley & Sons Ltd

Table 1. Studies on E/P Anomalies in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Basu (1977)	Performance comparison of quintile pfs (the separate low E/P pf including and negative earnings' firms is also included).	CRSP/ Compustat	Appr. 500 NYSE industrial firms per year	Apr 1957–Mar 1971	VP is 6.96% (6.75%) for the full (non-negative earnings' firms) sample. The value pf return is 16.3%. Jensen alphas are significantly positive (negative) for the top-two (bottom-two) quintile pfs.
Basu (1983)	Performance comparison of both E/P quintile pfs and double-sorted (based on E/P and size and vice versa) 5×5 quantile pfs (negative firms excluded).	CRSP/ Compustat	Appr. 900 NYSE firms per year	1963–1979	VP is .66% p.m. and the value pf return 1.38%. The value pf returns monotonically decrease when moving to bigger size quintiles. VP is the highest in the middle size quintile.
Cook and Rozeff (1984)	3-factor ANOVA with size, E/P and January anomalies as main effects. Pfs were formed on both independent 2-dimensional 5×5 size and E/P groupings and 5×5 bidirectional double sorts.	Compustat/ Moody's Manuals/ S&P's Stock Guide	900 (average) NYSE stocks	April 1968–Dec 1982	Equity returns are related to both size and E/P, as well as to January anomaly (size does not subsume E/P or vice versa). Instead, they are either two distinct anomalies or separate aspects of a single underlying effect.

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 1. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Senchack, Jr. and Martin (1987)	Performance comparison of S/P- and E/P-based quintile pfs. Hp lengths 3 months and 1 year.	Compustat	From 400 to 450 randomly selected NYSE and AMEX non-financial stocks per hp	1975–1984	Both E/P-based VP (2.65%) and value pf quintile return (26.87%) are lower than their S/P-based counterparts. However, on risk-adjusted basis E/P value pf outperforms comparable S/P pf.
Jaffe *et al*. (1989)	Performance comparison of double-sorted (based on E/P and size) 5×5 quantile pfs including negative earnings´ firms in 5 separate size pfs (resulting in 30 subgroup pfs) & SUR tests.	CRSP/ Compustat	From 352 (1950) to 1,309 (1974)	1951–1986	E/P anomaly exists but its intensity is higher among small-cap stocks than among large-cap stocks. In addition, evidence of consistently high returns for firms of all sizes with negative earnings.
Fama and French (1992)	EW decile pf returns & C-SRs.	CRSP/ Compustat	700–1,770	July 1963–Dec 1990	The E/P-based VP exists but it is not as wide as its B/P-based counterpart. Abnormal returns of stocks with negative earnings is explained by the size effect.
Fuller *et al*. (1993)	Performance comparison of EW quintile pfs (based on abnormal returns of industry-adjusted E/P quintiles).	Compustat/ BARRA (smallest cap stocks excluded)	From 887 (1973) to 1,179 (1990)	1973–1990	Significant alpha spread of 8.0% between high and low E/P quintiles.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 1. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (1993)	Performance comparison of 5 VW quintile pfs and pfs including negative yield stocks (pfs formed on E/P and D/P sorts).	CRSP/ Compustat	All non-financial NYSE, NASDAQ and AMEX stocks	July 1963–Dec 1991	Higher VP (.59% p.m.) than the comparable D/P-based VP (.17%). The same also holds for the value pf returns (.86% vs. .56%). Insignificant FF3 alphas for all E/P pfs.
Lakonishok *et al*. (1994)	Comparison of B/P, CF/P and E/P-based EW decile pf returns up to 5 years after pf formation.	CRSP/ Compustat	All NYSE and AMEX stocks	April 1968–April 1990	The lowest VP (3.9% (7.6%)), as well as the lowest value pf return (16.2% (19.0%)) for 1-year (5-year) hp length. Differences among IVRs are smaller for 5-year hp lengths.
Davis (1994)	C-SR and performance comparison of EW quintile pfs formed on β, B/P, market value, E/P, CF/P, sales growth and stock price.	CRSP and Moody's Industrial Manuals	Annual random samples of 100 firms included in the top half in CRSP files based on the market equity	July 1940–June 1963	The highest VP (9.56%), as well as the highest value quintile pf return (23.04%) among 3 IVRs examined, although the differences between E/P and CF/P criteria are small or marginal. E/P has significant explanatory power on subsequent returns.

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Table 1. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Roll (1995)	Performance comparison of 8 pfs including the stocks in the top and bottom halves based on 3 criteria (size, E/P and B/P) and monthly reformation frequency.	Roll and Ross Asset Management database/ CRSP	From 2,160 (appr.) to 3,160 (appr.) stocks listed in NYSE, AMEX and OTC	April 1984–Mar 1994	The best criterion is E/P. In raw return comparisons, the best 4 pfs are all high E/P pfs (The best pf is small-cap high E/P & high B/M pf with 21.2% return).
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, D/P & E/P.	CRSP/ Compustat	3,333–6,258 NYSE, NASDAQ and AMEX stocks	1975–1995	The second highest and significant VP (6.71%), as well as the second highest value pf return (14.09%).
Dhatt *et al.* (1999)	Performance comparison of B/P-, E/P- and S/P-based tertile pfs. Firms with negative IVRs excluded from tertiles but included in separate pf for each ratio.	Compustat/ Russell 2000 small-cap data	1,981 (average) small-cap stocks	July 1979–June 1997	The lowest VP (5.28%), but the highest value pf return (19.80%). However, E/P-based tertiles are smaller than B/P and S/P tertiles.
Dhatt *et al.* (2004)	Comparison of B/P-, CF/P- and E/P-, and S/P-based quintile pf performance.	CRSP/ Compustat	1,280–2,314 NYSE, NASDAQ and AMEX stocks	1980–1998	The second lowest VP (4.68%) and the lowest value pf return (18.36%) among 4 criteria being compared.
Desai *et al.* (2004)	Return comparison of B/P, CF/P, operating cash flow/P and E/P decile pfs & C-SR tests.	CRSP/ Compustat	2,823 (average) NYSE, AMEX and NASDAQ stocks per year	1973–1997	The lowest VP (10.2%) and the second lowest value pf return (27.4%) among 4 IVRs. The return differences between value pfs are small (max. 2.1%).

(*Continued*)

Table 1. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (2006)	25 VW pfs formed as intersections of independent sorts of U.S. stocks into five size groups and five E/P groups.	CRSP/ Compustat	3,858 (average)	June 1926–Dec 2004	Significant VP (return difference between the top-2 vs. the bottom-2 quintiles) exists for all 5 size groups. The VP is higher based on E/P than based on B/P in two smallest-size groups.
Barbee *et al*. (2008)	Return comparison of quintile pfs formed on 4 IVRs (E/P, CF/P, S/P, B/P) and C-SR tests based on 30 pfs formed on each of the 4 IVRs.	CRSP/ Compustat	990–1,860 NYSE, NASDAQ and AMEX firms	1981–2000	The second highest VP (3.78%) and value pf return (15.69%).
Li *et al.* (2009)	Performance comparison of VW decile pfs formed on CF/P, E/P and B/P.	French's data library (derived from CRSP /Compustat data)	NYSE, AMEX and NASDAQ stocks	July 1963–June 2006	The highest VP of (.6% p.m.), as well as the highest value pf return (1.42%). The differences among the 3 IVRs are very small.
Athanassakos (2011)	Return comparison of EW top and bottom quartile pfs formed on B/P and E/P separately of AMEX, NASDAQ and NYSE stocks.	CRSP/ Compustat	7,468 stocks (total within the sample period)	July 1985–June 2006	Lower VPs than based on B/P ratios in all three main U.S. stock exchanges (6.00% for NYSE, 11.40% for NASDAQ and 6.24% for AMEX stocks).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 1. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Loughran and Wellman (2011)	Return comparison of EW and VW decile pfs formed on 7 criteria (B/P, E/P, EBITDA/EV, D/P, sales growth, market leverage and prior 3-year returns) & C-SRs.	CRSP/ Compustat	2,280 (average) NYSE, AMEX and NASDAQ non-financial firms	June 1963–June 2009	The second lowest VP of .48% p.m. (.27%) and the value decile pf return 1.58% (1.21%) in comparison of EW (VW) returns.
Gray and Vogel (2012)	Performance comparison of EW and VW quintile pfs formed on 5 criteria (B/P, E/P, EBITDA/EV, FCF/EV and GP/EV).	CRSP/ Compustat	All NYSE, AMEX and NASDAQ non-financial firms with 8-year data for the calculation of IVRs	June 1971–Dec 2010	The lowest VP, and the lowest VW value pf return (13.62%; the latter shared with B/P criterion). Insignificant FF3 alphas for both E/P-based value pfs.
Penman and Reggiani (2013)	Return comparison of E/P decile pfs (negative E/P firms included).	CRSP/ Compustat	3,498 (average) U.S. listed firms	1963–2006	VP is 9.3%. The value decile pf return 25.3%.
Israel & Moskowitz (2013)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and past 5-year return-based VW decile pfs.	French's data library (derived from CRSP /Compustat data)	All US stocks included in decile pfs formed by French	July 1951–Dec 2011	The second highest VP (6.48% in terms of excess returns) and the second lowest value pf return (12.08% excess of risk-free rate) among 4 IVRs.
Hou *et al.* (2015)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and 7 other relative value-based VW decile pfs.	CRSP/ Compustat	All non-financial US stocks	July 1972–Dec 2012	The second highest and significant VP (.59% p.m.) among the 4 IVRs.

Table 1 summarizes the results from 24 studies on E/P anomalies in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period*, ANOVA to *analysis of variance*, C-SR to *cross-sectional regression*, SUR to *seemingly unrelated regression*, and FF3 alpha to the *Fama-French 3-factor alpha*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 2. Studies on E/P Anomalies in the Non-U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Levis (1989)	Performance comparison of D/P-, E/P-, share price- and size-quintile pfs and the corresponding double-sorted 5×5 pfs.	LSPD, London Stock Exchange	770–1,920 (1,150–2,300) UK stocks for E/P pfs (for other criteria)	April 1961–Mar 1985	The second lowest top-bottom quintile spread (7.0%) and the second lowest top quintile pf return (17.8%) among 4 criteria. However, the return differences between top pfs are marginal (max .96%). D/P anomaly is stronger than E/P anomaly, but together they subsume size and share price anomalies.
Chan *et al.* (1991)	Comparison of B/P-, CF/P- and E/P- and size-based quartile pf returns and C-SRs.	Hamao/ Daiwa Securities	1,215–1,246 Japanese stocks (range of average depending on the pf formation criteria) listed in TSE	June 1971–Dec 1988	The lowest VP (.4% p.m.) and value pf return (1.94%). However, the highest return among all pfs examined are reported for the pf of negative E/P stocks. The lowest explanatory power in C-SRs (size factor included).

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Miles and Timmermann (1996)	Performance comparison of EW decile pfs based on 3 IVRs (D/P, E/P & B/P) & monthly C-SRs.	Extel company accounts data/ LSPD	457 UK non-financial firms	May 1979–April 1991	E/P-based VP (.26% p.m.) is much lower than B/P-based VP (.79%), but higher than that based on D/P (-.16%). The lowest value pf return (1.21%).
Cai (1997)	Return comparison between the extreme decile pfs formed on 4 formation criteria (B/P, E/P, CF/P & sales growth) for 5-year hps & C-SR tests.	PACAP/Daiwa Securities/ Nihon Keizai Shimbun	1,186–1,651 Japanese non-financial and non-utility stocks in the first and second sections of TSE	June1971–Dec1993	No E/P-based VP. The weakest and insignificant explanatory power in C-SRs.
Doeswijk (1997)	Return comparison of B/P-, CF/P, D/P- and E/P-based quintile pfs. Hp lengths 1 year and 3 years.	Amsterdam Stock Exchange /Manual of Dutch Stocks	145–183 Dutch stocks per year	July 1976–June 1995	The highest VP (8.6%) and the second highest value pf return (19.2%) for 1-year hp lengths. The second highest VP (7.9%) and the highest value pf return (21.6%) for 3-year hp lengths.
Mukherji *et al.* (1997)	Return comparison of B/P, E/P and S/P tertile pfs.	PACAP complied by Research Center at the University of Rhode Island.	213 (median) Korean stocks	1982–1993	The lowest VP (2.55%) and value tertile pf return (17.26%) among 3 IVRs examined.

(*Continued*)

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Bauman *et al.* (1998)	Performance comparison of B/P-, CF/P-, D/P- & E/P-based EW quartile pf (examination with 4 most common FYEs).	Compustat Global Vantage	28,463 stocks (total) from 21 countries included in MSCI EAFE (Europe/Australia/ Far East) index + Canada, 28,463 stocks (total)	1985–1996	The second lowest (though significant) VP of 4.4%, and the second lowest value pf return (15%). The highest VP (6.4%) among the firms with FYE in June, whereas the lowest and negative (-1.7%) among those with FYE in September. The VP is positively related to firm size (no VP among companies in the lowest 25% size group, whereas for the top 25% group, it is the highest (9.4%)).
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, D/P and E/P.	MSCI EAFE data	948 (1975)–1,593 (1995) stocks from 12 developed national markets	1975–1995	Significant E/P-based VP in France, whereas insignificant in other 11 countries. The highest but insignificant E/P-based VPs in 2 out of 12 countries (in Netherlands and Sweden).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
	EV and VW return comparison between the top 30% of the high E/P stocks and the bottom 30% of the low E/P stocks of 16 emerging markets.	IFC Emerging Markets database	Stocks from 16 emerging stock markets	1987–1995	Significant positive E/P-based VP in 1 of 16 countries. Huge cross-country VP differences between B/P and E/P criteria. On average, E/P-based VP is insignificant.
Suzuki (1998)	Return comparison of yearly reformed pfs, each containing 100 highest B/P, E/P or S/P stocks.	Daiwa Institute of Research/ TSE	100 Japanese stocks in each value pf	1983–1996	The lowest value pf return (11.74%).
Gregory *et al.* (2001)	Return comparison of B/P, CF/P and E/P decile pfs based on 1- and 5-year hp lengths.	Datastream & LSPD	UK stocks (sample size N/A)	Jan 1975–Dec 1998	The lowest VP of 11.50% (3.98%) and the (second) lowest value pf return of 35.42% (24.62%) based on 1-year (5-year) hp.
Bird and Whitaker (2003)	Return comparison of quintile pfs formed on B/P, S/P, D/P and E/P. (Hp lengths 1, 3, 6, 9, 12, 24, 36 & 48 months).	Worldscope and GMO Woolley	2,219 firms (average) from 7 European developed countries	Jan 1990–Jun 2002	Negative VP for all the hp lengths longer than 6 months. The second lowest value pf returns for hp lengths from 3 months upwards.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
van der Hart *et al*. (2003)	Performance comparison of EW B/P, D/P and E/P pfs formed from the top and bottom 15% of stocks in each country. The internationally diversified pfs consist of EW country pfs rebalanced monthly using 6-month hps.	IFC Emerging Markets database	685 (average) emerging market stocks	Jan 1985–May 1999	The highest and significant VP (1.02% p.m.), as well as the highest value pf return (1.89%) among 3 IVRs.
Yen *et al.* (2004)	Performance comparison of 3 pairs of EW value and growth pfs formed on intersection of size and either B/P, CF/P or E/P by first dividing sample into size quartiles and then ranking stocks within them based on IVRs.	PACAP database	All Singaporean non-financial firms	July 1976–Dec 1998	The lowest VP (.33% p.m.) and size-adjusted value pf return (1.08%) among 3 IVRs. The return differences between value pfs are marginal (max. .06%).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
van der Hart *et al*. (2005)	Performance comparison of EW B/P and E/P pfs formed from the top and bottom 15%ile of stocks in each country. The internationally diversified pfs consist of EW country pfs that are rebalanced monthly using 6-month hps.	S&P/IFC Emerging Markets database	576 (on average) emerging market stocks	Dec 1988–June 2004	VP is slightly lower based on E/P (.68% p.m.) than based on B/P (.73%). The VP difference between the selection criteria varies over time.
Anderson and Brooks (2006)	Return comparison of EW decile pfs formed on E/Ps calculated from normalized earnings history up to 8 years. (Hps also extend from 1 to 8 years).	LSPD and Datastream	From 16,000 to 40,000 UK E/P observations depending on the length of earnings history requirement	1975–2003	The highest VP (11.62%) and value pf return (27.87%) is documented when E/Ps are calculated using 8-year historical average earnings. This practice nearly doubles the VP (5.98%) resulting from using the standard method in the calculation of E/Ps.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (2006)	25 VW pfs formed as intersections of independent sorts of the sample stocks into five size groups and five E/P groups.	MSCI merged data for 14 national markets outside USA	1,248 (average)	1975–2004	Significant overall VP (.65% p.m. return difference between the top-2 vs. the bottom-2 E/P quintiles) which is virtually identical among the biggest and the smaller stocks.
Kyriazis and Diacogiannis (2007)	Comparison of B/P-, D/P- and E/P-based tertile pf returns & C-SRs.	Datastream/ Athens Stock Exchange	Max. 260 Greek stocks	1995–2002	The second highest and significant VP (19.3%), as well as the second highest value pf return (19.2%). Highly significant explanatory power in C-SRs.
Bird and Casavecchia (2007b)	Return comparison of country-adjusted B/P-, S/P-, and E/P-based quintile pf for hp lengths of 3, 6, 12, 24 and 36 months.	Worldscope/ GMO Woolley London and IBES	Appr. 1,650 stocks from 15 European countries per year (total <8,000)	May 1989–April 2004	The lowest VP for hp lengths of 12 and 24 months. The lowest value pf returns for all hp lengths. The highest E/P-based VP and the highest E/P value pf return for 3-year hp length.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Leivo *et al.* (2009)	Performance comparison of quintile pfs formed on 4 IVRs (E/P, EBITDA/EV, B/P and S/P) and three hybrids of them. Hp length is 3 years.	Datastream /Opstock	74 (average) Finnish non-financial stocks	May 1991–Apr 2006	The second highest VP among IVRs (15.82%) and the second lowest value pf return (24.94%) among 4 IVRs.
Dissanaike and Lim (2010)	Comparison of long/short decile pfs formed on B/P, CF/P, operating cash flow/P, and E/P, and the Ohlson (1995) model and the residual income model.	Datastream/ Inland Revenue database	All listed UK firms (1,791, on average at some point or another) in LSE	1987–2001	The second lowest, but highly significant VP (10.57%) among 4 IVRs for all hp lengths examined (1, 2 and 3 years).
Hou *et al.* (2011)	Examination of size, D/P, E/P, CF/P, B/P, leverage, and momentum as potential explanatory factors for global stock returns. VP comparison based on VW quintile pfs formed on 4 IVRs examined.	Datastream / Worldscope	27,488 global common stocks from 49 countries	July 1981–Dec 2003	The highest VP (.75% p.m.). E/P has significant explanatory power on subsequent returns, but it works poorly as a value-based factor on global returns, unlike CF/P.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream /Opstock	85 (average) Finnish non-financial stocks	May 1993–Apr 2008	The third lowest, though moderate VP (13.83%) among all pf formation criteria examined. The second highest value sextile pf return (18.29%) among 6 IVRs.
Artmann *et al*. (2012)	Return comparison of decile pfs formed on 10 criteria (incl. B/P and E/P). C-SR and asset pricing tests using factor pfs formed on the same criteria.	Karlsrüher Kapitalmarkt daten bank and Saling/ Hoppenstedt Aktienführer	From 175 to 598 (955 altogether) German stocks listed in Frankfurt Stock Exchange	1963–2006	Significant and slightly higher VP (.99% p.m.) than based on B/P (.9%). The value pf return is also slightly higher based on E/P (1.31%) than based on B/P (1.20%). Both B/P and E/P are significant factors in cross-sectional and asset pricing tests.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 2. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Gharghori *et al.* (2013)	C-SR and return comparison of EW and VW decile (quintile*) pfs formed on 6 criteria (B/P, (E/P*), (CF/P*), S/P, debt-to-equity ratio and firm size).	CRIF/Aspect Huntley	778 (average) Australian stocks	Jan 1993–Dec 2009	The lowest VP among the quintile pfs. Positive E/Ps explain the subsequent returns better than negative E/Ps.
Pätäri *et al.* (2015)	Performance comparison of EBIT/EV-, B/P-, E/P- and S/P-based tertile pfs and all their size-, industry- and leverage-adjusted combinations and permutations.	Datastream /Opstock	97 (average) Finnish non-financial stocks	May 1996–Apr 2013	The second highest VP (6.16%) and the second lowest value pf return (12.05%) among 4 IVRs. However, industry-adjusted E/P criterion is one component in all of the best-performing combination strategies.

Table 2 summarizes the results from 25 studies on E/P anomalies in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period,* C-SR to *cross-sectional regression* and FYE to *fiscal year-end*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

To our knowledge, Stattman (1980) was the first to report a significant B/P anomaly in the U.S. stock market, although his results are prone to both survivorship and look-ahead biases due to the sample selection criteria employed. Chan *et al.* (1991) compared four portfolio formation criteria (i.e., the CF/P, E/P, B/P, and size criteria) in the Japanese stock market and concluded that B/P had the best discriminatory power on value and glamour stocks. In addition, the best performance in terms of both absolute and risk-adjusted returns was also reported for B/P value quartile portfolios. Parallel results from the U.S. markets were documented by Fama and French (1992), who found B/P to have the best explanatory power on expected returns in the U.S. markets. The authors further demonstrated that together with the market value of equity (i.e., firm size), these two variables captured the explanatory power of E/P. The dramatic dependence of returns on B/P was independent of β, suggesting either that high B/P firms are relatively underpriced, or that B/P is serving as a proxy for a risk factor that affects equilibrium expected returns. After controlling for the size and B/P effects, β seemed to have no power to explain average security returns, indicating that systematic risk seems not to matter, whereas the B/P ratio seems to be capable of predicting future returns.

In line with the seminal paper of Fama and French (1992), Capaul *et al.* (1993) documented the inverse relationship of return and β in most of the major stock markets when comparing the value premiums and their βs in six developed national markets. The authors also showed that the B/P anomaly was a global phenomenon and even stronger outside the United States. The overall evidence of the studies reviewed in our paper reinforces that this conclusion is also true at a more general level: Tables 3 and4 include 12 United States and 19 non-United States studies in which at least three alternative portfolio formation criteria based on individual valuation ratios, including B/P, have been compared to each other. For the U.S. sample data, in four of them, the greatest value premium was generated by B/P, whereas the B/P value portfolio return was the highest in three out of 11 studies. For the non-U.S. sample data, the corresponding proportions are 10 out of 18 and 10 out of 17, respectively. At the other extreme, in two out of 12 U.S. studies, the lowest value premium was generated by B/P rankings that also have resulted in the lowest value portfolio return in four of 11 cases. For the non-U.S. sample data, the lowest value premium was generated by B/P in five out of 18 studies, whereas the lowest value portfolio return was documented for the same criterion in only two out of 17 cases.

