Journal of Banking and Finance 114 (2020) 105785

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Does earnings growth drive the quality premium?[☆]

Georgi Kyosev[a,c,*], Matthias X. Hanauer[b,c], Joop Huij[a,c], Simon Lansdorp[c]

[a] Rotterdam School of Management, Erasmus University, the Netherland
[b] TUM (Technische Universität München), Germany
[c] Robeco, the Netherlands

ARTICLE INFO

Article history:
Received 16 May 2018
Accepted 21 February 2020
Available online 25 February 2020

JEL Classification:
C12
G11
G12
G14

Keywords:
Quality
Factor premiums
Earnings growth
Return-on-equity
Profit margins
Leverage
Earnings variability
Operating accruals
Investments
Gross profitability

ABSTRACT

High (low) quality stocks generate anomalously high (low) returns above and beyond expected returns based on betas, market sizes, valuations, and momentum. We provide a comprehensive