Based on these rough statistics, it seems that the relative efficacy of B/P compared to that of E/P has been somewhat stronger in the studies based on the non-U.S. sample data. If the condition of the inclusion of at least three alternative portfolio formation criteria based on individual valuation ratios is exempted, and the efficacy comparison is made among the studies that include B/P and E/P comparisons, Tables 1–4 include 35 such cases. In 21 of them, the B/P-based value premium has been higher than its E/P-based counterpart, whereas the reverse holds for 14 cases. In addition, the value portfolio return based on B/P was higher than that based on E/P in 20 out of 33 cases, whereas the reverse conclusion was drawn in 13 studies. If pooled results for the U.S. and non-U.S. samples are compared separately, the proportions based on comparisons of value portfolio returns reinforce the conclusion that B/P has in this sense worked somewhat better than E/P, particularly for the non-U.S. sample data, because the former has generated higher value portfolio returns in 14 out of 20 cases.[11] By contrast, based on the pooled results for the U.S. sample data, the E/P-based value portfolio return was higher in seven out of 14 cases. However, it should be noted that samples used as bases of quantile divisions based on B/P and E/P are not necessarily identical, because in many studies, negative earnings' stocks have been excluded from the E/P-based division, whereas they have been included in the B/P-based division unless their book values have also been negative (e.g., see Fama and French, 1992; Dhatt *et al.*, 1999).

Like the E/P anomaly, the B/P anomaly is also related to firm size. B/P-based value premium has often been documented to be at its highest among small-cap stocks and at its lowest among large-cap stocks (e.g., see Fama and French, 1993, 2012; Israel and Moskowitz, 2013; Hou *et al.*, 2015). However, this relationship cannot be generalized: for example, Bauman *et al.* (1998) documented for the large international sample that the value premium was higher among large-cap stocks than among small-cap

© 2015 John Wiley & Sons Ltd

Table 3. Studies on B/P Anomalies in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Stattman (1980)	Risk-adjusted return comparison of EW decile pfs.	CRSP/ Compustat (NYSE and AMEX firms)	308–643	April 1964–Mar 1978	Positive, large and consistent relationship between B/P and subsequent risk-adjusted returns. Average annual risk-adjusted return for the highest B/P pf is about 10%.
Fama and French (1992)	Comparison of EW decile pf returns and C-SRs.	CRSP/ Compustat	700–1770	July 1963–Dec 1990	Strong evidence of B/P-based VP. B/P generates higher VP, as well as higher value pf return than E/P. In C-SRs, B/P is stronger explanatory factor than firm size. The combination of size and B/P absorbs the apparent roles of leverage and E/P in C-SRs.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Capaul *et al.* (1993)	Performance comparison of high and low B/P pfs formed following 50%/50% division of S&P500 stocks based on their market cap (Hp 6 months).	S&P/ BARRA	500 largest-cap US stocks (incl. appr. 75 % of the largest-cap stocks)	Jan 1981–Jun 1992	VP is slightly positive (.11% p.m.) but insignicant and lower than in any of other 5 countries included in comparison. The breadth of value and growth pfs explains low VP.
Lakonishok *et al.* (1994)	Comparison of B/P, CF/P and E/P-based EW decile pf returns up to 5 years after pf formation.	CRSP/ Compustat (NYSE/AMEX firms)	All NYSE and AMEX stocks	April 1968–April 1990	Significant VP of 6.3% (10.0%) between extreme deciles when using 1-year (5-year) hp. The B/P-based VP, as well as value pf return is lower than those based on CF/P but higher than based on E/P.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Davis (1994)	C-SR and performance comparison of EW quintile pfs formed on β, B/P, market value, E/P, CF/P, sales growth and stock price.	CRSP and Moody's Industrial Manuals	Annual random samples of 100 firms included in the top half in CRSP files based on size	July 1940–June 1963	The lowest VP (6.82%), as well as the lowest value quintile pf return (21.06%) among 3 IVRs examined. CF/P subsumes B/P, though the latter is significant in univariate cross-sectional tests.
Roll (1995)	Performance comparison of 8 pfs including the stocks in the top and bottom halves based on 3 criteria (size, E/P and B/P) and monthly reformation frequency.	Roll and Ross Asset Management database/CRSP	From 2,160 (appr.) to 3160 (appr.) stocks listed in NYSE, AMEX and OTC	April 1984–Mar 1994	B/P is less efficient than E/P, but is a significant factor after accounting for multifactor risk with the APT, though less significant than E/P.
Chan *et al*. (1995)	Return comparison of quintile pfs formed from the top size-quintile stocks.	CRSP/ Compustat, Moody's Manuals & Value Line Investment Survey	From 404 (1,990) to 510 (1973) NYSE & AMEX stocks	1963–1992	Selection bias in Compustat is not a severe problem. The value quintile pf return is .1 pps higher (16.5%) for the complemented sample than for the Compustat sample. VP is equal (5.0%) for the both samples.

(*Continued*)

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, E/P and D/P.	CRSP/ Compustat	3,333–6,258 NYSE, AMEX and NASDAQ stocks	1975–1995	The highest and significant VP (6.79%) and the highest value pf return (14.55%). However, the VP differences are marginal among the first three criteria.
Dhatt *et al.* (1999)	Performance comparison of B/P-, E/P- and S/P-based tertile pfs. Firms with negative valuation IVRs excluded from tertile division but included in separate pf for each ratio.	Compustat/ Russell 2,000 small-cap stocks	1,981 (average) small-cap stocks	July 1979–June 1997	The second highest VP (6.60%). The lowest value pf return (18.12%) among 3 IVRs.
Davis *et al.* (2000)	C-SR and return comparison of VW pfs formed on 2×3 sorts based on size and B/P.	Moody's Industrial Manuals and Compustat/ CRSP	339 (June 1929) to 4,562 (June 1996) firms	1929–1997	The average VP is .46 % p.m. The highest return (1.53%) is reported for high B/P small-cap pf, whereas the lowest (.89%) for low B/P large-cap pf. The VP is relatively even among large- and small-cap stocks (.44% and .48%, respectively).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Dhatt *et al.* (2004)	Performance comparison of B/P-, CF/P-, E/P-, and S/P-based quintile pfs.	CRSP/ Compustat	1,280–2,314 NYSE, NASDAQ and AMEX stocks	1980–1998	The lowest VP (3.48%) and the second lowest value pf return (19.08%) among 4 IVRs.
Desai *et al.* (2004)	Return comparison of B/P, CF/P, operating cash flow/P and E/P decile pfs & C-SR tests.	CRSP/ Compustat	2,823 NYSE, AMEX and NASDAQ stocks per year (average)	1973–1997	The highest and significant VP (15.4%), but the lowest value pf return (26.2%) among 4 IVRs. However, the return differences between value pfs are small (max. 2.1%).
Bartov and Kim (2004)	Performance comparison of quintile pfs formed on B/P.	CRSP/ Compustat	2,065 NYSE, AMEX or NASDAQ firms (average)	May 1981–April 2000	VP 14.1 % (value quintile return 20.1% and growth quintile return 6.0%).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama & French (2006)	25 VW pfs formed as intersections of independent sorts of U.S. stocks into five size groups and five B/P groups.	CRSP/ Compustat	3,858 (average)	June 1926–Dec 2004	Significant B/P-based VP (return difference between the top-2 vs. the bottom-2 B/P quintiles) exists for all the five size groups, except for the largest-cap group. The VP is higher based on B/P than based on E/P in three biggest size groups.
Beaver *et al*. (2007)	Return comparison of decile pfs based on data with and without correcting delisting bias.	CRSP/ Compustat (NYSE, AMEX and NASDAQ stocks)	3,187 (3,445) (average for the delisting-biased (-adjusted) sample)	1962–2002	VP is higher for the delisting-adjusted sample than for the delisting-biased sample (14.6% vs. 9.4%), whereas the reverse holds for value decile pf returns (19.6% vs. 20.2%).

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Phalippou (2008)	VP comparison of double-sorted pfs (first into deciles based on institutional ownership, then within deciles into quartiles based on B/P).	CRSP/ Compustat, I/B/E/S	3,330 (average) stocks (NYSE, AMEX and NASDAQ)	July 1980–Dec 2001	VP is driven by 7% of the stock market. For 93% of the market capitalization held by institutional investors, VP is insignificant. For the sample of stocks held most by individual investors, it is huge (1.85% p.m.).
Barbee *et al*. (2008)	Return comparison of quintile pfs formed on 4 IVRs (E/P, CF/P, S/P, and B/P) and C-SR tests based on 30 pfs formed on each of the 4 IVRs.	CRSP/ Compustat	990–1,860 NYSE, NASDAQ and AMEX firms	1981–2000	The second lowest VP (1.93%) and value pf return (15.58%).
Li *et al*. (2009)	Performance comparison of VW decile pfs formed on CF/P, E/P and B/P.	French's data library (derived from CRSP/ Compustat data)	NYSE, AMEX and NASDAQ stocks	July 1963–June 2006	The second highest VP (.57% p.m.) and value pf return (1.39%). The differences among the 3 IVRs are very small.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Garcia-Feijóo and Jorgensen (2010)	Return comparison of B/P quintile pfs and double-sorted pfs (first into 5 size groups, then within them into B/P quintiles).	CRSP/ Compustat	2,080 (average) NYSE, AMEX and NASDAQ stocks	June 1987–Dec 2003	Average VP is .73% p.m. and at its largest among small caps (1.14%) and smallest among large caps (.13%). The average value quintile return is 1.38%, and at its highest among the second largest size group (1.67%).
Athanassakos (2011)	Return comparison of EW top and bottom quartile pfs formed on (B/P and E/P) separately for AMEX, NASDAQ and NYSE stocks.	CRSP/ Compustat	7,468 stocks (total within the sample period)	July 1985–June 2006	Higher VPs than based on E/P ratios in all three main U.S. stock exchanges (7.08% for NYSE, 13.44% for NASDAQ, and 14.88% for AMEX stocks).

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Loughran and Wellman (2011)	Return comparison of EW and VW decile pfs formed on 7 formation criteria (B/P, E/P, D/P, EBITDA/EV, sales growth, market leverage and prior 3-year returns) & C-SRs.	CRSP/ Compustat	2,280 (average) NYSE, AMEX and NASDAQ non-financial firms	June 1963–June 2009	The (second) highest VP of 1% p.m. (.51%) and the (second) highest value decile pf return of 1.84% (1.38%) in comparison of EW (VW) returns.
Gray and Vogel (2012)	Performance comparison of EW and VW quintile pfs formed on 5 IVRs (B/P, E/P, EBITDA/EV, FCF/EV and GP/EV).	CRSP/ Compustat	All NYSE, AMEX and NASDAQ non-financial firms with 8-year data for the calculation of IVRs	June 1971–Dec 2010	The lowest EW value pf return (15.03%). Negative (but not significant) FF3 alphas for both EW and VW B/P-based value pfs.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 3. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Israel and Moskowitz (2013)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and past 5-year return-based VW decile pfs.	French's data library (derived from CRSP/ Compustat data)	All US stocks included in decile pfs formed by French	July 1951–Dec 2011	The second lowest VP (5.89% in terms of excess returns), but the highest value pf return (12.54% excess of risk-free rate) among 4 IVRs.
Hou *et al.* (2015)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and 7 other relative value -based VW decile pfs.	CRSP/ Compustat	All non-financial US stocks	July 1972–Dec 2012	The highest and significant VP (.7% p.m.) among the first 4 IVRs.

Table 3 summarizes the results from 24 studies on B/P anomalies in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period,* pps to *percentage points*, APT to *Arbitrage Pricing Theory*, C-SR to *cross-sectional regression*, and FF3 alpha to the *Fama-French 3-factor alpha*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description).

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 4. Studies on B/P anomalies in the non-U.S. stock markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Chan *et al.* (1991)	Comparison of B/P-, CF/P- and E/P- and size-based quartile pf returns & C-SRs.	Hamao/Daiwa Securities	1,215–1,246 (range of average depending on the pf formation criteria) Japanese stocks	June 1971–Dec 1988	The highest VP (1.1% p.m.), as well as the highest value pf return (2.43%) among the 3 IVRs. However, higher returns were reported for all 3 pfs that consisted of stocks with negative IVRs. The best explanatory power in C-SRs (size factor included).
Capaul *et al.* (1993)	Performance comparison of high and low B/P pfs formed following 50%/50% division from MSCI country indices based on market capitalization of stocks for 5 developed countries (hp 6 months).	Morgan Stanley/UBS	The sample includes appr. 60 % of the total market cap of each country	Jan 1981–Jun 1992	VP is the highest in France (.53%), the lowest in Germany (.13%), and globally significant. The highest market-adjusted abnormal return for the value pf in Japan.
Miles and Timmermann (1996)	Performance comparison of EW decile pfs formed on 3 IVRs (D/P, E/P & B/P) and monthly C-SRs.	Extel Company Accounts data/ LSPD	457 UK non-financial firms	May 1979–April 1991	The highest VP (.79% p.m.) and value pf return (1.54%) among the 3 IVRs. In terms of risk-adjusted returns, B/P is the only IVR that generates significant outperformance for the value pf.
Cai (1997)	Return comparison between the top and bottom decile pfs formed on 4 criteria (B/P, E/P, CF/P and sales growth) for 5-year hps and C-SR tests.	PACAP/Daiwa Securities/Nihon Keizai Shimbun	From 1,186 to 1,651 Japanese non-financial and non-utility stocks	June1971–Dec1993	The highest VP (11.2%) and value pf return (25.5%) among the 3 IVRs. The best explanatory power in C-SRs (size factor included).
Doeswijk (1997)	Return comparison of B/P-, CF/P, D/P- and E/P-based quintile pfs. Hp lengths 1 year and 3 years.	Amsterdam Stock Exchange /Manual of Dutch Stocks	From 145 to 183 Dutch stocks per year	July 1976–June 1995	The second lowest VP (3.7%) and value pf return (15.3%) for 1-year hp lengths. The highest VP (9.2%) and the second highest value pf return (21%) for 3-year hp lengths.

(*Continued*)

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Mukherji *et al.* (1997)	Return comparison of B/P, E/P and S/P tertile pfs.	PACAP complied by Research Center at the University of Rhode Island.	213 (median) Korean stocks	1982–1993	The highest VP (22.53%) and the second highest value tertile pf return (30.39%).
Bauman *et al.* (1998)	Comparison of B/P-, CF/P, D/P and E/P-based EW quartile pf performance (examination with four most common FYEs).	Compustat Global Vantage	32,265 stocks (total) from 21 countries included in MSCI EAFE (Europe/Australia/Far East) index + Canada	1985–1996	The highest and significant VP of 5.7% and the highest value pf return (18.1%) among the 4 IVRs. The B/P-based VP much bigger (9.3%) among the firms with FYE in March than for those with FYE in December (2.7%). The only criterion that also finds the significant VP (5.4%) among 25% of the smallest-cap stocks.
Fama and French (1998)	VW return comparison between the top and the bottom 30% stocks included in MSCI country index and selected on the basis of either B/P-, CF/P, D/P and E/P.	MSCI EAFE data	948 (1975)–1,593 (1,995) stocks from 12 developed national markets	1975–1995	Significant B/P-based VP in 5 out of 12 countries (Japan, France, Belgium, Australia and Singapore), whereas insignificant and positive in 6 countries (negative in Italy). The highest B/P-based VP in 5 out of 12 countries (Japan, U.K., Belgium, Switzerland and Singapore). The highest value pf return in 7 out of 12 countries (Japan, France, Netherlands, Switzerland, Sweden, Australia and Singapore).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (1998)	EW and VW return comparison between the top and the bottom 30% stocks from emerging markets and selected on the basis of either B/P-, CF/P, D/P and E/P.	IFC Emerging Markets database	Stocks from 16 emerging stock markets	1987–1995	Significant positive (negative) VP in 3 (1) of 16 countries. Higher VP based on B/P than E/P in 11 countries. Higher value pf return based on B/P than E/P in 7 of 16 countries. Huge VP differences between B/P and E/P criteria. On average, higher and significant VP based on B/P than based on E/P (insignificant).
Arshanapalli *et al.* (1998)	Performance comparison of high and low B/P pfs formed following 50%/50% division based on market capitalization of stocks.	IIA database	1,554-2,629 stocks in 18 national markets (incl. USA)	1975–1995	The highest VP (17.69%) in Japan, and based on regional comparisons, in Pacific-Rim countries (17.27%, on average). International VP is 13.99%.
Suzuki (1998)	Return comparison of yearly reformed pfs, each containing 100 highest B/P, E/P or S/P stocks.	Daiwa Institute of Research/Tokyo Stock Exchange	100 Japanese stocks in each value pf	1983–1996	The highest value pf return (15.26%).
Rouwenhorst (1999)	Return comparison between EW top and bottom 30% B/P stocks.	IFC Emerging Markets database	1,564 firms from 20 emerging market countries	1982–1997	High B/P stocks outperform low B/P stocks in 16 of 20 countries. The average EW VP is .72% p.m. (.93% based on equal country weights).
Gregory *et al.* (2001)	Return comparison of B/P, CF/P and E/P decile pf based on 1- and 5-year hp lengths.	Datastream & London Share Price Database	UK stocks (sample size N/A)	Jan 1975–Dec 1998	The highest VP of 22.18% (8.95%) and the highest value pf return of 41.93% (27.42%) based on 1-year (5-year) hp.
Leledakis and Davidson (2001)	Return comparison of decile pfs formed on 4 criteria (B/P, S/P, debt-to-equity ratio and size) & C-SRs.	London Share Price Database/ Datastream	810 (average)	June 1980–June 1996	Slightly lower value pf return (31.92%) than for S/P value pf, but higher VP (18.84%) than for the S/P pfs. The cross-sectional explanatory power somewhat lower than that of S/P ratios.

© 2015 John Wiley & Sons Ltd

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Daniel *et al.* (2001)	Performance comparison of double-sorted 5 x 5 pfs (based on B/P and size).	PACAP database/Daiwa Securities/Nihon Keizai Shimbun	All Japanese common stocks listed on both sections of TSE	Oct 1975 –Dec 1997	VP is highest (1.03% p.m.) in the middle size quintile and significant in all size quintiles except in the smallest for which the value pf return is the highest (The only significant positive FF3 alpha is also documented for the small-cap value quintile pf).
Bird and Whitaker (2003)	Returns comparison of quintile pfs formed on B/P, S/P, D/P and E/P. Hps 1, 3, 6, 9, 12, 24, 36 & 48 months.	Worldscope and GMO Woolley	2,219 firms (average) from 7 European developed countries	Jan 1990–Jun 2002	The highest VP and value quintile pf return among 4 IVRs (holds also for 3-year hps). The largest VP in UK.
van der Hart *et al.* (2003)	Performance comparison of EW B/P, D/P & E/P pfs formed from the top & bottom 15% of stocks in each country. Internationally diversified pfs consist of EW country pfs rebalanced monthly (hp 6 months).	IFC Emerging Markets database	684 (average) emerging market stocks	Jan 1985–May 1999	The second highest and significant VP (.83% p.m.), as well as the second highest value pf return (1.83%) among 3 IVRs.
Yen *et al.* (2004)	Performance comparison of 3 pairs of EW value and growth pfs formed on intersection of size and either B/P, CF/P or E/P by first dividing sample into size quartiles and then ranking stocks within them based on IVRs.	PACAP database	All Singaporean non-financial firms	July 1976–Dec 1998	The highest VP (.48% p.m.) and size-adjusted value pf return (1.14%) among 3 IVRs. The return differences between value pfs are marginal (max. .06%).
van der Hart *et al.* (2005)	Performance comparison of EW B/P & E/P pfs formed from the top & bottom 15% of stocks in each country. Internationally diversified pfs consist of EW country pfs rebalanced monthly (hp 6 months).	S&P/IFC Emerging Markets database	576 (average) emerging market stocks	Dec 1988–June 2004	VP is slightly higher based on B/P (.73% p.m.) than based on E/P (.68%). However, the VP difference between the selection criteria varies over time.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (2006)	25 VW pfs formed as intersections of independent sorts of the sample stocks into five size groups and five B/P groups.	MSCI merged data for 14 national markets outside USA	1,421 (average)	1975–2004	Significant overall VP (.53% p.m. return difference between the top-2 vs. the bottom-2 B/P quintiles) that is almost equal among the biggest and the smaller stocks.
Bird and Casavecchia (2007b)	Return comparison of country-adjusted B/P-, S/P-, and E/P-based quintile pf for hp lengths of 3, 6, 12, 24 and 36 months.	Worldscope/ GMO Woolley London and I/B/E/S	Appr. 1,650 stocks from 15 European countries per year (total <8,000)	May 1989–April 2004	The lowest VP for hp lengths of 3, 6 and 36 months. The lowest value pf returns for all hp lengths. The highest B/P-based VP and the highest value pf return for 3-year hp length. VPs are significant for hp lengths longer than 1 year. VPs increase monotonically with hp length.
Kyriazis and Diacogiannis (2007)	Comparison of B/P-, D/P- and E/P-based tertile pf returns & C-SRs.	Datastream/ Athens Stock Exchange.	Max. 260 Greek stocks	1995–2002	The lowest but significant VP (15.0%) and the lowest value pf return (14.3%). No significance in multivariate C-SRs.
Leivo *et al.* (2009)	Performance comparison of quintile pfs formed on 4 IVRs (E/P, EBITDA/EV, B/P and S/P) and three hybrids of them. 3-year hp.	Datastream/ Opstock	74 (average) Finnish non-financial stocks	May 1991–Apr 2006	The highest VP (22.57%), and value pf return (27.83%) among 4 IVRs.
Dissanaike and Lim (2010)	Comparison of long/short decile pfs formed on B/P, CF/P, operating cash flow/P, and E/P, and the Ohlson (1995) model and the residual income model.	Datastream/Inland Revenue database	All listed UK firms (1,791, on average at some point or another) in London Stock Exchange	1987–2001	The lowest, but highly significant VP (8.65%) among 4 IVRs for all hp lengths examined (1,2 and 3years).

(*Continued*)

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Hou *et al.* (2011)	Examination of size, D/P, E/P, CF/P, B/P, leverage, and momentum as potential explanatory factors for global stock returns. VP comparison based on VW quintile pfs formed on 4 IVRs examined.	Datastream / Worldscope	27,488 global common stocks from 49 countries	July 1981–Dec 2003	The lowest VP (.51% p.m.). B/P has significant explanatory power on subsequent returns, but it works poorly as a value-based factor on global returns, unlike CF/P. The incremental explanatory power of a B/P factor-mimicking pf, over and above that based on CF/P, is also negligible.
Huang (2011)	Return comparison of EW and VW quintile pfs formed on B/P.	Taiwan Stock Exchange/GTSM/ Taiwan Economic Journal	Averagely 396 Taiwanese firms per month (range 65–1,117 firms per month)	July 1985–June 2009	The VP is .59% p.m. (.32%) based on EW (VW) returns but insignificant in both cases. The EW (VW) value pf return is 1.04% (.67%).
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream /Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The lowest VP (5.21%) and the second lowest value sextile pf return (15.96%) among all pf formation criteria examined.
Fama and French (2012)	VP comparison of double-sorted pfs (first into size quantiles, then within them into B/P quantiles; 2 x 3 and 5 x 5 sorts on size and B/P).	Bloomberg (supplemented by Datastream and Worldscope)	23 developed stock markets divided into 4 regional markets (North America, Japan, Asia Pacific, and Europe)	Nov 1989–Mar 2011	VP is the largest in Asia Pacific countries based on 2 x 3 sorts among large-cap stocks (.62% p.m.) and the lowest in US (.33%), whereas in 5 x 5 sorts for the full sample it is largest in Japan (.68%) among large-cap stocks, and in Asia Pacific among small-cap stocks (1.22%). VPs are generally larger for small-cap stocks, except in Japan.
Artmann *et al.* (2012)	Return comparison of decile pfs formed on 10 criteria (incl. B/P and E/P). C-SR and asset pricing tests using factor pfs formed on the same criteria.	Karlsrüher Kapitalmarktdaten bank and Saling/ Hoppenstedt Aktienführer	From 175 to 598 (955 altogether) German stocks listed in Frankfurt Stock Exchange	1963–2006	Significant, but slightly lower VP (.9% p.m.) than based on E/P (.99%). The decile return is also slightly lower for B/P value pf (1.20%) than for E/P value pf (1.31%). Both B/P and E/P are significant factors in cross-sectional and asset pricing tests.

(*Continued*)

© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Table 4. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Cakici *et al.* (2013)	VP comparison of double-sorted pfs (first into size quintiles, then within them into B/P quintiles).	Datastream	18 emerging stock markets divided to 3 regional markets (Asia, Latin America and Eastern Europe)	Jan 1990–Dec 2011	VP is larger in emerging markets than in developed markets. The largest VP for both large-cap and small-cap stocks is documented in Eastern Europe markets (2.49% and 3.46% p.m., respectively). For pooled emerging markets, VP is almost identical for the large-cap and small-cap samples.
Gharghori *et al.* (2013)	C-SR and return comparison of EW and VW decile (quintile*) pfs formed on 6 criteria (B/P, E/P*, CF/P*, S/P, debt-to-equity ratio and firm size).	CRIF/Aspect Huntley	778 Australian stocks (average)	Jan 1993–Dec 2009	The highest EW VP (1.5% p.m.) and the highest EW value decile pf return (1.28%) among 4 IVRs. The best explanatory power on subsequent returns among 6 criteria.
Pätäri *et al.* (2015)	Performance comparison of EBIT/EV-, B/P-, E/P- and S/P-based tertile pfs and all their size-, industry- and leverage-adjusted combinations and permutations.	Datastream /Opstock	97 Finnish non-financial stocks (average)	May 1996–Apr 2013	The second lowest VP (5.25%) and the second highest value pf return (12.55%) among 4 IVRs. However, size-adjusted B/P criterion is one component in the best-performing 3-combination strategy, as is also leverage-adjusted B/P criterion in the best 4-combination strategy.

Table 4 summarizes the results from 31 studies on B/P anomalies in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period*, FYE to *fiscal year-end*, C-SR to *cross-sectional regression*, and FF3 alpha to the *Fama-French 3-factor alpha*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

stocks. Instead, Daniel *et al.* (2001) reported the highest value premium in the middle-size quintile for the comprehensive sample of Japanese stocks. According to their results, the B/P-based value premium was significant in all size quintiles other than the smallest-cap quintile, for which the value portfolio return was the highest, indicating that the small-cap anomaly existed among the sample stocks. Parallel results based on recent data were also provided by Fama and French (2012) who, for the sample of 23 developed stock markets divided into 4 regional stock markets, also documented that unlike in other countries, the value premiums in Japan were higher in large-cap quantiles than in small-cap quantiles. By contrast, Fama and French (2006) found in their earlier study that in the U.S. markets, significant B/P-based value premiums (return difference between the top-2 vs. the bottom-2 B/P quintiles) existed for all five size groups, except for the largest-cap group. They also reported that the U.S. value premiums were higher based on B/P than based on E/P in the three biggest size groups, whereas the reverse held for the two smallest-cap groups. Instead, for the merged data from 14 non-U.S. developed national markets, the value premium was almost equal among the largest-cap and the smallest-cap groups, according to the same authors. Cakici *et al.* (2013) also supported the previous finding in their sample of 18 emerging stock markets.

Recent U.S. evidence of B/P-based value premium is provided by Loughran and Wellman (2011) and Israel and Moskowitz (2013), whereas the corresponding global evidence is offered by Fama and French (2012) from developed markets and by Cakici *et al.* (2013) from emerging markets. Both last-mentioned studies document strong and significant value premium in most of the regional markets examined, but there are some fundamental differences between the developed and emerging markets. According to Fama and French's results, the value premium decreased when moving from small-cap samples toward the larger-cap ones. Instead, Cakici *et al.* (2013) documented for the pooled emerging market sample that the value premium is almost identical among the large-cap and small-cap samples. The value premiums, as well as the top value portfolio returns, were higher in emerging markets than in developed markets.

4. Dividend Yield (D/P) Anomaly

The hypothesis that D/P predicts returns has been the subject of considerable theoretical (e.g., see Boudoukh *et al.*, 2008) and empirical research (e.g., see Ball, 1978; Hodrick, 1992; Goetzmann and Jorion, 1993; Kothari and Shanken, 1997; Ang and Bekaert, 2006). Actually, there are at least three central competing hypotheses: the tax-effect hypothesis, the dividend-neutrality hypothesis, and the signalling hypothesis. The tax-effect hypothesis proposed by Brennan (1970) states that investors receive higher before-tax, risk-adjusted returns on stocks with higher anticipated dividend yields to compensate for the historically higher taxation of dividend income relative to capital gain income. In contrast, the dividend-neutrality hypothesis proposed by Black and Scholes (1974) states that if investors require higher returns for holding higher dividend-yield stocks, firms would adjust their dividend policy to restrict the quantity of dividends paid, lower their cost of capital, and thus increase their stock price. Correspondingly, if investors required a lower return on high dividend-yield stocks, value-maximizing firms would increase their dividend payouts to increase their stock price. In equilibrium, value-maximizing behaviour would result in an aggregate supply of dividends to equal the aggregate demand for dividend income from investors that prefer dividends at least as much as capital gains. As a consequence, a predictable relationship between anticipated dividend yields and risk-adjusted stock returns should not exist. According to the signaling hypothesis, dividend yields and their changes reflect the management's beliefs about the future prospects of the firm (e.g., see Dielman and Oppenheimer, 1984; Denis *et al.*, 1994; Sant and Cowan, 1994). Therefore, higher D/Ps could be assumed to signal the management's trust in the continuity of dividend-paying ability.

The prediction power of D/P on stock returns can also be reasoned on the basis of the dividend growth model (Gordon and Shapiro, 1956), according to which the total return of a stock will be determined

© 2015 John Wiley & Sons Ltd

based on its initial dividend yield and the dividend growth rate. In an efficient market, if all stocks with equal risk offered the same total return, the stocks with a low dividend growth would have to offer higher initial dividend yields. However, if investors are incapable of assessing growth prospects correctly, it is possible that the growth rate assumed for high growth rate stocks will be too high and that for low growth rate stocks will be too low. Therefore, overoptimistic growth extrapolations might explain why high D/P stocks would offer a higher total return. A related explanation is that investors do not necessarily behave according to the dividend growth model when pricing stocks. For example, investors might not be indifferent between a stock with a 2% higher initial yield and a stock with a 2% faster growth rate, as put forth by Lofthouse (2001).

Evidence of differences in returns among stocks with high and low or zero-dividend yields has been mixed. Blume (1980) and Keim (1985) documented a U-shaped relationship between risk-adjusted returns and D/Ps, with zero-dividend stocks generating larger returns than dividend-paying stocks and higher D/P stocks realizing larger risk-adjusted returns than lower D/P stocks. By contrast, Christie (1990) showed that the anomalous returns of zero-dividend stocks were largely due to the performance of stocks with a price of less than two dollars during the 1930s. By comparing the returns between zero-dividend and dividend-paying stocks of equal market capitalization, he documented significantly higher size-adjusted returns for the latter type of stocks. Elton *et al.* (1983) also documented a strong positive relationship between D/P and expected returns. In addition, Keim (1985, 1986) found a significant though not linear relationship between D/P and abnormal returns (i.e., Jensen's alphas) in the U.S. market.

In the U.K., Levis (1989) examined the relationship between D/Ps and returns and found that a high D/P and a high return were monotonically positively related. According to his results, the D/P anomaly was the strongest in relation to size, E/P, and stock price anomalies. Although a large degree of interdependency between all four anomalies was documented, Levis reported the D/P and E/P anomalies to subsume the size and stock price anomalies. Controlling for firm size, intra-year seasonality, and market risk, Morgan and Thomas (1998) also found a significant positive relation between D/Ps and returns in the U.K. stock markets. Parallel results from the same stock market were also reported by Chan and Chui (1996) for the 1973–1990 period on the basis of monthly cross-sectional regressions, whereas based on the annual data, the explanatory power of D/P was insignificant. By contrast, the results of Miles and Timmermann (1996) showed that the explanatory power of D/P on the subsequent returns was not significant even in monthly cross-sectional regressions, and in addition, that the D/P-based value premium was negative.[12] McManus *et al.* (2004) introduced the payout ratio into the empirical relationship between stock returns and D/P and found that it had an important impact on the statistical significance of the dividend yield itself in explaining returns, and furthermore, that it conveys signaling information beyond that of D/P.

Naranjo *et al.* (1998) found that both absolute and risk-adjusted returns for NYSE stocks increased with an increasing dividend yield. Consistent with Blume (1980) and Keim (1985), the authors documented higher absolute returns for zero-dividend stocks than for low-dividend stocks, but after the Fama–French 3-factor risk-adjustment, the former stocks performed the worst. Naranjo *et al.* (1998) showed further that tax effects could not explain their findings. Fama and French (1998) compared the value premiums obtained from using four different portfolio formation criteria (i.e., B/P, CF/P, E/P, and D/P) in 13 major stock markets. According to their results, the D/P criterion resulted in the greatest value premium in only one out of 13 national stock markets during the 1975–1995 period. Moreover, the value premium based on D/P was statistically significant in only two national markets. Instead, a comparison of the same four valuation ratios by Bauman *et al.* (1998) documented the greatest value premium based on D/P for a large pooled sample of international stocks whose fiscal year end (FYE) was in March. However, the Sharpe ratios (Sharpe, 1966) of the value quartile portfolios formed on CF/P and B/P were slightly higher than that of the D/P value portfolio for this subsample. Instead, for the subsample of the stocks with FYE in December, the highest Sharpe ratio was shared with the E/P and D/P value quartile portfolios. Based on that evidence, the relative performance of value portfolios based on different valuation ratios also seems to be dependent on the time of the FYE of the sample companies (however, when interpreting the results

© 2015 John Wiley & Sons Ltd

of Bauman *et al.* (1998), it should also be noted that due to varying FYE practices across countries, the different FYE subsamples are "country-biased," because for example, the most common FYE for the Japanese companies is in the end of March, whereas in Australia it is in the end of September, whereas for the majority of U.S. companies the fiscal year equals the calendar year).

Tables 5 and 6 include four U.S. and nine non-U.S. studies in which at least three alternative portfolio formation criteria based on individual valuation ratios, including D/P, have been compared to each other. Somewhat surprisingly, such a literature is very scant, particularly for the U.S. sample data, which might be at least partially explained by the relatively high ratio of non-dividend-paying stocks, as well as by huge variability in the proportion of dividend-paying stocks in the U.S. over time (e.g., see Christie, 1990; Fama and French, 2001). In all such U.S. studies, D/P has generated the lowest value premium as well as the lowest value portfolio return. Instead, for the non-U.S. sample data, the results are more mixed. In two out of nine studies, the highest value premium has been generated by D/P, whereas the highest value portfolio return has been documented for the D/P criterion in three out of eight studies. It is noteworthy that all of the evidence for the superiority of D/P is from the small European national stock markets. At the other end, the lowest value premium and the lowest value portfolio return based on D/P have been documented in three non-U.S. studies.

In addition, when comparing the evidence on the D/P anomaly, the reader should also note that many different practices have been employed in the calculation of dividend yields. For example, Naranjo *et al.* (1998) multiplied a firm's most recently declared quarterly dividend by four and divided the resulting product by the previous month's closing price, whereas Keim (1985) and Christie (1990) used the ratios of the sum of dividend paid over the last 12-month period to the stock price in the beginning of that period. Instead, Fama and French (1993) and Hou *et al.* (2015) divided the similarly calculated sum of total dividends by the stock price at the end of the period preceding the moment of portfolio formation. Moreover, the calculation practices of dividend yields can also vary across the countries in the same databases. For example, according to Hou *et al.* (2011), Worldscope presents all price and per share data (including dividends) on a calendar year-end basis for U.S. firms, but on a fiscal year-end basis for non-U.S. firms. In addition, the group of non-dividend-paying stocks makes the sizes of D/P portfolios very unequal compared to the quantile portfolios formed on some other individual valuation ratios. In spite of weak overall evidence on the efficacy of D/P as a single value measure, D/P may still add value to the investor as a complementary value measure, as argued by Dimson *et al.* (2003). There is also some evidence that high D/P stocks might be less risky (e.g., see Naranjo *et al.*, 1998; Leivo and Pätäri, 2011[13]).

5. Sales-to-Price (S/P) Anomaly

Influenced by Fisher (1984), the use of the S/P criterion became popular during the era of the new economy in the change of the millennium. In those days, analysts found it hard to justify their recommendations on the basis of earnings and book value multiples because negative earnings and low book values were very common among many information, communication, and technology companies. Instead, sales multiples could be calculated even for the most distressed and for very new firms. The use of sales multiples is also motivated in the financial literature by their stability compared to other valuation multiples (e.g., see Bodie *et al.*, 2010), or by the fact that sales are relatively difficult to manipulate, unlike earnings and book values (e.g., see Damodaran, 2012). The biggest disadvantage of using sales multiples is that if a firm generates a high sales growth while simultaneously losing significant amounts of money, S/P could erroneously indicate a low relative valuation for such a firm. Sales can also be increased by increasing debt, which in most cases increases S/P.[14] However, the sales multiples do not indicate whether the sales have been generated without leverage or with maximum leverage, which certainly makes a difference to risks of the firms being compared. In spite of the above-mentioned disadvantages, evidence has shown

© 2015 John Wiley & Sons Ltd

Table 5. Studies on D/P Anomalies in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Blume (1980)	C-SR tests for double-sorted (based on β and D/P and vice versa) 5×5 quantile pfs including firms with zero dividends in separate group.	CRSP/ Compustat	NYSE stocks (sample size is not mentioned)	1936–1976	U-shaped relationship between risk-adjusted returns and D/Ps of pfs. The positive linearity of the relationship is violated by zero-dividend stocks generating the highest returns among the groups. However, high time-variability in prediction power of D/P on subsequent returns exists during 10-year subperiods.
Elton *et al.* (1983)	Performance comparison of 20 D/P pfs and C-SRs.	CRSP	419–880 (average 740) NYSE stocks	1937–1976	Persistent positive relationship between D/P and abnormal returns (ie. alphas of zero-β CAPM) with the exception of zero-dividend stocks.
Keim (1985)	Performance comparison of EW quintile pfs with positive D/P and the zero-dividend stock pf.	CRSP	429–1,289 NYSE stocks	Jan 1931–Dec 1978	U-shaped relationship between returns and D/Ps of pfs. The highest return (21.36%) for zero-dividend stock pf. However, based on Jensen alphas, the top-2 value quintile pfs are better, the top quintile being the best.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 5. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Keim (1986)	Performance comparison of EW quintile pfs with positive D/P and the zero-dividend stock pf and double-sorted (size and D/P) 6×5 pfs, & cross-sectional monthly regressions.	CRSP	920 (average appr.) NYSE stocks	Jan 1931–Dec 1978	After controlling for the size effect, the explanatory power of D/P on subsequent abnormal returns is significant only in January.
Christie (1990)	Size-adjusted return comparison of double-sorted (based on size and D/P) 10×5 quantile pfs including firms with zero dividends in separate group.	CRSP	607 (avrg)–1,486 (avrg) NYSE stocks	1926–1985	The previously documented abnormal returns of zero-dividend stocks are explained by their superior performance before World War II (WWII) period and particularly, during the period following the Great Depression. For the post WWII-period, the relationship between size-adjusted returns and D/Ps of pfs is linear over the full sample.
Fama and French (1993)	Performance comparison of 5 VW quintile pfs and pfs including negative yield stocks (pfs formed on E/P and D/P sorts).	CRSP/ Compustat	All non-financial NYSE, NASDAQ and AMEX stocks	July 1963–Dec 1991	Insignificant VP of .17% p.m. based on return difference between top and bottom quintile pfs. Insignificant FF3 alpha for the top D/P quintile, whereas the corresponding CAPM alpha is significant.

(*Continued*)

Table 5. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Fama and French (1998)	VW return comparison between top and bottom 30% pfs formed on B/P, CF/P, D/P and E/P.	CRSP/ Compustat	3,333–6,258 NYSE, NASDAQ and AMEX stocks	1975–1995	The lowest and significant VP (3.73%), as well as the lowest value pf return (11.75%).
Naranjo *et al.* (1998)	Performance comparison of D/P decile pfs and double-sorted pfs (first based on size into 4 quartiles, then based on D/P into 5 groups of which one is for zero-yield stocks resulting in 4×5 pfs).	CRSP/ Compustat (NYSE stocks)	1,125 (average)	July 1963–Dec 1994	VP of 2.95% based on return difference between top and bottom decile pfs. However, the third highest D/P decile pf generates the highest return (17.12%). VP is slightly higher among small-cap stocks than among large-caps.
Loughran and Wellman (2011)	Return comparison of EW and VW decile pfs formed on 7 formation criteria (B/P, E/P, EBITDA/EV, D/P, sales growth, market leverage and prior 3-year returns) & C-SRs.	CRSP/ Compustat	2,280 (average) NYSE, AMEX and NASDAQ non-financial firms	June 1963- June 2009	The lowest VP of .26% p.m. (.25%) and the value decile pf return of 1.33% (1.1%) in comparison of both EW and (VW) returns.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 5. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Israel and Moskowitz (2013)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and past 5-year return-based VW decile pfs.	French's data library (derived from CRSP /Compustat data)	All US stocks included in decile pfs formed by French	July 1951–Dec 2011	Clearly the lowest VP (2.14% in terms of excess returns), as well as the lowest value pf return (8.73% excess of risk-free rate).
Hou *et al.* (2015)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and 7 other relative value -based VW decile pfs.	CRSP/ Compustat	All non-financial US stocks	July 1972–Dec 2012	Clearly the lowest and insignificant VP (.27% p.m.) among the 4 first IVRs.

Table 5 summarizes the results from 11 studies on D/P anomalies in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period,* C-SR to *cross-sectional regression*, and FF3 alpha to the *Fama-French 3-factor alpha*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

Table 6. Studies on D/P Anomalies in the Non-U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Levis (1989)	Performance comparison of D/P-, E/P-, share price- and size-quintile pfs and the corresponding double-sorted pfs.	LSPD, London Stock Exchange	1,150–2,300 UK stocks	April 1961–Mar 1985	The highest VP (9.4%) and value pf return (18.6%) among 4 criteria examined for the sample that is consistent with that employed for E/P criterion. The D/P anomaly is the strongest of 4 anomalies examined. D/P and E/P anomalies subsume size and stock price anomalies.
Chan and Chui (1996)	C-SR tests of subsequent returns on size, B/P, leverage, share price and D/P.	LSPD/ Exstat Company Accounts Database	851 (895) UK non-financial firms annual (monthly) data (average)	July 1973–June 1990	D/P has explanatory power based on monthly regressions but not based on annual regressions.
Miles and Timmermann (1996)	Performance comparison of EW decile pfs formed on 3 IVRs (D/P, E/P & B/P) & monthly C-SRs.	LSPD/Extel company accounts data	457 UK non-financial firms	May 1979–April 1991	Negative D/P-based VP (-.16% p.m.) The value pf return almost equal to that based on E/P (1.22% vs. 1.21%, respectively). D/P is insignificant variable in monthly C-SRs.
Doeswijk (1997)	Return comparison of B/P-, CF/P, D/P- and E/P-based quintile pfs. Hp lengths 1 year and 3 years.	Amsterdam Stock Exchange/ Manual of Dutch Stocks	From 145 to 183 Dutch stocks per year	July 1976–June 1995	The second highest VP of (7.6%) and the highest value pf return (19.5%) for 1-year hp length.

(*Continued*)

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 6. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Bauman *et al.* (1998)	Comparison of B/P-, CF/P, D/P and E/P-based EW quartile pf performance (examination with four most common FYEs).	Compustat Global Vantage	25,394 stocks (total) from 21 countries included in MSCI EAFE (Europe/Far East/Australia) index + Canada	1985–1996	The second highest, significant VP of 4.8%, but the lowest value pf return (14.1%). The best return/risk ratio among value stocks with FYE in December (shared with E/P) in spite of the lowest return. The highest VP (9.9%) among the 4 IVRs for the firms with FYE in March. VP is positively related to firm size (No VP among companies in the lowest 25% group whereas for the top 25% group, it is 7.6%).
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, D/P and E/P.	MSCI EAFE data	948–1,593 stocks from 12 developed national markets	1975–1995	Significant D/P-based VP in 2 out of 12 countries (Japan & France) whereas insignificant and positive in 7 countries (negative in Germany, Italy & Singapore). The highest VP in 1 out of 12 countries (France).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 6. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Morgan and Thomas (1998)	Performance comparison of EW quintile pfs with positive D/P stocks and the zero-dividend stock pf.	LSPD	<800 - >1500 UK stocks	Jan 1975–Dec 1993	Non-linear relationship between returns and D/Ps. The highest return (2.51% p.m.) is documented for high D/P quintile pf. Jensen alpha of the zero-dividend pf is negative (-.17%), though insignicant in spite of its high absolute return (2.06%). Instead, alphas of the 2 highest D/P pfs are significantly positive (The highest (.53%) for the top D/P pf).
Bird and Whitaker (2003)	Comparison of B/P, S/P, D/P and E/P-based quintile pf returns. The hp lengths 1, 3, 6, 9, 12, 24, 36 and 48 months.	Worldscope and GMO Woolley	2,219 firms (average) from 7 European developed countries	Jan 1990–Jun 2002	Negative VP for 1- and 2-year hp lengths. The lowest value pf returns for all hp lengths examined.
van der Hart *et al.* (2003)	Performance comparison of EW B/P, D/P and E/P pfs formed from the top and bottom 15% of stocks in each country. The internationally diversified pfs consist of EW country pfs rebalanced monthly using 6-month hps.	IFC Emerging Markets database	684 (on average) emerging market stocks	Jan 1985–May 1999	The lowest but significant VP (.42% p.m.), as well as the lowest value pf return (1.76%) among 3 IVRs. The differences between the value pf returns among 3 IVRs are small (max .14 pps).

(*Continued*)

© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90

Table 6. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Kyriazis and Diacogiannis (2007)	Comparison of B/P, S/P, D/P and E/P-based tertile pf returns & C-SRs.	Datastream/ Athens Stock Exchange.	Max. 260 Greek stocks	1995–2002	The highest and significant VP (23.2%), as well as the highest value decile pf return (20.6%). Highly significant explanatory power in C-SRs.
Hou *et al.* (2011)	Examination of size, D/P, E/P, CF/P, B/P, leverage, and momentum as potential explanatory factors for global stock returns. VP comparison based on VW quintile pfs.	Datastream/ Worldscope	27,488 global common stocks from 49 countries	July 1981–Dec 2003	The second lowest VP (.69% p.m.) among the 4 IVRs, though the VP differences between the top-3 IVRs are marginal (max. .06%). D/P has significant explanatory power on subsequent returns, but it works poorly as a value-based factor on global returns, unlike CF/P.
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream/ Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The highest VP (17.51%) among IVRs and the second highest among all pf formation criteria. The best-performing value pf among all pf formation criteria in terms of both absolute (22.76%) and risk-adjusted returns.

Table 6 summarizes the results from 12 studies on D/P anomalies in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, pps to *percentage points*, hp to *holding period*, C-SR to *cross-sectional regression*, and FYE to *fiscal year-end*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

that S/P has worked surprisingly well as a relative valuation criterion (e.g., see Bird and Casavecchia, 2007b; Barbee *et al.*, 2008; Gharghori *et al.*, 2013).

Senchack and Martin (1987) were the first to examine the efficacy of S/P for value portfolio selection. They compared the relative performance of high S/P and high E/P strategies among NYSE and AMEX stocks and found that high S/P stocks produced abnormal risk-adjusted returns compared to both low S/P stocks and the market portfolio. However, high E/P stocks dominated high S/P stocks in terms of both absolute and risk-adjusted returns because the relative performance of the high E/P stocks was more consistent than that of the high S/P stocks. Instead, Barbee *et al.* (1996) found that S/P explained U.S. stock returns better than B/P or firm size. The authors stated further that S/P captures the role of the debt/equity ratio in explaining the returns. By contrast, Mukherji *et al.* (1997), who found evidence of the S/P and the B/P anomaly in the Korean stock markets, showed that the positive relationship of the debt-to-equity ratio persisted in portfolios formed on B/P and S/P.

Dhatt *et al.* (1999) found that for small-cap U.S. stocks, S/P was a better indicator of value than B/P, which in turn was superior to E/P. The same authors also reported the superiority of S/P over B/P, E/P, and CF/P in terms of both value premium and value portfolio returns for the sample of larger-cap U.S. stocks, although the composite value measure based on combining the S/P and E/P criteria generated marginally higher returns (Dhatt *et al.*, 2004). Leledakis and Davidson (2001) also reported higher value decile returns and higher value premium for the portfolios formed on S/P than those based on B/P. For the U.K. sample data, the authors also found that S/P had the best explanatory power on subsequent returns among the four criteria examined (the others were B/P, firm size, and debt-to-equity ratio).

Bird and Casavecchia (2007b) documented the superiority of S/P in the European markets, but found no evidence of added value from combining valuation ratios. Barbee *et al.* (2008) found that in the U.S. stock markets, S/P has the most consistently significant positive relation and the highest explanatory power with subsequent annual returns. According to the authors, S/P is an undervalued value measure, because investors may tend to focus more on E/P and B/P than on CF/P or S/P, resulting in the information contained in the first two multiples being more efficiently incorporated into stock returns than the information in the last two multiples. Recent evidence for the S/P-based value premium has also been provided from the Australian stock market by Gharghori *et al.* (2013), who reported the highest value-weighted value premium based on S/P, as well as the highest value-weighted return for the value decile portfolio formed on S/P. The support for S/P was also given by cross-sectional regression tests in which it appeared to be a significant explanatory factor, although not as significant as B/P.

Table 7 indicates that so far there are only three such U.S. studies in which at least three alternative portfolio formation criteria based on individual valuation ratios, including S/P, have been compared to each other. The number of such studies is surprisingly low in the light of the fact that in every one of these three studies, the highest value premium has been generated by S/P, and in addition, the highest value portfolio return is based on the same criterion in two out three comparable cases.[15] In addition to the evidence based on the portfolio formation approach, S/P was also documented to have the best explanatory power on subsequent returns among the individual valuation ratios in cross-sectional regression tests by Barbee *et al.* (1996, 2008).

For the non-U.S. sample data, the evidence is less favourable for S/P. Table 8 includes eight such comparable studies that have employed non-U.S. sample data. In two out of seven of them, the highest value premium has been generated by S/P,[16] whereas the highest value portfolio return is based on the same criterion in three out of eight cases. By contrast, the S/P-based value premium and the corresponding value portfolio return have been the lowest in an equal number of cases, but it is noteworthy that all this evidence is from the Finnish stock market and based on overlapping sample periods.

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 7. Studies on S/P Anomalies in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Senchack, Jr. and Martin (1987)	Performance comparison of S/P- and E/P-based quintile pfs. Hp lengths 3 months and 1 year.	Compustat	From 400 to 450 randomly selected NYSE and AMEX non-financial stocks per hp	1975–1984	Both S/P-based VP (11.64%) and value pf quintile return (29.63%) are higher than their E/P-based counterparts. In terms of risk-adjusted returns, E/P value pf outperforms the comparable S/P pf.
Barbee *et al.* (1996)	C-SR tests of the prediction power of 4 explanatory variables (ie. S/P, B/P, book debt-to-market equity and firm size) on next year's returns.	CRSP/ Compustat	All NYSE and AMEX non-financial firms	April 1979–Dec 1991	The S/P has the best explanatory power and is the only variable that consistently has a significant role in explaining stock returns in the models being tested. S/P absorbs the role of all the other explanatory variables included in the comparisons.
Dhatt *et al.* (1999)	Comparison of B/P-, E/P- and S/P-based tertile pfs. Firms with negative valuation IVRs excluded from tertile division but included in separate pf for each ratio.	Compustat/ Russell 2000	1,981 small-cap stocks (average)	July 1979–June 1997	The highest VP (8.40%). Instead, the highest value pf return for E/P tertile pf. However, E/P-based tertiles are narrower than B/P and S/P tertiles.

(*Continued*)

Table 7. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Dhatt *et al.* (2004)	Performance comparison of B/P, CF/P and E/P, and S/P-based quintile pfs.	CRSP/ Compustat	1,280–2,314 NYSE, NASDAQ and AMEX stocks	1980–1998	Significant and the highest VP (8.64%) among 16 pf formation criteria including 11 combination criteria. The highest value pf return (20.88%) among 4 IVRs.
Barbee *et al.* (2008)	Return comparison of quintile pfs formed on 4 IVRs (E/P, CF/P, S/P, and B/P) and cross-sectional regression tests based on 30 pfs formed on the basis of each of the 4 IVRs.	CRSP/ Compustat	990–1,860 NYSE, NASDAQ and AMEX firms	1981–2000	S/P has the best explanatory power on subsequent returns. The highest VP (4.69%) and value pf return (15.93%).

Table 7 summarizes the results from five studies on S/P anomalies in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, C-SR to *cross-sectional regression*, and hp to *holding period*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 8. Studies on S/P Anomalies in the Non-U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Mukherji *et al.* (1997)	Return comparison of B/P, E/P and S/P tertile pfs.	PACAP complied by Research Center at the University of Rhode Island.	213 (median) Korean stocks	1982–1993	The second highest VP (15.06%) and the second highest value tertile pf return (25.16%) among 3 IVRs examined.
Suzuki (1998)	Return comparison of yearly reformed pfs, each containing 100 highest B/P, E/P or S/P stocks.	Daiwa Institute of Research / TSE	100 Japanese stocks in each value pf	1983–1996	The second highest value pf return (13.39%). The highest value pf in 6 out of 13 years.
Leledakis and Davidson (2001)	Return comparison of decile pfs formed on 4 criteria (B/P, S/P, debt-to-equity ratio and size) & cross-sectional regression tests.	LSPD/ Datastream	810 UK stocks (average)	June 1980–June 1996	Slightly higher value decile return (33%) than for the B/P value pf. Marginally lower VP (18.6%) for the S/P pfs than for the B/P pfs. The cross-sectional explanatory power of S/P is better than that of B/P.
Bird and Whitaker (2003)	Comparison of B/P, S/P, D/P and E/P-based quintile pf returns. The hp lengths 1, 3, 6, 9, 12, 24, 36 and 48 months.	Worldscope and GMO Woolley	2,219 firms (on average) from European developed countries	Jan 1990–Jun 2002	The second highest VP and value pf return among 4 criteria. The highest S/P-based value quintile pf return, as well the highest VP for 3-year hps.
Bird and Casavecchia (2007b)	Return comparison of country-adjusted B/P-, S/P-, and E/P-based quintile pf for hp lengths of 3, 6, 12, 24 and 36 months.	Worldscope and GMO Woolley London and IBES	Appr. 1,650 stocks from 15 European countries per year (total <8000)	May 1989–April 2004	The highest VP and value pf return for all hp lengths. VPs are significant for hp lengths longer than 6 months. Both VP and value pf return increase monotonically with hp length.

(*Continued*)

Table 8. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Leivo *et al.* (2009)	Performance comparison of the quintile pfs formed on 4 IVRs (E/P, EBITDA/EV, B/P and S/P) and three hybrids of them. 3-year hp.	Datastream/ Opstock	74 Finnish non-financial stocks (average)	May 1991–Apr 2006	The lowest VP (13.85%) and value pf return (23.93%) among 4 IVRs. However, S/P is one component in the composite value measure that performs best in terms of risk-adjusted returns.
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on the basis of 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream/ Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The second lowest VP (5.83%) and the lowest value sextile pf return (13.03%) among all pf formation criteria examined.
Gharghori *et al.* (2013)	Cross-sectional regression and return comparison of EW and VW decile (quintile*) pfs formed on the basis of 6 criteria (B/P, E/P*, CF/P*, S/P, debt-to-equity ratio and firm size).	CRIF/Aspect Huntley	778 Australian stocks (average)	Jan 1993–Dec 2009	The highest VW VP (.96% p.m.) and the highest VW decile pf return (1.16%) among 4 IVRs for 1-year hp lengths. The significant explanatory power on subsequent returns.
Pätäri *et al.* (2015)	Performance comparison of EBIT/EV-, B/P-, E/P- and S/P-based tertile pfs and all their size-, industry- and leverage-adjusted combinations and permutations.	Datastream/ Opstock	97 Finnish non-financial stocks (average)	May 1996–Apr 2013	The lowest VP (.48%) as well as the lowest value pf return (11.22%) among 4 IVRs. However, leverage-adjusted S/P criterion is one component in all of the best-performing combination strategies.

Table 8 summarizes the results from nine studies on S/P anomalies in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, and hp to *holding period*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

6. Cash Flow-to-Price (CF/P) Anomaly

Many financial analysts as well as scholars are cautious with earnings figures because of differences in companies' practices to calculate accruals such as depreciations and amortizations (e.g., see Bildersee *et al.*, 1990; Chan *et al.*, 2006), and, in addition, differences over time in calculation principles of those figures stemming from accounting standards (e.g., Cheng *et al.*, 1996). Many scholars have also shown that accounting losses (i.e. negative earnings) can be regarded as temporary by nature; therefore, they are not reflected in cash flow expectations (e.g., see Hayn, 1995; Martikainen, 1997; Kallunki *et al.*, 1998). The shortcomings of accounting earnings have motivated a number of scholars to explore the relationship between cash flow yields and stock returns (for the first attempts, see e.g., Wilson, 1986; Bernard and Stober, 1989).

To the best of our knowledge, the use of CF/P as a basis of a value investment strategy was first adopted by Chan *et al.* (1991), who compared the efficacy of the CF/P criterion with the E/P, B/P, and size criteria in the Japanese stock market. Their results showed that of the four criteria considered, B/P and CF/P had the most significant positive impact on expected returns. Lakonishok *et al.* (1994) documented parallel results from the U.S. stock markets, with the exception that CF/P was somewhat more efficient for their sample data than B/P, whereas the opposite held for the Japanese sample employed by Chan *et al.* (1991). Both of these cornerstone studies concluded that the observed value premium was not explained by the higher risk (measured by volatility) of value stocks. In the cross-country comparison of value premiums based on four different individual valuation ratios, Fama and French (1998) reported the highest value premium for CF/P in four out of 13 national stock markets. The highest value portfolio return was documented for CF/P in five out of 13 national stock markets, whereas for E/P, it was not documented in any of the markets examined.

The strong performance of CF/P-based strategies relative to E/P-based strategies is also consistent with recent evidence. For example, for the large sample of tradable NYSE and NASDAQ stocks, Dhatt *et al.* (2004) found that among 16 different portfolio formation criteria, which included the size criterion, the B/P, CF/P, E/P, and S/P, and 11 combination criteria formed as the combinations of the four last-mentioned ratios, the lowest risk and the highest return/risk ratio were documented for the CF/P value portfolio. Desai *et al.* (2004) noted that the average annual return for the simple E/P-based market-neutral long/short strategy was 10.2%, whereas for the corresponding CF/P-based strategy, it was 15.3%. Dissanaike and Lim (2010) compared the performance of value strategies based on relatively simple measures, such as B/P, CF/P, E/P, and past return, and some more sophisticated measures, such as those based on the Ohlson (1995) model and the residual income model (suggested by Dechow *et al.*, 1999). For the comprehensive sample of U.K. stocks, the authors found that simple cash flow-to-price measures appeared to perform almost as well as, and in some cases even better than, the more sophisticated alternatives. The value premium based on both standard CF/P, as well as operating cash flow/price, were substantially higher than that based on either B/P or E/P. By contrast, Gregory *et al.* (2001) reported the superiority of B/P and E/P over CF/P in selecting value portfolios in the same stock market.

Hou *et al.* (2011) examined a large number of firm-level characteristics that might explain the cross-sectional and time-series variation in global stock returns. Their analysis included the size, D/P, E/P, CF/P, B/P, leverage, and momentum factors using monthly returns for over 27,000 individual stocks from 49 countries from 1981 to 2003. Using cross-sectional Fama and MacBeth (1973) tests of individual stock returns and time-series regression-based tests of multifactor models, the authors confirmed the strong and reliable explanatory power of a value-based factor in global stock returns. In contrast to almost all preceding comparable studies, this factor was somewhat surprisingly based on CF/P instead of B/P, E/P, or D/P. In addition, the incremental explanatory power of a B/P factor-mimicking portfolio, over and above that based on CF/P, turned out to be negligible.

Table 9 summarizes the studies on CF/P anomaly executed with the U.S. sample data; it introduces eight such papers in which at least three alternative portfolio formation criteria based on individual valuation

© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90

Table 9. Studies on CF/P Anomalies in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Lakonishok *et al.* (1994)	Comparison of B/P, CF/P and E/P-based EW decile pf returns up to 5 years after pf formation.	CRSP/ Compustat	All NYSE and AMEX stocks	Apr 1968–Apr 1990	Significant VP of 9.9% (11.0%) between extreme deciles when using 1-year (5-year) hp. The CF/P-based VP, as well as value pf return are higher than those based on B/P or E/P.
Davis (1994)	Performance comparison of EW quintile pfs formed on β, B/P, market value, E/P, CF/P, sales growth and stock price & C-SRs.	CRSP and Moody's Industrial Manuals	Annual random samples of 100 firms included in the top half in CRSP files based on the market equity	July 1940–June 1963	The second highest VP (9.53%) and value pf return (22.07%) among 3 IVRs examined, although the differences between E/P and CF/P criteria are small or marginal. CF/P has significant explanatory power on subsequent returns and subsumes that of B/P in multivariate cross-sectional tests.
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, D/P and E/P.	CRSP/ Compustat	3,333–6,258 NYSE, NASDAQ and AMEX stocks	1975–1995	The second lowest, but significant VP (6.66%), as well as the second lowest value pf return (13.74%).

(*Continued*)

Table 9. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Dhatt *et al.* (2004)	Comparison of B/P, CF/P and E/P, and S/P-based quintile pf performance.	CRSP/ Compustat	1,280-2,314 NYSE, NASDAQ and AMEX stocks	1980–1998	Significant VP (6.60%). Among 16 pf formation criteria, the best risk-return ratio, as well as the lowest risk, is documented for CF/P. The CF/P-based VP, as well as value pf return (19.44%) are higher than those based on B/P and E/P, but lower than based on S/P.
Desai *et al.* (2004)	Return comparison of B/P, CF/P, operating cash flow/P and E/P decile pfs and C-SR tests.	CRSP/ Compustat	Average 2,823 observations per year, NYSE, AMEX and NASDAQ stocks	1973–1997	The second highest and significant VP (15.3%) and the highest value pf return (28.3%) among 4 IVRs. Operating cash flow/P subsumes abnormal returns related to the 3 other IVRs in cross-sectional tests.
Barbee *et al.* (2008)	Return comparison of quintile pfs formed on 4 IVRs (E/P, CF/P, S/P, and B/P) and C-SR tests based on 30 pfs formed on each of the 4 IVRs.	CRSP/ Compustat	990-1,860 NYSE, NASDAQ and AMEX firms	1981–2000	The lowest VP (1.67%) and value pf return (14.84%).

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 9. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Israel and Moskowitz (2013)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and past 5-year return-based VW decile pfs.	French's data library (derived from CRSP /Compustat data)	All US stocks included in decile pfs formed by French	July 1951–Dec 2011	The highest VP (6.67% in terms of excess returns) and the second highest value pf return (12.15% excess of risk-free rate) among 4 IVRs.
Hou *et al.* (2015)	Performance comparison of B/P-, CF/P-, D/P-, E/P-, and 7 other relative value -based VW decile pfs.	CRSP/ Compustat	All non-financial US stocks	July 1972–Dec 2012	The second lowest but significant VP (.52% p.m.) among 4 IVRs.

Table 9 summarizes the results from 8 studies on CF/P anomalies in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, C-SR to *cross-sectional regression*, and hp to *holding period*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

ratios, including CF/P, have been compared to each other. In two out of eight such studies, CF/P has generated the highest value premium, whereas the CF/P-based value portfolio return has been the highest in two out of seven cases.[17] By contrast, only one of the U.S. studies included in Table 9 has documented the lowest value premium and the lowest value portfolio return for CF/P (i.e., Barbee *et al.*, 2008).

For the non-U.S. sample data (Table 10), CF/P has generated the highest value premium in one out of 10 cases[18]; in addition, based on size-adjusted returns, it has done this in another case, which is actually the only case where the CF-based value portfolio return has also been the highest outside the U.S. markets (i.e., Gregory *et al.*, 2001, for the U.K. data). At the other end, the lowest CF/P-based value premium has been documented in two out of 10 non-U.S. studies, whereas the lowest value portfolio return has been based on the same criterion in only one out of eight cases. Thus, evidence for the relative efficacy of CF/P is slightly stronger in the United States than in elsewhere in the world. Based on the overall evidence and compared to E/P, CF/P has generated higher value premium in 11 out of 17 cases, whereas the value portfolio return has been higher based on CF/P in 10 out of 15 cases.[19]

Analogous to E/P-based value premium, the CF/P-based value premium can also vary remarkably depending on whether the firms with negative cash flows are included in the sample, which makes the comparison of CF/P-based results with the results based on other individual valuation ratios complicated. Because cash flows are in most cases higher than the corresponding earnings, the samples including non-negative cash flows' stocks are generally larger than the samples including only non-negative earnings' stocks (e.g., Chan *et al.*, 1993), which makes the former samples more consistent with B/P samples than the latter samples (i.e., samples of non-negative E/P stocks). Based on the overall evidence, the comparison between CF/P and B/P shows the superiority of CF/P-based value premiums in seven out of 17 cases, whereas based on value portfolio returns, the superiority of B/P has been documented in 11 out of 15 cases.[20] Although the overall evidence slightly favours B/P, the same does not hold for the U.S. evidence, according to which CF/P has resulted in higher value premium in four out of seven studies, whereas based on value portfolio returns, B/P has performed better in an equal number of cases.

7. Enterprise Value-Based Anomalies

So far, enterprise value-based multiples have been seldom examined as a basis for value investing strategy. The reason for the increasing popularity of enterprise value multiples is that they can be compared more easily across firms with diverging leverage, because enterprise value also takes the net value of a company's debt into account. The use of enterprise value multiples as a basis of a value investment strategy has been justified by the fact that if an acquirer is going to buy the entire company, he/she must also take the responsibility of the payment of a debt, therefore allowing for debt in the purchase price. Correspondingly, an investor should not ignore debt, because by buying stocks of a company, he/she is actually buying a part of the company. The most commonly used enterprise value-based multiples introduced in valuation textbooks are earnings before interests and taxes-to-enterprise value (EBIT/EV), earnings before interests, taxes, depreciations and amortizations-to-enterprise value (EBITDA/EV), and sales-to-enterprise value (S/EV). However, their relative efficacy in separating value stocks from glamour or growth stocks has been discussed in only a handful of studies.

One reason supporting EBITDA/EV as a measure of relative valuation is in its use of operating income before depreciation as the profitability measure. Differences in depreciation methods across different companies can cause differences in net income, but they do not affect EBITDA. Of course, the limitations of EBITDA as a measure of profitability should also be borne in mind. As stated by Penman (2013), depreciation is a real economic cost; therefore, pricing a company without considering plant, copyright, and patent expenses would imply that a business could be run without these expenses. Therefore, some scholars have argued for EBIT/EV because EBIT includes depreciations and amortizations, which reflect a firm's capital expenditure in previous years. According to Chan and Lui (2011), EBIT figures can give

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 10. Studies on CF/P Anomalies in the Non-U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Chan *et al.* (1991)	Comparison of B/P-, CF/P- and E/P- and size-based quartile pf returns & C-SRs.	Hamao/ Daiwa Securities	1,215–1,246 Japanese stocks (range of average depending on the formation criteria)	June 1971–Dec 1988	The second highest VP (.79% p.m.) and value pf return (2.22%) among the 3 IVRs. However, higher returns are reported for all 3 pfs that consisted of stocks with negative IVRs, as well as for the small-cap quartile pf. The second best explanatory power in C-SRs (size factor included).
Cai (1997)	Return comparison between the top and bottom decile pfs formed on 4 criteria (B/P, E/P, CF/P and sales growth) for 5-year hps & C-SR tests.	PACAP/ Daiwa Securities/ Nihon Keizai Shimbun	1,186–1,651 Japanese non-financial and non-utility stocks in the first and second sections of TSE	June1971–Dec1993	High VP (11%) among the largest 50%ile of stocks. Positive CF/P ratios have significant explanatory power on subsequent 1-year returns, but it is absorbed by B/P and size factors.
Doeswijk (1997)	Return comparison of B/P-, CF/P, D/P- and E/P-based quintile pfs. Hp lengths 1 year and 3 years.	Amsterdam Stock Exchange/ Manual of Dutch Stocks	From 145 to 183 Dutch stocks per year	July 1976–June 1995	The lowest VP (1.8%) and value pf return (14.4%) for 1-year hp length. For 3-year hp length, the differences in the results for the 4 IVRs are small.

(*Continued*)

Table 10. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Bauman *et al.* (1998)	Comparison of B/P-, CF/P, D/P and E/P-based EW quartile pf performance (examination with four most common FYE).	Compustat Global Vantage	30,240 stocks (total) from 21 countries included in MSCI EAFE (Europe/ Australia/Far East) index + Canada	1985–1996	The lowest but significant VP (4.3%) but the second highest value pf return (15.5%). The best return/risk ratio for the sample of March FYE companies, whereas the lowest for the sample of December FYE companies. The VP is positively related to firm size (no VP among companies in the lowest 25% market-cap group, whereas for the top 25% group, it is 8.2%).
Fama and French (1998)	VW return comparison between the top and the bottom 30% pfs formed on B/P, CF/P, D/P and E/P.	MSCI EAFE data	948 (1975)–1,593 (1995) 12 developed national markets	1975–1995	Significant CF/P-based VP in 4 out of 12 countries (Japan, France, Germany and Australia), whereas insignificant and positive in 7 countries (negative in Netherlands). The highest CF/P-based VPs in 4 out of 12 countries (Germany, Italy, Australia and Hong Kong). The highest value pf return in 5 out of 12 countries (U.K., Germany, Italy, Belgium and Hong Kong).

(*Continued*)

© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Table 10. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Gregory *et al.* (2001)	Return comparison of B/P, CF/P and E/P decile pfs based on 1- and 5-year hp lengths.	Datastream/ LSPD	UK stocks (sample size N/A)	Jan 1975–Dec 1998	The second highest (the lowest) VP of 20.22% (3.92%) as well as the second highest (the lowest) value pf return of 36.49% (24.49%) based on 1-year (5-year) hp.
Yen *et al.* (2004)	Performance comparison of 3 pairs of EW value and growth pfs based on intersection of size and either B/P, CF/P or E/P by first dividing sample into size quartiles and then ranking stocks within them based on IVRs.	PACAP database	All Singaporean non-financial firms	July 1976–Dec 1998	Somewhat lower VP (.34% p.m.) than based on the best IVR (B/P .48%). The value quartile pf return is 1.12%. However, the return difference between value pfs based on CF/P and B/P is marginal (.02% in favour of B/P).
Dissanaike and Lim (2010)	Comparison of long/short decile pfs formed on B/P, CF/P, operating cash flow/P, and E/P, and the Ohlson (1995) model and the residual income model.	Datastream/ Inland Revenue database	All UK firms (1,791, on average at some point or another) listed in LSE	1987–2001	The second highest significant VP (14.42%) among 4 IVRs (The highest (18.44%) based on operating CF/P). Both CF -based criteria worked as well as, and in some cases better than the more sophisticated alternatives.
Hou *et al.* (2011)	Examination of size, D/P, E/P, CF/P, B/P, leverage, and momentum as potential explanatory factors for global stock returns. VP comparison based on VW quintile pfs.	Datastream / Worldscope	27,488 global common stocks from 49 countries	July 1981–Dec 2003	The strongest and reliable explanatory power of CF/P among the factors based on VPs of IVRs examined. CF/P is related to a global covariance risk factor. The second highest VP (.7% p.m.)

(*Continued*)

Table 10. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream/ Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The third highest VP (14.53%) and value sextile pf return (18.10%) among 6 IVRs.
Gharghori *et al.* (2013)	C-SR and return comparison of EW and VW decile (quintile*) pfs formed on 6 criteria (B/P, (E/P*), (CF/P)*, S/P, debt-to-equity ratio and firm size).	CRIF/ Aspect Huntley	778 Australian stocks (average)	Jan 1993–Dec 2009	The value quintile pf return is clearly higher for the sample of stocks with positive CF/Ps than that with positive E/Ps. The same also holds for the corresponding VPs. The best explanatory power on subsequent returns among all 6 criteria for the sample of stocks with positive earnings.

Table 10 summarizes the results from 11 studies on CF/P anomalies in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period*, C-SR to *cross-sectional regression*, and FYE to *fiscal year-end*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

investors better guidance on profit growth and future sustainability, which makes EBIT/EV ratios more reflective of a company's true profitability than EBITDA/EV ratios. However, that does not necessarily imply that EBITDA/EV would be inferior to EBIT/EV as a basis of selection criterion for value portfolios. Actually, the recent results of Gray and Vogel (2012) show that even the more rudimentary profitability measures than EBITDA can be used as a profit component in enterprise value-based multiples. Based on value-weighted returns calculated for quintile portfolios of U.S. stocks, the authors documented both the highest value premium and the highest value portfolio return for the *gross profit*-to-*enterprise value* (GP/EV) ratio, in which gross profit is calculated by subtracting the cost of goods sold from total sales. Their finding is not so surprising in the light of the fact that S/P has also worked well as a relative valuation criterion despite that S/P does not account for the leverage differences between the firms being compared. Although S/EV ratios do so, it is probable that they would also work well as a relative valuation criterion. As EBITDA/EV also worked well for the same purpose in Gray and Vogel (2012) (particularly based on equally-weighted quintile portfolios), it is not surprising that GP/EV also did so (because GP is between sales and EBITDA in income statement). Instead, based on value-weighted quintile returns, *free cash flow*-to-*enterprise value* (FCF/EV) was not as successful as GP/EV and EBITDA/EV as a value portfolio selection criterion, although based on equally weighted portfolios, it slightly outperformed the GP/EV criterion.

To our best knowledge, the first published journal article that examined the performance of EBITDA/EV-ranked quantile portfolios and compared it to the performance of portfolios formed on more commonly used valuation ratios is Leivo *et al.* (2009). Among 20 quintile portfolios formed on four individual valuation ratios (i.e., EBITDA/EV, E/P, B/P, and S/P), the best return/total risk ratios (i.e., the Sharpe and the Sortino ratios)[21] in the Finnish stock markets were documented for the EBITDA/EV value portfolio. The results are consistent with Gray and Vogel (2012) who found the top-quintile EBITDA/EV portfolio the best-performing one among 25 quintile portfolios formed on five individual valuation ratios (i.e., EBITDA/EV, free cash flow/EV, E/P, B/P, and GP/EV). So far the latest evidence of the superiority of enterprise value-based multiples over price-based multiples is provided by Pätäri *et al.* (2015) who, in a performance comparison of tertile portfolios formed on four individual valuation ratios (i.e., EBIT/EV, E/P, B/P, and S/P), documented the best performance for the EBIT/EV value portfolio. The value premium was also higher based on EBIT/EV than on the basis of any of the price-based valuation multiples examined.

Table 11 summarizes the findings of the studies that have examined enterprise value-based anomalies. Both of the studies based on U.S. data included here have documented very promising results on the applicability of EBITDA/EV for portfolio selection criterion. According to Loughran and Wellman (2011), EBITDA/EV-based value premium was the highest in comparison of value-weighted decile portfolios formed on four individual valuation ratios. Based on the corresponding comparison of equally weighted portfolios, EBITDA/EV was the second best after the B/P criterion. EBITDA/EV-based value portfolios also generated the second highest returns after their B/P-based counterparts on both equally and value-weighted bases. Gray and Vogel (2012) reported the highest return based on equally weighted portfolios for EBITDA/EV, followed by FCF/EV and GP/EV. For value-weighted portfolios, the same enterprise value-based multiples also dominated the two price-based multiples (i.e., E/P and B/P), but the rank order based on absolute returns was in this case GP/EV, followed by EBITDA/EV and FCF/EV, whereas based on the Fama-French three-factor alphas, the rank order of the three best individual valuation ratios was the same as in case of equally weighted portfolios. Among three studies from the Finnish stock market, Leivo *et al.* (2009) documented the highest risk-adjusted return among the value portfolios formed on individual valuation ratios for the EBITDA/EV value portfolio, whereas in Pätäri *et al.* (2015), the EBIT/EV-based value portfolio was the best in terms of both absolute returns, Sharpe ratios, skewness- and kurtosis-adjusted Sharpe ratios[22] and size-adjusted alphas among all tertile portfolios formed on four individual valuation ratios. In Leivo and Pätäri (2011), EBITDA/EV generated the second highest value premium after D/P, but the corresponding value sextile portfolio return was only the fourth best

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 11. Studies on Enterprise Value-Based Anomalies in the U.S. and Non-U.S. Stock markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Leivo *et al.* (2009)	Performance comparison of quintile pfs formed on 4 IVRs (E/P, EBITDA/EV, B/P and S/P) and their 3 hybrids (3-year hp).	Datastream /Opstock	74 Finnish non-financial stocks (average)	May 1991–Apr 2006	The second lowest VP (13.92%) but the second highest value pf return (27.34%) among 4 IVRs. Among 20 quintile pfs the best on risk-adjusted basis is EBITDA/EV value pf.
Loughran and Wellman (2011)	Return comparison of EW and VW decile pfs formed on 7 criteria (B/P, E/P, EBITDA/EV, D/P, sales growth, market leverage and prior 3-year returns) & cross-sectional regressions.	CRSP/ Compustat	2,280 NYSE, AMEX and NASDAQ non-financial firms (average)	June 1963–June 2009	The (third) highest VP of .64% p.m. (.82%) in comparison of VW (EW) pfs. EBITDA/EV is slightly better explanatory factor in cross-sectional regressions than B/P.
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream/ Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The second highest VP (15.28%) but the third lowest value sextile pf return (17.55%) among 6 IVRs.
Gray and Vogel (2012)	Performance comparison of EW and VW quintile pfs formed on 5 IVRs (B/P, E/P, EBITDA/EV, FCF/EV and GP/EV).	CRSP/ Compustat	All NYSE, AMEX and NASDAQ non-financial firms with 8-year data for the calculation of IVRS	June 1971–Dec 2010	Two highest EW and VW VPs are generated by EV-based valuation multiples (EBITDA/EV and GP/EV). All 3 EV-based EW and VW value quintile returns are higher than their E/P- and B/P-based counterparts.
Pätäri *et al.* (2015)	Performance comparison of EBIT/EV-, B/P-, E/P- and S/P-based tertile pfs and all their size-, industry- and leverage-adjusted combinations and permutations.	Datastream/ Opstock	97 Finnish non-financial stocks (average)	May 1996–Apr 2013	The highest VP (7.74%), as well as the highest value pf return (14.35%) among 4 IVRs.

Table 11 summarizes the results from 5 studies on enterprise value-based anomalies in the U.S. and non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, and hp to *holding period*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

Table 12. Summary of the Numbers of Top and Bottom Rankings Among Individual Valuation Ratios.

	Value premium			Value portfolio return		
Criterion Panel A (U.S.)	No. of best cases	No. of worst cases	No. of comparable studies	No. of best cases	No. of worst cases	No. of comparable studies
E/P	2	4	12	3	3	11
B/P	4	2	12	3	4	12
D/P	0	4	4	0	4	4
S/P	3	0	3	2	0	3
CF/P	2	1	8	2	1	7
EV	2	0	2	1	0	2
Panel B (non-U.S.)						
E/P	3	6	18	1	6	17
B/P	10	5	18	10	2	17
D/P	2	3	9	3	3	9
S/P	2	2	9	3	3	9
CF/P	1 (2)[a]	2	10	0 (1)[a]	1	8
EV	1	0	3	1(2)[b]	0	3

Table 12 summarizes the numbers of top and bottom rankings given to each individual valuation ratio (the corresponding numbers for all the enterprise value-based multiples are documented in the same row).
[a]Refers to size-adjusted findings.
[b]Refers to risk-adjusted findings.

in comparison of the six selection criteria based on individual valuation ratios (the other four were E/P, CF/P, B/P, and S/P).

Table 12 summarizes the numbers of the top and bottom rankings given to each valuation ratio among such studies in which at least three different individual valuation ratios have been compared. Panel A provides the numbers for the studies that are based on U.S. sample data, whereas Panel B does the same for the studies that employed non-U.S. sample data.

8. Evidence of the Benefits of Composite Value Criteria

The idea of combining value indicators to enhance the value portfolio performance and/or value premium is not new (e.g., see Graham, 1973; Rosenberg *et al.*, 1985). The combination may add value if the value indicators are not highly correlated. Early empirical evidence of the benefits from combining individual valuation ratios was provided by Guerard, Jr. and Takano (1992). However, the existing literature on the empirical tests of the benefits of composite value criteria is relatively young (see Table 13 for the U.S. evidence and Table 14 for the non-U.S. evidence). To the best of our knowledge, Dhatt *et al.* (1999) were the first to report the results of performance comparisons between value portfolios based on both individual valuation ratios and a composite value criterion. The authors formed tertile portfolios of stocks included in the Russell 2000 Index, which is the commonly used U.S. small-cap benchmark, on E/P, B/P, and S/P. The portfolios based on the composite value criterion were formed by combining stocks with consistently low positive values by all three aforementioned valuation ratios into one portfolio, consistently medium positive values into another portfolio, and consistently high positive values into a third portfolio.[23] The composite value portfolio generated the highest return/risk ratios among all of the portfolios that were compared. However, the performance of the tertile portfolios based on individual

© 2015 John Wiley & Sons Ltd

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90

Table 13. Studies on the Efficacy of Composite Value Criteria in the U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Dhatt *et al.* (1999)	Performance comparison of the pfs that included the stocks in the top and bottom tertiles based on all 3 IVRs employed (B/P, E/P and S/P).	Russell 2000 small-cap data	536 (average)	July 1979–June 1997	The average annual return of the combined value pf slightly higher (1.56%) than for the value pf based on the best IVR (E/P). No gain in VP compared to the best IVR (S/P based on VP). No reduction in risk compared to E/P and B/P value pfs.
Piotroski (2000)	Performance comparison between EW B/P value quintile pf and more concentrated double-sorted sub-pfs formed of the same value stocks (first based on B/P, and then based on F_Score).	CRSP/ Compustat	669 firm-year observations per year (average)	1976–1996	Double-sorted sub-pf performed clearly better than single-sorted B/P quintile pf (the corresponding returns were 31.3% vs. 23.9%). However, the double-sorted pf included (on average) only 1/10 of the stocks of the B/P value quintile pf.
Chan and Lakonishok (2004)	Cross-sectional models for predicting future returns based on preceding B/P, CF/P, E/P, and S/P ratios. The estimated slope coefficients determine the weights for each IVR in the composite value indicator.	Russell database	The market-cap deciles of US stocks listed in NYSE, AMEX or NASDAQ (the lowest decile excluded)	1969–2001	VP 11.9% (10.4%) based on return differences between the extreme deciles (quintiles) of the sample of 60 %ile largest-cap stocks. VP 21.1% (16.5%) based on return differences between the extreme deciles (quintiles) of the small-cap stocks (included in the 6th through 9th size deciles).

(*Continued*)

Table 13. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Dhatt *et al.* (2004)	Performance comparison of the quintile pfs formed on 4 IVRs (B/P, E/P, CF/P and S/P) and all their median-adjusted 2-, 3- and 4-combinations.	CRSP/ Compustat	1,280 (1980)–2,314 (1997)	1980–1998	Composite value measures expand the set of efficient pf in the mean-variance framework. Among the 11 combination criteria, the highest VP (8.52%), as well as the highest value pf return (21.12%) is generated by the 2-combination of S/P and E/P.
Bartov and Kim (2004)	Return comparison between B/P top- and bottom quintile pfs and their counterpart sub-pfs formed on 2-dimensional independent quintile sorts (according to B/P and accruals).	CRSP/ Compustat	2,065 NYSE, AMEX or NASDAQ firms (average)	May 1981–April 2000	VP 20.6 % (return for pf of high B/P and low accruals stocks is 22.5%, as for pf of low B/P and high accruals it is only 1.9%).
Bird and Gerlach (2006)	EW and VW performance comparison between B/P value quartile pf and more concentrated double-sorted sub-pfs formed of the same value stocks (first based on B/P, and then based on accounting variables).	Compustat	The quartile of all non-financial US firms in database (229 value stocks, on average)	Apr 1986–Mar 2002	Double-sorted quintile sub-pfs performed slightly better than single-sorted B/P quartile pfs. For more concentrated sub-pfs (1/16 pfs) the returns further increased, but so did also total risk.
Novy-Marx (2013)	Performance comparison between VW B/P value quintile pf and independently sorted 5×5 pfs (based on GP/A, and B/P).	CRSP/ Compustat	3,619 U.S. non-financials firms (average)	July 1963–Dec 2010	The pf of high B/P and high GP/A stocks outperformed the high B/P quintile pf. The VPs were higher in all GP/A quintiles compared to VP based on B/P only.

Table 13 summarizes the results from 7 studies on composite value criteria in the U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period,* and GP/A to the *gross profitability/assets*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

Journal of Economic Surveys (2015) Vol. 00, No. 0, pp. 1–90
© 2015 John Wiley & Sons Ltd

Table 14. Studies on the Efficacy of Composite Value Measures in the Non-U.S. Stock Markets.

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
van der Hart *et al.* (2003)	Performance comparison between EW B/P, D/P and E/P pfs formed from the top and bottom 15%ile of stocks in each country and the corresponding pfs formed on the combination of normalized B/Ps and E/Ps.	IFC Emerging Markets database	684 (on average) emerging market stocks	Jan 1985–May 1999	By combining B/P and E/P, VP is increased by .2% p.m., and the value pf return by .1% compared to the best IVR criterion (ie. E/P).
Chan and Lakonishok (2004)	Cross-sectional models for predicting future returns based on preceding B/P, CF/P, E/P, and S/P ratios. The estimated slope coefficients determine the weights for each IVR in the composite value indicator.	Morgan Stanley database	The largest-cap stocks in the MSCI EAFE Index of developed non-U.S. countries	1989–2001	VP 16.8% (13.5%) based on return differences between the extreme deciles (quintiles).
Ding *et al.* (2005)	Return comparison of 12 VW double-sorted pfs (first into quartiles based on average IVR rankings (B/P, D/P or S/P, and E/P), and then into 1/12 pfs based on 12 3-class criteria).	PACAP database	All non-financial firms in PACAP database	1975–1997	Significant positive VP in Japan, Malaysia and Singapore, insignificant or zero VP in Indonesia and Taiwan, whereas negative in Thailand. Mixed results from the Hong Kong stock market.

(*Continued*)

Table 14. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Guerard, Jr. (2006)	Cross-sectional regression tests for predicting future returns based on E/P, B/P, CF/P and S/P, and in addition, based on their relative historical levels (ie. current level compared to their 5-year average).	PACAP database	All Japanese stocks in both sections of Tokyo Stock Exchange	1982–2001	A slight improvement in prediction power (compared to either each IVR alone or all 4 of them as explanatory factors) when the relative historical levels are also included in the model. No improvement (compared to the best IVRs) without their inclusion.
Bird and Gerlach (2006)	Performance comparison between EW B/P value quartile (for UK sample) or tertile (for Australian sample) pf and more concentrated double-sorted sub-pfs formed of the same value stocks (first based on B/P, and then based on accounting variables).	Compustat/ GMO's proprietary database	The value quartile (tertile) of all non-financial firms in database (for UK (Australian) sample, 89 (50) value stocks per year, on average)	Apr 1990–Mar 2002 (Oct 1990–Sept 2001) for UK (Australian) sample	UK: double-sorted quintile sub-pf as well as more concentrated sub-pf performed clearly better than single-sorted B/P quartile pfs (generating higher return with lower total risk). Australia: double-sorted 27-%ile sub-pf as well as more concentrated sub-pf earned higher return than single-sorted B/P tertile pf but with higher total risk.
Bird and Casavecchia (2007b)	Return comparison of country-adjusted B/P-, S/P-, and E/P-based quintile pf for hp lengths of 3, 6, 12, 24 and 36 months.	Worldscope, GMO Woolley London and IBES from 15 European countries.	Almost 8,000 firms/ appr. 1,650 stocks in each year	May 1989–April 2004	No improvement in performance (over that reported using IVRs) from combinations and permutations of IVRs examined.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 14. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Brown *et al.* (2008b)	Return comparison of top- and bottom-quintile pfs formed on the basis average rankings based on 4 IVRs (B/P, CF/P, E/P and D/P) for the 4 national Asian markets (Hp lengths 1, 2 and 3 years).	Datastream /Woldscope	114 (average) Hongkongese, Korean, Singaporean and Taiwanese stocks that simultaneously belong to top- 50% in terms of market liquidity, market cap and stock price (except that stock price filter in Singapore is top 70% instead of 50%)	1993–2005	For EW returns, the significant VP in Korea (1.77% p.m.) and Hong Kong (.93%), whereas based on VW returns, VP is significant only in Singapore (1.67%). VPs are significant for all hp lengths. VP is significantly negative in Taiwan for 1-year hps for both weighting schemes.
Brown *et al.* (2008a)	Return comparison of top- and bottom-quintile pfs formed on the basis average rankings based on 4 country-adjusted IVRs (B/P, CF/P, E/P and D/P) for the integrated sample of stocks from 4 (3) markets (Hp lengths are 1, 2 and 3 years).	Datastream /Woldscope		1993–2005	Significant VP (1.33% p.m. for EW returns and 1.53% for VW returns) for the 3-country sample (excl. Taiwan) and for 1-year hp. VP is higher and significant among the higher market cap half of the sample, but insignificant among the lower size half.
Leivo *et al.* (2009)	Performance comparison of the quintile pfs formed on 4 IVRs (E/P, EBITDA/EV, B/P and S/P) and three hybrids of them. 3-year hp.	Datastream/ Opstock	74 Finnish non-financial stocks (average)	May 1991–Apr 2006	Among the 7 pf selection criteria, highest value quintile pf return (32.25%) is generated by 2-composite value measure based on median-scaled B/P and EBITDA/EV.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 14. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Pätäri *et al.* (2010)	Performance comparison of tertile pfs formed on DEA scale efficiency scores that implicitly combine E/P, D/P and B/P ratios into a single score. Hp lengths from 1 to 5 years with annual frequency.	Datastream / Opstock	85 Finnish non-financial stocks (average)	1993–2008	VP is at its highest (8.63%) based on 1-year hps, whereas the highest value pf return (20.25%) is achieved based on 4-year hps. However, the performance of value pfs is very even in terms of both absolute and risk-adjusted returns for all hp lengths examined.
Leivo and Pätäri (2011)	Performance comparison of sextile pfs formed on 6 IVRs (E/P, CF/P, EBITDA/EV, B/P, D/P and S/P) and three combination criteria.	Datastream/ Opstock	85 Finnish non-financial stocks (average)	May 1993–Apr 2008	The best composite value pf cannot outperform the D/P value pf, though VP can be somewhat widened by combining median-scaled D/P and EBITDA/EV ratios.
Huang (2011)	Return comparison between B/P quintile pfs and those formed on the basis average rankings based on 4 IVRs (B/P, CF/P, E/P and D/P).	Taiwan Stock Exchange/ GTSM /Taiwan Economic Journal	396 (average) Taiwanese firms per month (range 65–1,117 firms per month)	July 1985–June 2009	Negative but insignificant VP, whereas based on only B/P quintile division it is positive, though still insignificant.
Pätäri *et al.* (2012)	Performance comparison of tertile pfs formed on three variants of DEA efficiency (CRS- (constant returns to scale), cross- and super-efficiency) scores.	Datastream /Opstock	91 Finnish non-financial stocks (average)	May 1994–Apr 2010	CRS-efficiency score produces the highest VP (8.22%), as well as the highest value pf return among the value-only strategies examined. The discriminatory power of DEA efficiency score is very strong.

(*Continued*)

© 2015 John Wiley & Sons Ltd

Table 14. *Continued*

Author(s) and publ. year	Method(s)	Sample database	Sample description and size (if available)	Sample period	Findings
Pätäri *et al.* (2015)	Performance comparison of EBIT/EV-, B/P-, E/P- and S/P-based tertile pfs and all their size-, industry- and leverage-adjusted combinations and permutations.	Datastream /Opstock	97 Finnish non-financial stocks (average)	May 1996–Apr 2013	The best combinations can enhance both VP (by 3.65 pps at their best) and the value pf return (by 2.17 pps at their best) compared to the best IVR. On average, the combination pays off either due to return increase or/and risk decrease.

Table 14 summarizes the results from 14 studies on composite value criteria in the non-U.S. stock markets (EW refers to *equally-weighted*, VW to *value-weighted*, VP to *value premium*, IVR to *individual valuation ratio*, pf *to portfolio*, hp to *holding period*, and pps to *percentage points*). The returns are given on an annual basis, unless indicated by p.m. in the first context within each summary description of the findings (In those cases, the returns are given on a monthly basis throughout the description). The length of holding periods is one year unless otherwise indicated.

© 2015 John Wiley & Sons Ltd

valuation ratios and those based on the composite value criterion is not totally comparable, because the amount of stocks was remarkably lower in the latter case.[24]

A different methodological approach was introduced by Piotroski (2000), who examined whether the performance of the B/P-based value strategy could be boosted by accounting-based fundamental variables. He used the sum of nine individual binary signals (referred as F_Score) as the measure for three areas (i.e., profitability, financial leverage or liquidity, and operating efficiency) of the firm's financial condition. Piotroski (2000) tested the F_Score and its ability to separate "winners from losers" within a quintile portfolio of high B/P U.S. companies. He showed that the mean return earned by a high B/P quintile investor could be increased by 7.5 percentage points annually through the selection of financially strong high B/P firms. However, in spite of numerous robustness tests, Piotroski did not compare the performance of the portfolio formed from the highest F_Score stocks picked from the top-quintile B/P stocks to the same size of portfolios formed on the highest B/P stocks.[25]

To the best of our knowledge, the first study comparing the performance of value portfolios formed from individual valuation ratios with composite value portfolios of equal size is van der Hart *et al.* (2003). For the global sample of emerging market stocks, they used the combination of normalized B/Ps and E/Ps as a basis of composite value measure and found a slight enhancement both in value premium and average return of the top value portfolio in comparison with the results based on the individual valuation ratios. However, the added value from using the composite value measure was rather small for this particular sample. Some other studies analyzing emerging market data have also employed combined value measures, but they have used somewhat simpler nonparametric methodology for allocation of stocks into quantile portfolios. For example, Ding *et al.* (2005) and Brown *et al.* (2008a,b) allocated the stocks into portfolios in accordance with their average rankings based on three or four individual valuation ratios. However, no comparison of the results based on that kind of allocation system and those based on individual valuation ratios is made in any of these three studies.

Chan and Lakonishok (2004) examined the efficacy of combining B/P, CF/P, E/P, and S/P ratios. By employing robust regression methods, they first estimated cross-sectional models that predicted future annual returns from the beginning-year values of each valuation ratio. The estimated slope coefficients determined the weights to be applied to the valuation ratio to arrive at the composite value criterion. The authors tested the efficacy of the above-described portfolio formation criterion with the three different samples, of which the first consisted of six largest-cap deciles of NYSE stocks, the second of the stocks that were in the sixth through the ninth deciles in the same stock exchange, and the third of largest-cap stocks in the MSCI EAFE Europe (Europe/Australasia/Far East) Index of non-U.S. countries. For all the samples examined, Chan and Lakonishok (2004) concluded that the use of the composite value criterion boosted the performance of the value strategy. The authors showed further that the outperformance was not explained by greater risk of value portfolios.

Unfortunately, Chan and Lakonishok (2004) did not report the results based on individual valuation ratios, which would have been an interesting extension for comparability purposes. In contrast, Dhatt *et al.* (2004) did so for the same valuation ratios employed by the former authors. However, the latter authors used a somewhat simpler methodology in constructing the quintile portfolios instead of the decile portfolios employed by the former ones. At the first stage, Dhatt *et al.* (2004) standardized each of the valuation ratios of a firm in a particular year by the median value of that ratio for all the firms in their final sample in that year. At the second stage, composite value measures were computed as simple averages of different combinations of these relative valuation ratios. According to their results, the highest return during was reported for the value portfolio that was based on the combination of E/P and S/P ratios. Although the best return/risk trade-off, as well as the lowest risk, was documented for the CF/P value portfolio, the authors concluded that using composite value measures can expand the set of efficient portfolios, enabling investors to achieve a wider range of return/risk trade-offs.

Bartov and Kim (2004) examined whether the return of the B/P value portfolio could be boosted by including accruals as another selection criterion beside B/P. The authors showed that by picking the

© 2015 John Wiley & Sons Ltd

stocks whose accruals are in the lowest quintile (within the full sample) from the highest B/P quintile, the average annual return for the value portfolio could have been enhanced by 2.4% p.a. The impact of adding accruals as another independent sorting dimension on the value premium was even more dramatic, because the value premium between the aforementioned value portfolio and the comparable glamour portfolio (consisting of those high-accruals quintile stocks that simultaneously belong to the low B/P quintile) was 20.6%, whereas based on the return difference between the top and bottom B/P quintile portfolios, it was 14.1% for the same sample period. However, Bartov and Kim (2004) did not compare the performance of portfolios formed on two independent sorts with that of the quantile portfolios of same size, which leaves the real benefits of the methodology employed open, like in the majority of studies on composite value criteria.

Bird and Gerlach (2006) examined the extent to which fundamental accounting information can be used to better identify truly undervalued stocks. They used Gibbs sampling and model averaging in a logistic regression setting, employing fundamental accounting information as explanatory variables to enhance the performance of value investment strategies in the United States, the United Kingdom, and the Australian stock markets. According to the results, the stocks in the value portfolio that were most likely to show positive abnormal (i.e., above market) returns could be predicted more successfully through the use of fundamental accounting information. The highest added value from including accounting variables beside B/P in portfolio selection was documented in the United Kingdom, whereas in Australia, it was the lowest. Again, the authors did not compare the double-sorted (first based on B/P and then based on accounting variables) sub-portfolios with such B/P value portfolios that would have included the same number of constituent stock series as the former sub-portfolios.

Guerard, Jr. (2006) tested the methodology introduced in Guerard, Jr. and Takano (1992), which they developed to improve the prediction power of valuation ratios on subsequent returns in Japanese stock markets. By using several variants of cross-sectional regression models, he documented a slight improvement in prediction accuracy (compared to that based on the use of either each individual valuation ratio alone or all four of them as explanatory factors) when the relative historical levels of individual valuation ratios quoted as the ratios of their current level to their corresponding five-year averages are added as explanatory factors to the regression model. Instead, no improvement in prediction power (compared to that of the best individual valuation ratios) was documented without them. Moreover, Guerard, Jr. (2006) showed that the prediction power of the model could be further increased by also including one- and two-year projected E/Ps and their monthly revisions in the model.

Using a comprehensive sample of Finnish stocks, Leivo *et al.* (2009) compared the performance of tertile portfolios formed on the two combinations of B/P and EBITDA/EV, the three combinations of B/P, EBITDA/EV, and S/P, and the inverse of the Graham ratio (i.e., the product of E/P and B/P) with the performance of corresponding portfolios formed on individual valuation ratios. The first two combinations appeared to improve risk/return ratios of value portfolios. In addition, the authors noted that abnormal returns (i.e., Jensen's alphas) of value portfolios formed on composite value measures were generally less sensitive to changing stock market sentiment than those based on individual valuation ratios.

Following the methodology developed by Dhatt *et al.* (2004), Leivo and Pätäri (2011) reported the results for nine different value-only strategies that included three combination strategies. The authors compared the performance of sextile portfolios formed on the two combinations of D/P and two three combinations of which the first is based on D/P, B/P, and EBITDA/EV, and the second on D/P, B/P, and E/P. However, none of the combinations examined added value to value portfolio selection because of the superior performance of the D/P-based value portfolio.

To our best knowledge, Pätäri *et al.* (2010) were the first to show the applicability of data envelopment analysis (DEA)[26] for separating value stocks from other types of stocks (i.e., neutral and glamour) on the basis of input and output factors derived from the components of three traditional valuation ratios (i.e., E/P, B/P, and D/P). In their further study, Pätäri *et al.* (2012) examined the added value of using DEA as formation criteria for equity portfolio selection by including several new variables. Although most of

© 2015 John Wiley & Sons Ltd

their results are for criteria that combine the value and price momentum criteria, two criteria are based on composite value-only measures. The results for the Finnish sample data showed that these two criteria were very selective in identifying the best-performing stocks of the future to the extent that not only the DEA glamour tertile portfolio but also the DEA middle tertile portfolio was significantly outperformed by the corresponding value portfolio. However, neither of these studies compares the performance between DEA value portfolios and value portfolios based on other selection criteria.

Novy-Marx (2013) stated that controlling for gross profitability dramatically increases the performance of value strategies. He formed 5×5 portfolios based on independent sorts of *gross profit*-to-*assets* (GP/A) and B/P; furthermore, he documented the best performance for the portfolio that consists of stocks with both high B/P and high GP/A and the worst for the low B/P and low GP/A portfolio. The performance enhancement was particularly evident among the largest, most liquid stocks. However, Novy-Marx (2013) did not report the performance statistics for the B/P portfolios of comparable size. Moreover, the excess returns of the stocks with the lowest profitability in the highest B/P quintile were higher than those of the stocks with the highest profitability in the lowest B/P quintile, although the first-mentioned returns were the lowest among the highest B/P quintile and the latter returns the highest among the lowest B/P quintile.

As several previous papers have shown that firm size, financial leverage, and industry can affect the relative valuation of stocks, Pätäri *et al.* (2015) introduced a new and innovative methodology for equity portfolio selection by adjusting the traditional individual valuation ratios on the basis of firm size, financial leverage, and/or industry classification and combining them as hybrid selection criteria after harmonizing the resulting valuation scores. Their results showed that the simultaneous combining of valuation scores based on different multiples coupled with size, leverage, and industry adjustments pays off to the value investor. The top-tertile portfolios formed on the best combination criteria significantly outperformed not only the stock market portfolio but also the corresponding middle and bottom tertile portfolios. Moreover, the middle-tertile portfolios outperformed the comparable bottom-tertile portfolios for the same criteria. Consistent with previous literature, the division of the full sample period into bullish and bearish periods revealed that the outperformance of top-tertile value portfolios against the market portfolio is mostly attributed to the fact that they lose far less of their value during the bearish conditions than the market portfolio, or the middle- or bottom-tertile portfolios.[27] Thus, the use of multidimensional criteria offered better downside protection against stock market declines than one-dimensional valuation criteria or traditional valuation ratios.

Although almost all of the studies on the efficacy of composite value criteria have concluded that the combination adds value to the investor, only a few of them have provided impartial performance comparisons between portfolios based on such criteria and those formed on individual valuation ratios. To the best of our knowledge, the only exceptions are van der Hart *et al.* (2003), Guerard, Jr. (2006), Bird and Casavecchia (2007b), Leivo *et al.* (2009), Leivo and Pätäri (2011), and Pätäri *et al.* (2015). However, the overall evidence of the added value of combined strategies is relatively weak, because in only one of these six studies did the combination result in both higher returns and lower risk in comparison with the best criterion based on individual valuation ratios (i.e., in Pätäri *et al.*, 2015). Nevertheless, this branch of literature is relatively young; therefore, further empirical studies are required to see the real potential of composite value criteria.

9. Explanations of Value Premium

The reasons for both value anomalies and value premium are widely discussed in the financial literature. The explanatory studies are usually divided into three categories. The first of these relies on risk-based explanations stating that the value premiums and/or related anomalies stem from hidden risk factors that have not been taken into account in the risk-adjustment procedures employed in the value premium studies. The second group of explanations is based on the assumption of irrational behavior of investors,

© 2015 John Wiley & Sons Ltd

which results in the mispricing of assets. The third group of studies asserts that value premium or related anomalies are artefacts of data snooping bias or other biases related to data or data processing. Of course, these explanations can also be intertwined so that the value premium and/or related anomalies are partially explained by omitted risk-based factors and partially by behavioural reasons, as well as by data-related biases. This section reviews and synthesizes the literature on explanations of value premium starting from risk-based reasons, which are often connected to the analysis of B/P-based value premium.

Fama and French (1992) reported in their seminal study that size and B/P explain most of the anomalous differences in future stock returns. However, Daniel and Titman (1997) showed that, after controlling for size and B/P, returns are not strongly related to market βs calculated on the basis of the Fama–French 3-factor model (for a contrary view on this inference, see Davis *et al.*, 2000). In contrast, Ang and Chen (2007) argued that when the tests allow for time-varying market βs, no evidence against a CAPM story for the value premium is left. However, Fama and French (2006) showed that the inferences of Ang and Chen (2007) were valid only for the 1926–1963 period, and furthermore, that during the 1963–2004 period, the value stocks had lower βs than the growth stocks, contrary to the CAPM requirements for explaining the value premiums. Moreover, contradicting the findings of Loughran (1997), Fama and French (2006) showed that the value premium is not restricted to small-cap stocks by rejecting the CAPM pricing formed on size, B/P, and market β during the 1928–2004 period. Daniel and Titman (2006) argued that the B/P anomaly is driven by overreaction to the part of the B/P ratio that is not related to accounting fundamentals. The other part of the B/P ratio that is related to the fundamentals did not appear to forecast returns, thus casting doubts on the explanation according to which violations of the CAPM could be captured by controlling for size and B/P effects that have been interpreted to represent proxies for distress risk by the proponents of market efficiency.

Fama and French (1993) suggested that the value premium exists to compensate investors for the risks inherent in value stocks relative to growth stocks, which are not captured by the traditional CAPM of Sharpe (1964), Lintner (1965), and Mossin (1966). Using the neoclassical framework with rational expectations and competitive equilibrium, Zhang (2005) came to a parallel conclusion, but explained the value premium with the difference between value and growth companies in their ability to adjust the level of production to match the demand in varying economic conditions (i.e., the differences in operating leverage). The conclusion that the B/P anomaly relates to operating leverage is also supported by Carlson *et al.* (2004) and García-Feijóo and Jorgensen (2010). Moreover, Petkova and Zhang (2005) showed that the economic fundamentals of value firms responded negatively to economic shocks, whereas the same did not hold for growth stocks. They interpreted this as evidence that value stocks are riskier than growth stocks, at least in the adverse states of the world. Cooper (2006), Li *et al.* (2009), and Gulen *et al.* (2011) also agreed that the value premium is explained by the lesser flexibility of value firms in adjusting to worsening economic conditions compared to growth firms. Empirical evidence of the "operating leverage hypothesis" is also given by Novy-Marx (2011), who found that the value premium is driven by intra-industry differences in firms, and not by industry characteristics. Instead, Fong (2012) found no evidence of macroeconomic risks explaining the value premium. Guthrie (2013) also showed that the value premium exists even without operating leverage or an industry-wide investment effect, according to which investments by other firms in the same industry buffers demand shocks, reducing the risk of assets in place in high-demand states (e.g., see Kogan, 2004; Aguerrevere, 2009). Though both operating leverage and industry-wide investment contribute to the observed B/P anomaly, they are not entirely satisfactory as explanations of this phenomenon (Guthrie, 2013).

When seeking further explanations for the B/P anomaly, Fama and French (1995) concluded that low B/P firms typically have high average returns on capital, and moreover, that high B/P companies are relatively financially distressed. The authors showed further that low B/P companies remained more profitable for at least five years after the portfolio formation, but that the earnings growth rates of high B/P firms became more similar to low B/P firms after the portfolio formation. They also found evidence that the market does not understand this convergence of earnings growth and that the market seems merely

© 2015 John Wiley & Sons Ltd

to extrapolate the strong earnings growth of low B/P firms and the weaker growth of high B/P firms. Similar findings were also reported by Chan *et al.* (2003), who showed that the market estimates the earnings growth of high B/P stocks too low, leading to a mispricing of stocks due to an over-pessimistic extrapolation of previous growth. The interpretation is also consistent with the conclusion of Penman (1996), who used the residual income valuation framework to illustrate expectations embedded in the price of a high B/P company.

Vassalou and Xing (2004) showed that the B/P anomaly only exists in the two quintiles of the highest default risk, indicating that the B/P anomaly is for the most part related to the financial distress variable. In contrast, the role of financial distress as an explanation for the B/P anomaly is questioned or rejected in many papers. For example, Dichev (1998) found that the relation between bankruptcy risk and book-to-market is not monotonic. Although distressed firms generally have high B/Ps, the most distressed firms have lower B/Ps; therefore, a return premium related to default risk cannot fully explain the B/P anomaly even if distress was rewarded by higher returns. Similar conclusions were also drawn by Griffin and Lemmon (2002) and Campbell *et al.* (2008), who used Ohlson's (1980) O-score model and a dynamic logit model, respectively, to predict defaults.

Piotroski (2000) also supported the argument made by Fama and French (1995) and suggested that accounting fundamentals such as leverage, liquidity, profitability trends, and cash flow adequacy could also be used in discriminating between companies within the high B/P set of firms. Previous literature has also shown that an average high B/P firm is in many cases neglected by the market and followed by fewer investors or analysts (see, e.g., Griffin and Lemon, 2002; Jegadeesh *et al.*, 2004; Doukas *et al.*, 2005). This would also support the usefulness of financial statement analysis on high B/P firms, because the market is more likely to misprice companies that are not actively followed by investors.

Penman *et al.* (2007) suggested that the B/P ratio could be decomposed into an enterprise B/P ratio and a leverage component reflecting financial risk. The authors also showed that as high B/Ps are associated with high returns, the leverage component is negatively associated with the returns. This suggests that the B/P-based value premium could be further enhanced if the leverage-related factors could be taken into account in the portfolio formation. However, this result is contrary to the belief that a higher amount of leverage and risk should yield higher excess return as a reward for the leverage risk, when the effect is in fact the opposite (for recent evidence of this, see e.g., Campbell *et al.*, 2008; George and Hwang, 2010; Obreja, 2013). Penman *et al.* (2007) suggested that this finding could be due to one or more of the following explanations: measurement error in leverage, omitted operating risk factors that are negatively correlated with leverage, or mispricing of leverage by the market. Although the reason for the leverage effect was not explained, its existence at least on some level supported the conclusion that mispricing could happen within high B/P companies and that it might be exploited.

Fama and French (2007a) traced three sources of the value premium: first, it is contributed by the value stocks that improve in type because their companies are acquired by other companies or because they earn high returns and migrate to a neutral or growth portfolio. Secondly, the value premium is attributed to the poor performance of some growth stocks earning low returns and thus moving to a neutral or value portfolio. The third reason for the value premium is the slightly higher returns of value stocks that do not migrate compared with the returns of corresponding growth stocks. In another related study, Fama and French (2007b) found the convergence in B/Ps of value and growth portfolios, which is caused by mean reversion in profitability and expected returns: B/Ps of value portfolios tend to fall as some value companies become more profitable, whereas B/Ps of growth portfolios rise as growth companies cannot reach the profitability level that is expected from them.

An alternative explanation for the value anomalies is based on the irrational behavior of investors, first proposed in the 1930s by Graham and Dodd (1934). This interpretation was re-launched in the theory of investments in the form of De Bondt and Thaler's (1985) overreaction hypothesis. The conclusion was supported by the results of Chopra *et al.* (1992) and Lakonishok *et al.* (1994), who applied it in the context of examining the value premium and drew conclusions parallel to the reasoning of the original

© 2015 John Wiley & Sons Ltd

authors. Moreover, Lakonishok *et al.* found little, if any, support for the view that value strategies are fundamentally riskier than glamour strategies. The similar conclusion was also drawn by Haugen and Baker (1996) and Brennan *et al.* (1998). Barberis *et al.* (1998) also stated that the naïve extrapolation of past growth causes stock prices to overreact in both directions, resulting in return predictability on the basis of valuation ratios.

In contrast, Doukas *et al.* (2002) showed that the value premium is not explained by over-optimism in analysts' EPS forecasts, thus rejecting their non-risk-based explanation of the value premium. In their follow-up paper, the same authors found support for the risk factor explanation as the source of value premium when using the standard deviation of analysts' EPS forecasts as a risk proxy (Doukas *et al.*, 2004). The authors suggested that the abnormal returns of value stocks reflect compensation for higher risk as measured by the dispersion in analysts' EPS forecasts.

According to Daniel *et al.* (2001), investors' overconfidence induces overreaction, and extreme B/P ratios are caused by overreactions to private signals. Phalippou (2008) showed that the value premium is concentrated in stocks mostly held by individual investors and declines from the lowest to the largest institutional ownership decile, consistent with behavioural explanations, the value premium. Parallel conclusions were also drawn by Bartov and Kim (2004), who found that the B/P anomaly is stronger among firms held primarily by small (unsophisticated) investors and followed less closely by market participants than among firms with considerable institutional ownership and analyst coverage. The authors divided their sample into two subsamples based on stock price; they documented clearly higher value premiums for the sample of stocks whose price was below $10 than for those whose price was at least $10.[28] Thus, the mispricing explanation for the B/P anomaly held primarily to a subset of stocks with unsophisticated ownership, as investment professionals are typically unable to invest in firms with a stock price less than $10 due to institutional restrictions. Consistently with the mispricing explanation, Ali *et al.* (2003) showed that the B/P anomaly is stronger among stocks with a higher idiosyncratic return volatility, higher transaction costs, and lower investor sophistication.

The recent results of Piotroski and So (2012) and Hwang and Rubesam (2013) also support the mispricing hypothesis. The first-mentioned authors found that prices of glamour (value) firms reflect systematically optimistic (pessimistic) expectations. Thus, the value/glamour effect should be concentrated (absent) among firms with (without) *ex ante* identifiable expectation errors. Classifying firms based upon whether expectations implied by current valuation multiples are congruent with the strength of their fundamentals, Piotroski and So (2012) documented that value/glamour returns and *ex post* revisions to market expectations were predictably concentrated (absent) among firms with *ex ante* biased (unbiased) market expectations. Hwang and Rubesam (2013) found that the large and positive average value premium comes from the correction of mispricing, which is accelerated during the bearish market sentiment. According to their results, the correction is more severe for value stocks than for growth stocks due to the higher uncertainty of value stocks. As a result, value stocks tend to outperform growth stocks during bad times, because the underpricing of value stocks is corrected faster during bear markets when volatility increases. Arnott and Hsu (2008) also argued that both size and value anomalies are driven by pricing noise, but the authors did not exclude the possibility that such anomalies could also be partially driven by hidden risk factors or behavioural irrationalities.

A third group of explanations for the existence of the value premium relies on the data snooping bias or other biases related to data (e.g., see Black, 1993; Conrad and Kaul, 1993; Ball *et al.*, 1995; Kothari *et al.*, 1995; Conrad *et al.*, 2003). However, in the light of recent results on the value premium documented all around the world, it seems unlikely that all of the evidence of its existence could be explained by these types of biases (e.g., see Markowitz and Xu, 1994; Guerard, Jr. *et al.*, 2012). As the ongoing academic debate on the reasons for the value premium indicates, the research community is still far from a consensus in this respect.

© 2015 John Wiley & Sons Ltd

10. Conclusions

This paper provides an extensive literature review on value premium and the related anomalies. The current literature is based mainly on two different approaches, the first of which uses cross-sectional regression models in explaining future returns with the variables based on the most recent information from stock markets and/or financial statements. The other approach divides the full sample of stocks into quantile portfolios based on the same type of information and then compares the performance of quantile portfolios to each other and/or to the stock market average. It is noteworthy that these two approaches can produce different conclusions about relative efficacy of individual valuation ratios even for the same sample data because the outperformance of a quantile portfolio may stem from the superior performance of the minority of stocks included in that specific quantile, whereas the cross-sectional regression may not find significant causality due to the diversity of performance of individual stocks included in the same quantile.

Among individual valuation ratios, the two most-examined are clearly B/P and E/P, of which the former has in most cases generated both the higher value premium and the higher value portfolio return. However, the pairwise efficacy comparisons of individual valuation ratios based on overall evidence are troublesome, because the samples used for the E/P criterion are often much smaller than those used for B/P due to the commonly used methodological choice to exclude negative earnings' firms from the former samples. The same kind of dilemma is also faced in portfolio divisions based on D/Ps, due to a remarkable subsample of zero-dividend stocks. However, based on overall evidence, D/P has been clearly the least efficient criterion among individual valuation ratios, particularly in the U.S. markets, though some exceptions from small European national markets where it has been the best have been documented. By contrast, the S/P criterion has proven to be very successful in the U.S. markets in those few studies in which it has been employed. In the light of those results, the number of studies that have included S/P as one portfolio formation criterion is surprisingly low, which provides a clear research gap for forthcoming studies. For the non-U.S. sample data, the evidence on the efficacy of S/P is more mixed than it is for the U.S. sample data.

Based on overall evidence, the CF/P criterion has clearly performed better than E/P, particularly in that it has very seldom been the worst criterion in such studies where three or more individual valuation ratios have been compared. Analogous to E/P-based value premiums, the CF/P-based value premiums can also vary remarkably depending on whether or not the firms with negative cash flows are included in the sample. In comparison with B/P, CF/P has performed approximately equally in U.S. markets, whereas in non-U.S. markets, overall evidence is somewhat inclined toward B/P.

For the present, the academic literature on the efficacy of enterprise value-based valuation ratios has been scant, but the results on their relative efficacy have been promising. In the light of those few results, it is very likely that enterprise value-based multiples will get more emphasis in forthcoming studies. One clear research gap related to enterprise value-based multiples is that so far, to the best of our knowledge, none of the studies have included the S/EV criterion in their comparisons. This is even more surprising in the light of the good performance of the S/P criterion, because S/EV has more solid theoretical foundations than S/P, and in addition, may benefit from the lack of academic research on the related anomaly.[29]

This literature survey reviewed 10 such studies that reported the performance enhancement for value portfolios formed on composite value criteria compared to their counterparts formed on individual valuation ratios. However, in the majority of these studies, the value portfolios based on the former criteria were much narrower than the portfolios of individual valuation ratios. Thus, the real performance enhancement stemming from the use of composite value criteria is questionable in all such studies. After careful review, we identified only one study in which the use of composite measures had resulted in both higher return and lower risk without narrowing the value portfolios being compared. However, the literature on composite value criteria is young and based for the most part on very simple methods of combination. Therefore, it is too early to draw such a conclusion that the composite value criteria

© 2015 John Wiley & Sons Ltd

could not add value to the investor,[30] although on the other hand, the use of more sophisticated portfolio formation criteria does not necessarily result in better performance. However, it is likely that their use in academic research will proliferate in the near future. In conclusion, overall evidence shows that the best criteria may vary over time[31] and across markets. The efficacy ranking of the valuation criteria may also depend on methodological choices such as general research design (cross-sectional regression or portfolio formation approach), frequency of quantile division of portfolios, return calculation practice (equally or value-weighting), performance metrics employed, treatment of firms with negative valuation ratios (to include or exclude), and the lengths of the holding and selection periods, for example.

Notes

1. By definition, enterprise value-based multiples include all such valuation ratios in which some output variable is contrasted with enterprise value (the most common output variables used as numerators of enterprise value-based multiples are earnings before interests, taxes, depreciations and amortizations (EBITDA), earnings before interests, and taxes (EBIT) and sales). The denominator of the ratio is usually calculated as the sum of the market value of the firm's equity and the net debt of the firm. Correspondingly, net debt is calculated by subtracting liquid assets (i.e., total cash and cash equivalents) from interest-bearing debt.
2. In some studies, valuation ratios are expressed as input/output ratios, such as P/E (e.g., see Dhatt *et al.*, 2004), P/B (e.g., see Fama and French, 2007a), or EV/EBITDA (e.g., see Loughran and Wellman, 2011). For the sake of clarity and ease of comparability, we have converted the notation of such studies to match with the practice of expressing them as output/input ratios while reviewing such literature. Naturally, we have also taken this into account when interpreting the results of such studies.
3. Recently, parallel results about the insignificance of the size factor and the significance of the E/P factor were also reported by Artmann *et al.* (2012) for the large sample of German stocks over the 1963–2006 period.
4. The formula for the Fama–French three-factor model is as follows:
$$r_{it} - r_{ft} = \alpha_i + b_i(r_{mt} - r_{ft}) + s_i SMB_t + h_i HML_t + \varepsilon_{it},$$
where r_{it} is the return of a portfolio; r_{ft} is the risk-free rate of return; alpha is the abnormal return over and above what might be expected based on the three-factor model employed; r_{mt} is the stock market return; SMB_t is the return of size factor (i.e., the return difference between small- and large-cap portfolios); HML_t is the return of book-to-market (B/P) factor (i.e., the return difference between high and low B/P portfolios); and b_i, s_i, and h_i are factor sensitivities to stock market, SMB, and HML factors, respectively. ε_i is the residual term.
5. Throughout all of the tables of this study, the results documented are those that are the most relevant for comparability purposes. Therefore, the reported results are not necessarily the main results of the studies. For example, in efficacy comparisons of different individual valuation ratios, the sample sizes often vary depending on what valuation criterion is used. (Typically, the samples employed in testing the E/P anomaly are narrower than those used for testing the B/P anomaly. If the original authors have also reported the results for the samples that are of equal size, we have picked those results in tables for the sake of comparability). In addition, many of the results included in the tables have arisen as by-products of the studies whose primary goal was something other than the efficacy comparison of value indicators.
6. In addition to Davis (1994), Dhatt *et al.* (1999), and Li *et al.* (2009) have documented such results. However, the results of Dhatt *et al.* are for the sample of small-cap U.S. stocks. Moreover, their results can be explained by the exclusion of negative earnings' stocks from the sample, which has

© 2015 John Wiley & Sons Ltd

made the top tertile portfolio formed on E/P smaller than the corresponding portfolios formed on B/P and S/P).

7. The total number of comparable studies included in this comparison is reduced by one because Suzuki (1998) did not report the value premiums. In addition, Fama and French (1998) reported the highest value premium based on E/P in two of 13 developed stock markets.
8. In van der Hart *et al.* (2003). The difference in the number of comparable studies stems from the fact that Dissanaike and Lim (2010) and Hou *et al.* (2011) report only the value premiums, but not the returns of value portfolios.
9. In addition, Fama and French (1998) reported the lowest value premium based on E/P in three of 13 developed national stock markets.
10. For example, Basu (1977) reported higher returns for the low E/P U.S. stocks when negative earnings' stocks were excluded than in the case when they were included. By contrast, Jaffe *et al.* (1989) found evidence of consistently high returns for firms of all sizes with negative earnings. For an extreme example, Chan *et al.* (1991) documented the highest returns for the portfolio that included only negative earnings' firms in the Japanese stock markets.
11. Furthermore, Fama and French (1998) documented that in nine out of 12 national developed stock markets, the value premium was higher based on B/P than on the basis of E/P. In eight out of 12 cases, the B/P-based value portfolio return was higher than its E/P-based counterpart.
12. Different taxation on dividends and capital gains must also be taken into account when comparing the results from different national markets. U.S. tax law has treated capital gains more favourably than dividends, and therefore, taxable investors may have demanded a higher pre-tax return on higher-yielding stocks to compensate for the increased tax liability. By contrast, in some other countries, such as the United Kingdom, for example, the tax rates have been lower for dividends than for capital gains. On the other hand, many of the results for U.K. sample data are consistent with U.S. results (e.g., see Morgan and Thomas, 1998), whereas contradictory results have been reported from the same national market. Therefore, it seems that the results are in this sense more dependent on the sample period or methodological issues than on taxation differences between dividends and capital gains.
13. Moreover, Leivo and Pätäri (2011) showed that selecting value stocks based on the D/P criterion resulted in the largest proportion of stocks that have provided a higher total return than the stock market average during the subsequent one-year holding period.
14. In exceptional cases, it may happen that, as a result of higher leverage, the stock price may rise by a higher percentage than that of sales growth generated by increased debt. However, in most cases, the impact of increased debt on the S/P ratio is positive (e.g., see Damodaran, 2012 for details).
15. In Dhatt *et al.* (2004) and in Barbee *et al.* (2008). The reader should also note that in the only exceptional case (i.e., Dhatt *et al.*, 1999), in which the S/P-based value portfolio has not generated the highest value portfolio return despite that the S/P-based value premium has been the highest, the size differences of the value portfolios can explain the superiority of the E/P value portfolio that is distinctly narrower than its B/P- and S/P-based counterparts.
16. See footnote 7 for the reason for the difference in the number of comparable studies.
17. The difference in the number of comparable studies stems from the fact that Hou *et al.* (2015) report only the value premiums, but not the returns of value portfolios.
18. In Dissanaike and Lim (2010), who documented the two highest value premiums in the United Kingdom for the two cash flow-based multiples (i.e., the highest based on operating cash flow/price and the second highest for the simple cash flow/price in which cash flow is calculated as current earnings plus depreciation/amortization plus provisions. The reader should note that there are several ways to calculate cash flow component in CF/P ratios, for example, Lakonishok *et al.* (1994), Fama and French (1998), and Yen *et al.* (2004) used earnings plus depreciations, whereas Hou *et al.* (2011) used net income plus depreciation and amortization plus income statement deferred taxes from the

© 2015 John Wiley & Sons Ltd

sum of which Dhatt *et al.* (2004) subtracted preferred dividends. Instead, Israel and Moskowitz (2013) and Hou *et al.* (2015) used income before extraordinary items plus equity's share of depreciation plus deferred taxes).

19. In addition, Fama and French (1998) documented that in nine out of 12 developed non-U.S. stock markets, the value premium was higher based on CF/P than on the basis of E/P, whereas in comparisons of value portfolio returns, CF/P was superior to E/P in eight out of 12 cases.
20. According to Fama and French (1998), in five out of 12 national developed non-U.S. stock markets, the value premium was higher based on CF/P than based on B/P, whereas in comparisons of value portfolio returns, CF/P was superior to B/P in six out of 12 cases.
21. The Sortino ratio is calculated by dividing the average excess return of the portfolio by the semi-standard deviation of the corresponding period returns (see Sortino and van der Meer, 1991, for details).
22. The skewness- and kurtosis-adjusted Sharpe ratio was first introduced in a hedge fund study of Pätäri and Tolvanen (2009).
23. Stocks with negative values for any of the three valuation ratios were excluded from the sample. The average total number of companies included in three portfolios formed on the composite value measure decreased to 536 from 1958, indicating that the great majority of companies did not have consistently low, medium, or high values of the three ratios used as the basis of the composite value criterion.
24. If the value portfolios based on individual valuation ratios had included only the same amount of stocks as the composite value portfolio, their performance might have been as good as or even better than the performance of the composite value portfolio. In the light of the fact that the value premium is usually stronger when the frequency of quantile division is higher, and on the other hand, that the reported performance differences between tertile portfolios based on individual valuation ratios and those based on the composite value criterion are small, this is likely.
25. This kind of comparison would have been particularly valid and interesting in the light of the fact that, on average, the former portfolios included only one-tenth of the stocks included in the highest B/P quintile portfolios (see Table 3 in Piotroski 2000). Therefore, the seemingly dramatic added value of the use of F_scores beside B/P criterion might be fully or partially explained by the size differences of the portfolios being compared, analogous to the aforementioned case of Dhatt *et al.* (1999).
26. DEA was originally developed by operations researchers for the efficiency comparison purposes (e.g., see Charnes *et al.*, 1978).
27. In this sense, the results of Pätäri *et al.* (2015) are parallel to those of Lakonishok *et al.* (1994), Bird and Whitaker (2003), and Hwang and Rubesam (2013).
28. For the first subsample, the value premium based on the return difference of top and bottom quintile portfolios was 20.7%, whereas for the second subsample it was only 5.2%. The high B/P quintile return for the first subsample was 20.9% and 16.3% for the second, indicating that the value premium difference is for the most part explained by poor performance of low B/P stocks whose price is below $10.
29. For example, McLean and Pontiff (2014) documented the degradation of profits from anomalies after their publication.
30. For example, see the empirical evidence in Leivo *et al.* (2009), Israel and Moskowitz (2013), and Pätäri *et al.* (2015), according to whom the combination strategies make the value portfolio performance more stable over time and less prone to varying stock market conditions.
31. See also Davis (2001) for an evidence of time-variability in relative value premiums calculated on the basis of four individual valuation ratios (B/P, CF/P, E/P, and S/P).

© 2015 John Wiley & Sons Ltd

References

Aguerrevere, F.L. (2009) Real options, product market competition, and asset returns. *Journal of Finance* 64: 957–983.

Ali, A., Hwang, L.S. and Trombley, M.A. (2003) Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics* 69: 355–373.

Anderson, K.P. and Brooks, C. (2006) The long term price-earnings ratio. *Journal of Business Finance & Accounting* 33: 1063–1086.

Ang, A. and Bekaert, G. (2006) Stock return predictability. *Review of Financial Studies* 20: 651–707.

Ang, A. and Chen, J. (2007) The CAPM over the long-run: 1926–2001. *Journal of Empirical Finance* 14: 1–40.

Arnott, R. and Hsu, J. (2008) Noise, CAPM and the size and value effects. *Journal of Investment Management* 6: 1–11.

Arshanapalli, B.G., Coggin, T.D. and Doukas, J. (1998) Multifactor asset pricing analysis of international value investment strategies. *Journal of Portfolio Management* 24: 10–23.

Artmann, S., Finter, P. and Kempf, A. (2012) Determinants of expected stock returns: large sample evidence from the German Market. *Journal of Business Finance and Accounting* 39: 758–784.

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

Athanassakos, G. (2011) The Performance, pervasiveness and determinants of value premium in different us exchanges: 1985–2006. *Journal of Investment Management* 9: 33–73.

Ball, R. (1978) Anomalies in relationships between securities' yields and yield-surrogates. *Journal of Financial Economics* 6: 103–126.

Ball, R., Kothari, S.P. and Shanken, J. (1995) Problems in measuring portfolio performance: an application to contrarian investment strategies. *Journal of Financial Economics* 38: 79–107.

Banz, R. (1981) The relationship between return and market value of common stocks. *Journal of Financial Economics* 9: 3–18.

Banz, R. and Breen, W.J. (1986) Sample-dependent results using accounting and market data: some evidence. *Journal of Finance* 41: 779–793.

Barbee, W.C., Jeong, J.G. and Mukherji, S. (2008) Relations between portfolio returns and market multiples. *Global Finance Journal* 19: 1–10.

Barbee, W.C. Jr, Mukherji, S. and Raines, G.A. (1996) Do sales-price and debt-equity explain stock returns better than book-market and firm size? *Financial Analysts Journal* 52: 56–60.

Barberis, N., Shleifer, A. and Vishny, R. (1998) A model of investor sentiment. *Journal of Financial Economics* 49: 307–343.

Bartov, E. and Kim, M. (2004) Risk mispricing and value investing. *Review of Quantitative Finance and Accounting* 23: 353–376.

Basu, S. (1975) The information content of price-earnings ratios. *Financial Management* 4: 53–64.

Basu, S. (1977) Investment performance of common stocks in relation to their price earnings ratios: a test of the efficient market hypothesis. *Journal of Finance* 32: 663–682.

Basu, S. (1983) The relationship between earnings' yield, market value and return for NYSE common stocks: further evidence. *Journal of Financial Economics* 12: 129–156.

Bauman, S.W., Conover, M.C. and Miller, R.E. (1998) Growth versus value and large-cap versus small-cap stocks in international markets. *Financial Analysts Journal* 54: 75–89.

Beaver, W., McNichols, M. and Price, R. (2007) Delisting returns and their effect on accounting-based market anomalies. *Journal of Accounting and Economics* 43: 341–368.

Bernard, V.L. and Stober, T.L. (1989) The nature and amount of information in cash flows and accruals. *Accounting Review* 64: 624–652.

Bettman, J., Sault, S. and Schultz, E. (2009) Fundamental and technical analysis: substitutes or complements? *Accounting and Finance* 49: 21–36.

Bildersee, J., Cheh, J. and Lee, C. (1990) The International price-earnings ratio phenomenon: a partial explanation. *Japan and the World Economy* 2: 263–282.

Bird, R. and Casavecchia, L. (2007a) Sentiment and financial health indicators for value and growth stocks: the European experience. *European Journal of Finance* 13: 769–793.

© 2015 John Wiley & Sons Ltd

Bird, R. and Casavecchia, L. (2007b) Value enhancement using momentum indicators: the European experience. *Le* 3: 229–262.
Bird, R. and Gerlach, R. (2006) A Bayesian model averaging approach to enhance value investment. *International Journal of Business and Economics* 5: 111–127.
Bird, R. and Whitaker, J. (2003) The performance of value and momentum investment portfolios: recent experience in the major European markets. *Journal of Asset Management* 4: 221–246.
Black, F. (1993) Beta and return. *Journal of Portfolio Management* 20: 8–18.
Black, F. and Scholes, M.S. (1974) The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics* 1: 1–22.
Blume, M.E. (1980) Stock returns and dividend yields: some more evidence. *Review of Economics and Statistics* 62: 567–577.
Bodie, Z., Kane, A. and Marcus, A.J. (2010) *Investments*. Singapore: McGraw-Hill/Irwin.
Boudoukh, J., Richardson, M. and Whitelaw, R. (2008) The myth of long-horizon predictability. *Review of Financial Studies* 21: 1577–1605.
Breen, W. (1968) Low price-earning ratios and industry relatives. *Financial Analyst Journal* 24: 125–127.
Brennan, M. (1970) Taxes, market valuation and corporate financial policy. *National Tax Journal* 23: 417–427.
Brennan, M., Chordia, T. and Subrahmanyam, A. (1998) Alternative factor specifications, security characteriristics and the cross-section of expected stock returns. *Journal of Financial Economics* 49: 345–373.
Brown, S., Du, D., Rhee, S. and Zhang, L. (2008a) The returns to value and momentum in Asian markets. *Emerging Markets Review* 9: 79–88.
Brown, S., Rhee, S. and Zhang, L. (2008b) The returns to value in Asian stock markets. *Emerging Markets Review* 9: 194–205.
Cai, J. (1997) Glamour and value strategies on the Tokyo stock exchange. *Journal of Business Finance and Accounting* 24: 1291–1310.
Cakici, N., Fabozzi, F.J. and Tan, S. (2013) Size, value, and momentum in emerging market stock returns. *Emerging Markets Review* 16: 46–65.
Campbell, J.Y. (2000) Asset pricing at the millennium. *Journal of Finance* 55: 1515–1567.
Campbell, J.Y., Hilscher, J. and Szilagyi, J. (2008) In search of distress risk. *Journal of Finance* 63: 2899–2939.
Capaul, C., Rowley, I., and Sharpe, W.F. (1993) International value and growth stock returns. *Financial Analysts Journal* 49: 27–36.
Carlson, M., Fisher, A. and Giammarino, R. (2004) Corporate investment and asset price dynamics: implications for the cross-section of returns. *Journal of Finance* 59: 2577–2603.
Chan, A. and Chui, A.P.L. (1996) An empirical re-examination of the cross-section of expected returns: UK evidence. *Journal of Business Finance and Accounting* 23: 1435–1452.
Chan, K., Chan, L.K.C., Jegadeesh, N. and Lakonishok, J. (2006) Earnings quality and stock returns. *Journal of Business* 79: 1041–1082.
Chan, L.K.C., Hamao, Y. and Lakonishok, J. (1991) Fundamentals and stock returns in Japan. *Journal of Finance* 46: 1739–1765.
Chan, L.K.C., Hamao, Y. and Lakonishok, J. (1993) Can fundamentals predict Japanese stock returns?. *Financial Analysts Journal* 49: 63–69.
Chan, L.K.C., Hameed, A. and Tong, W. (2000) Profitability of momentum strategies in the international equity markets. *Journal of Financial and Quantitative Analysis* 35: 153–172.
Chan, L.K.C., Jegadeesh, N. and Lakonishok, J. (1995) Evaluating the performance of value versus glamour stocks: the impact of selection bias. *Journal of Financial Economics* 38: 269–296.
Chan, L.K.C., Jegadeesh, N. and Lakonishok, J. (1996) Momentum strategies. *Journal of Finance* 51: 1681–1713.
Chan, L.K.C., Karceski, J. and Lakonishok, J. (2003) The level and persistence of growth rates. *Journal of Finance* 58: 643–684.
Chan, L.K.C. and Lakonishok, J. (2004) Value and growth investing: review and update. *Financial Analysts Journal* 60: 71–86.
Chan, L.K.C., Yasushi H. and Lakonishok, J. (1991) Fundamentals and stock returns in Japan, *Journal of Finance* 46: 1739–1764.
Chan, R.W. and Lui, B.C. (2011) EV/EBIT ratio: the best of both worlds. *Better Investing* 60: 27–28.

© 2015 John Wiley & Sons Ltd

Charnes, A., Cooper, W.W. and Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429–444.
Chen, P. and Zhang, G. (2007) How do accounting variables explain stock price movements? Theory and evidence. *Journal of Accounting and Economics* 43: 219–244.
Cheng, C.S.A., Liu, C. and Schaefer, T. (1996) Earnings permanence and the incremental information content of cash flows from operations. *Journal of Accounting Research* 34: 173–181.
Chopra, N., Lakonishok, J. and Ritter J. (1992) Measuring abnormal performance: do stocks overreact? *Journal of Financial Economics* 31: 235–268.
Christie, W. (1990) Dividend yield and expected returns: the zero-dividend puzzle. *Journal of Financial Economics* 28: 95–125.
Chui, A.C.W., Titman, S. and Wei, K.C.J. (2010) Individualism and momentum around the world. *Journal of Finance* 65: 361–392.
Conrad, J. and Kaul, G. (1993) Long-term market overreaction or biases in computed returns? *Journal of Finance* 48: 39–63.
Conrad, J., Cooper, M. and Kaul, G. (2003) Value versus glamour. *Journal of Finance* 58: 1969–1995.
Cook T.J. and Rozeff, M.S. (1984) Size and earnings/price ratio anomalies: one effect or two? *Journal of Financial and Quantitative Analysis* 19: 449–466.
Cooper, I. (2006) Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance* 61: 139–170.
Damodaran, A. (2012) *Investment Valuation*. Hoboken, NJ: John Wiley & Sons.
Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (2001) Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance* 56: 921–965.
Daniel, K. and Titman, S. (1997) Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52: l-33.
Daniel, K. and Titman, S. (2006) Market reactions to tangible and intangible information. *Journal of Finance* 61: 1605–1643.
Davis, J. (1994) The cross-section of realized stock returns: the pre-compustat evidence. *Journal of Finance* 49: 1579–1593.
Davis, J. (2001) Is There Still Value in the Book-to-Market Ratio? Available at http://www.dfaus.com/2009/05/is-there-still-value-in-the-book-to-market-ratio.html accessed May 23, 2014.
Davis, J., Fama, E. and French, K. (2000) Characteristics, covariances and average returns: 1929 to 1997. *Journal of Finance* 55: 389–406.
DeBondt, W.F.M. and Thaler, R.H. (1985) Does the stock market overreact? *Journal of Finance* 40: 793–805.
Dechow, P.M., Hutton, A.P., and Sloan, R.G. (1999) An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics* 26: 1–34.
Denis, D.J., Denis, D.K. and Sarin, A. (1994) The information content of dividend changes: cash flow signaling, overinvestment and dividend clienteles. *Journal of Financial and Quantitative Analysis* 29: 567–587.
Desai, H., Rajgopal, S. and Venkatachalam, M. (2004) Value-glamour and accruals mispricing: one anomaly or two? *Accounting Review* 79: 355–385.
Dhatt, M.S., Kim, Y.H. and Mukherji, S. (1999) The value premium for small-capitalization stocks. *Financial Analysts Journal* 55: 60–68.
Dhatt, M.S., Kim, Y.H. and Mukherji, S. (2004) Can composite value measures enhance portfolio performance? *Journal of Investing* 13: 42–48.
Dichev, I. (1998) Is the risk of bankruptcy a systematic risk? *Journal of Finance* 53: 1131–1148.
Dielman, T.E. and Oppenheimer, H.R. (1984) An examination of investor behavior during periods of large dividend changes. *Journal of Financial and Quantitative Analysis* 19: 197–216.
Dimson, E., Nagel, S. and Quigley, G. (2003) Capturing the value premium in the United Kingdom. *Financial Analysts Journal* 59: 35–45.
Ding, D.K., Chua, J.L. and Fetherston, T.A. (2005) The performance of value and growth portfolios in east asia before the asian financial crisis. *Pacific-Basin Finance Journal* 13: 185–199.
Dissanaike, G. and Lim, K.H. (2010) The sophisticated and the simple: the profitability of contrarian strategies. *European Financial Management* 16: 229–255.

© 2015 John Wiley & Sons Ltd

Doeswijk, R.Q. (1997) Contrarian investment in the Dutch stock market. *De Economist* 145: 573–598.
Doukas, J.A., Kim, C.F. and Pantzalis, C. (2002) A test of the errors-in-expectations explanation of the value/glamour stock returns performance: evidence from analysts' forecasts. *Journal of Finance* 57: 2143–2165.
Doukas, J.A., Kim, C.F. and Pantzalis, C. (2004) Divergence of opinions and the performance of value stocks. *Financial Analysts Journal* 60: 55–64.
Doukas, J.A., Kim, C.F. and Pantzalis, C. (2005) The two faces of analyst coverage. *Financial Management* 34: 99–125.
Elton, E.J., Gruber, M.J. and Rentzler, J. (1983) A simple examination of the empirical relationship between dividend yields and deviations from the CAPM. *Journal of Banking and Finance* 7: 135–146.
Fama, E.F. (1970) Efficient capital markets: a review of theory and empirical work. *Journal of Finance* 25: 383–417.
Fama, E.F. (1991) Efficient capital markets: II. *Journal of Finance* 46: 1575–1617.
Fama, E.F. (1998) Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49: 283–306.
Fama, E.F. and French, K.R. (1992) The cross-section of expected stock returns. *Journal of Finance* 47: 427–465.
Fama, E.F. and French, K.R. (1993) Common risk factors in the returns on stock and bonds. *Journal of Financial Economics* 33: 3–56.
Fama, E.F. and French, K.R. (1995) Size and book to market factors in earnings and returns. *Journal of Finance* 50: 131–155.
Fama, E.F. and French, K.R. (1998) Value versus growth: the international evidence. *Journal of Finance* 53: 1935–1974.
Fama, E.F. and French, K.R. (2001) Disappearing dividends: changing firm characteristics or lower propensity to pay?. *Journal of Financial Economics* 60: 3–43.
Fama, E.F. and French, K.R. (2006) The value premium and the CAPM. *Journal of Finance* 61: 2163–2185.
Fama, E.F. and French, K.R. (2007a) Migration. *Financial Analysts Journal* 63: 48–58.
Fama, E.F. and French, K.R. (2007b) The anatomy of value and growth stock returns. *Financial Analysts Journal* 63: 44–54.
Fama, E.F. and French, K.R. (2012) Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105: 457–472.
Fama, E.F. and MacBeth, J.D. (1973) Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* 81: 607–636.
Fisher, K.L. (1984) *Super Stocks*. Homewood, IL: Dow-Jones-Irwin.
Fong, W.M. (2012) Do expected business conditions explain the value premium? *Journal of Financial Markets* 15: 181–206.
Fuller, R.J., Huberts, L.C. and Levinson, M.J. (1993) Returns to E/P strategies, higgeldy piggeldy growth, analysts' forecast errors, and omitted risk factors. *Journal of Portfolio Management* 19: 13–24.
García-Feijóo, L. and Jorgensen, R.D. (2010) Can operating leverage be the cause of the value premium? *Financial Management* 39: 1127–1154.
George, T.J. and Hwang, C.Y. (2010) A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics* 96: 56–79.
Gharghori, P., Stryjkowski, S. and Veeraraghavan, M. (2013) Value versus growth: australian evidence. *Accounting and Finance* 53: 393–417.
Goetzmann, W.N. and Jorion, P. (1993) Testing the predictive power of dividend yields. *Journal of Finance* 48: 663–679.
Gordon, M.J. and Shapiro, E. (1956) Capital equipment analysis: the required rate of profit. *Management Science* 3: 102–110.
Graham, B. (1973) *The Intelligent Investor*. New York, NY: Harper & Row Publishers.
Graham, B. and Dodd, D. (1934) *Security Analysis*. New York, NY: McGraw-Hill.
Gray, W.R. and Vogel, J. (2012) Analyzing valuation measures: a performance horse-race over the past 40 years. *Journal of Portfolio Management* 39: 112–121.

© 2015 John Wiley & Sons Ltd

Gregory, A., Harris, R.D.F. and Michou, M. (2001) An analysis of contrarian investment strategies in the UK. *Journal of Business Finance and Accounting*, 28: 1192–1228.
Griffin, J.M. and Lemmon, M.L. (2002) Book-to-market equity, distress risk, and stock returns. *Journal of Finance* 57: 2317–2336.
Guerard, J.B., Jr. (2006) Quantitative stock selection in Japan and the United States: some past and current issues. *Journal of Investing* 15: 43–49.
Guerard, J.B., Jr, and Takano, M. (1992) The development of mean-variance efficient portfolios in Japan and the U.S. *Journal of Investing* 1: 48–51.
Guerard, J.B., Jr, Xu, G. and Gültekin, M. (2012) Investing with momentum: the past, present, and future. *Journal of Investing* 21: 68–80.
Gulen, H., Xing, Y. and Zhang, L. (2011) Value versus growth: time-varying expected stock returns. *Financial Management* 40: 381–407.
Guthrie, G. (2013) A value premium without operating leverage. *Finance Research Letters* 10: 1–11.
Gutierrez, R.C. and Kelley, E.K. (2008) The long-lasting momentum in weekly returns. *Journal of Finance* 63: 415–447.
van der Hart, J., Slagter, E. and van Dijk, D. (2005) The success of stock selection strategies in emerging markets: is it risk or behavioral bias? *Emerging Markets Review* 6: 238–262.
van der Hart, J., deZwart, G. and van Dijk, D. (2003) Stock selection strategies in emerging markets. *Journal of Empirical Finance* 10: 105–132.
Haugen, R. and Baker, N. (1996) Commonality in the determinants of expected stock returns. *Journal of Financial Economics* 41: 401–439.
Hayn, C. (1995) The information contents of losses. *Journal of Accounting and Economics* 20: 125–153.
Hodrick, R.J. (1992) Dividend yields and expected stock returns: alternative procedures for inference and measurement. *Review of Financial Studies* 5: 357–386.
Hou, K., Karolyi, G.A. and Kho, B.C. (2011) What factors drive global stock returns? *Review of Financial Studies* 24: 2527–2574.
Hou, K., Xue, C. and Zhang, L. (2015) Digesting anomalies: an investment approach. *Review of Financial Studies* 28: 650–705.
Huang, I.H. (2011) The Cyclical Behavior of the Risk of Value Strategy: Evidence from Taiwan. *Pacific-Basin Finance Journal* 19: 404–419.
Hwang, S. and Rubesam, A. (2013) A behavioral explanation of the value anomaly based on time-varying return reversals. *Journal of Banking and Finance* 37: 2367–2377.
Israel, R. and Moskowitz, T.J. (2013) The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics* 108: 275–301.
Jaffe, J., Keim, D.B. and Westerfield, R. (1989) Earnings yield, market values and stock returns. *Journal of Finance* 44: 135–148.
Jegadeesh, N., Kim, J., Krische, S. and Lee, C. (2004) Analyzing the analysts: when do recommendations add value? *Journal of Finance* 59: 1083–1124.
Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48: 65–91.
Jegadeesh, N. and Titman, S. (2001) Profitability of momentum strategies: an evaluation of alternative explanations. *Journal of Finance* 56: 699–720.
Kallunki, J-P., Martikainen, M. and Martikainen, T. (1998) Accounting income, income components and market-to-book equity ratios: Finnish evidence. *International Journal of Accounting* 33: 359–375.
Keim, D.B. (1985) Dividend yields and stock returns: implications of abnormal january returns. *Journal of Financial Economics* 14: 473–489.
Keim, D.B. (1986) Dividend yields and the january effect. *Journal of Portfolio Management* 12: 54–60.
Kogan, L. (2004) Asset prices and real investment. *Journal of Financial Economics* 73: 411-431.
Korajczyk, R.A. and Sadka, R. (2004) Are momentum profits robust to trading costs? *Journal of Finance* 59: 1039–1082.
Kothari, S.P. and Shanken J., (1997) Book-to-market, dividend yield, and expected market returns: a time-series analysis. *Journal of Financial Economics* 44: 169–203.

© 2015 John Wiley & Sons Ltd

Kothari, S.P., Shanken J. and Sloan, R.G. (1995) Another look at the cross-section of expected stock returns. *Journal of Finance* 50: 185–224.
Kyriazis, D. and Diacogiannis, C. (2007) Testing the performance of value strategies in the athens stock exchange. *Applied Financial Economics* 17: 1511–1528.
Lakonishok, J., Vishny, R.W. and Shleifer, A. (1994) Contrarian investment, extrapolation and risk. *Journal of Finance* 49: 1541–1577.
Leivo, T.H. and Pätäri, E.J. (2011) Enhancement of value portfolio performance using momentum and the long-short strategy: the Finnish evidence. *Journal of Asset Management* 11: 401–416.
Leivo, T.H., Pätäri, E.J., and Kilpiä, I.J.J. (2009) Value enhancement using composite measures: the Finnish evidence. *International Research Journal of Finance and Economics* 33: 7–30.
Leledakis, G.N. and Davidson, I. (2001) Are two factors enough? The U.K. evidence. *Financial Analysts Journal* 57: 96–105.
Levis, M. (1989) Stock market anomalies: a re-assessment based on UK evidence. *Journal of Banking and Finance* 13: 675–696.
Li, X., Brooks, C. and Miffre, J. (2009) The value premium and time-varying volatility. *Journal of Business Finance and Accounting* 36: 1252–1272.
Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47: 13–37.
Lofthouse, S. (2001) *Investment Management*. Chichester, UK: John Wiley & Sons.
Loughran, T. (1997) Book-to-market across firm size, exchange, and seasonality: is there an effect? *Journal of Financial and Quantitative Analysis* 32: 249–268.
Loughran, T. and Wellman, J. (2011) New evidence on the relation between the enterprise multiple and average stock returns. *Journal of Financial and Quantitative Analysis* 46: 1629–1650.
Markowitz, H.M. and Xu, G.L. (1994) Data mining corrections. *Journal of Portfolio Management* 21: 60–69.
Martikainen, M. (1997) Accounting losses and earnings response coefficients: the impact of leverage and growth opportunities. *Journal of Business Finance and Accounting* 24: 277–291.
McLean, R.D. and Pontiff, J. (2014) Does Academic Research Destroy Stock Return Predictability?. Working Paper. Boston College, Boston, MA.
McManus, I.D., ap Gwilym, O. and Thomas, S.H. (2004) The role of payout ratio in the relationship between stock returns and dividend yield. *Journal of Business, Finance and Accounting* 31: 1355–1387.
McWilliams, J.D. (1966) Price, earnings and P-E ratios. *Financial Analyst Journal* 22: 137–142.
Miles, D. and Timmermann, A. (1996) Variations in expected stock returns: evidence on the pricing of equities from a cross-section of UK companies. *Economica* 63: 369–382.
Morgan, G. and Thomas, S. (1998) Taxes, dividend yields and returns in the UK equity market. *Journal of Banking and Finance* 22: 405–423.
Mossin, J. (1966) Equilibrium in a capital asset market. *Econometrica* 34: 768–783.
Mukherji, S., Dhatt, M. and Kim, Y. (1997) A fundamental analysis of Korean stock returns. *Financial Analysts Journal* 53: 75–80.
Naranjo, A., Nimalendran, M. and Ryngaert, M. (1998) Stock returns, dividend yields, and taxes. *Journal of Finance* 53: 2029–2057.
Nicholson, S.F. (1960) Price-earnings ratios. *Financial Analysts Journal* 16: 43–45.
Nicholson, S.F. (1968) Price ratios in relation to investment results. *Financial Analysts Journal* 24: 105–109.
Novy-Marx, R. (2011) Operating leverage. *Review of Finance* 15: 103–134.
Novy-Marx, R. (2013) The other side of value: the gross profitability premium. *Journal of Financial Economics* 108: 1–28.
Obreja, I. (2013) Book-to-market equity, financial leverage, and the cross-section of stock returns. *Review of Financial Studies* 26: 1146–1189.
Ohlson, J.A. (1980) Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18: 109–131.
Ohlson, J. (1995) Earnings, book values and dividends in equity valuation. *Contemporary Accounting Research* 11: 661–687.
Pätäri, E.J., Leivo, T.H. and Honkapuro, J.V.S. (2010) Enhancement of value portfolio performance using data envelopment analysis. *Studies in Economics and Finance* 27: 223–246.

© 2015 John Wiley & Sons Ltd

Pätäri, E.J., Leivo, T.H. and Honkapuro, J.V.S. (2012) Enhancement of equity portfolio performance using data envelopment analysis. *European Journal of Operational Research* 220: 786–797.
Pätäri, E.J., Karell, V. and Luukka, P. (2015) Can size-, industry-, and leverage-adjustment of valuation ratios benefit the value investor? *International Journal of Business Innovation and Research, Forthcoming*.
Pätäri, E.J. and Tolvanen, J. (2009) Chasing performance persistence of hedge funds – comparative analysis of evaluation techniques. *Journal of Derivatives and Hedge Funds* 15: 223–240.
Penman, S. (1996) The articulation of price-earnings and market-to-book ratios and the evaluation of growth. *Journal of Accounting Research* 34: 235–259.
Penman, S.H. (2013) *Financial Statement Analysis and Security Valuation*. Singapore: McGraw-Hill.
Penman, S. and Reggiani, F. (2013) Returns to buying earnings and book value: accounting for growth and risk. *Review of Accounting Studies* 18: 1021–1049.
Penman, S., Richardson, S.A. and Tuna, I. (2007) The book-to-price effect in stock returns: accounting for leverage. *Journal of Accounting Research* 45: 427–467.
Petkova, R. and Zhang, L. (2005) Is value riskier than growth? *Journal of Financial Economics* 78: 187–202.
Phalippou, L. (2008) Where is the value premium? *Financial Analysts Journal* 64: 41–48.
Piotroski, J. (2000) Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research* 38: 1–41.
Piotroski, J. and So, E.C. (2012) Identifying expectation errors in value/glamour strategies: a fundamental analysis approach. *Review of Financial Studies* 25: 2841–2875.
Reinganum, M.R. (1981) The Misspecification of capital asset pricing empirical anomalies based on earnings yield and market values. *Journal Financial Economics* 9: 19–46.
Roll, R. (1995) Style return differentials: illusions, risk premiums, or investment opportunities. In Coggin, T.D. and Fabozzi, F.J. (eds) *The Handbook of Equity Style Investment*. New Hope, USA: Frank J. Fabozzi Associates.
Rosenberg, B., Reid, K. and Lanstein, R. (1985) Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11: 9–16.
Rouwenhorst, K.G. (1998) International momentum strategies. *Journal of Finance* 53: 267–284.
Rouwenhorst, K.G. (1999). Local return factors and turnover in emerging stock markets. *Journal of Finance* 54: 1439–1464
Sant, R. and Cowan, A.R. (1994) Do dividends signal earnings? The case of omitted dividends. *Journal of Banking and Finance* 18: 1113–1133.
Senchack, A.J., Jr. and Martin, J.D. (1987) The relative performance of the PSR and PER investment strategies. *Financial Analysts Journal* 43: 46–56.
Sharpe, W.F. (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425–442.
Sharpe, W.F. (1966) Mutual fund performance. *Journal of Business* 39: 119–138.
Sortino, F. A. and van der Meer, R. (1991) Downside risk. *Journal of Portfolio Management*, 17: 27–31.
Stattman, D. (1980) Book values and stock returns. *The Chicago MBA: A Journal of Selected Papers* 4: 25–45.
Suzuki, M. (1998) PSR—an efficient stock-selection tool? *International Journal of Forecasting* 14: 245–254.
Vassalou, M. and Xing, Y. (2004) Default risk in equity returns. *Journal of Finance* 59: 831–868.
Wilson, G.P. (1986) The Relative Information content of accruals and cash flows: combined evidence at the earnings announcement and annual report release date. *Journal of Accounting Research* 24: 165–200.
Yen, J.Y., Sun, Q. and Yan, Y. (2004) Value versus growth stocks in Singapore. *Journal of Multinational Financial Management* 14: 19–34.
Zhang, L. (2005) The value premium. *Journal of Finance* 60: 67–103.

© 2015 John Wiley & Sons Ltd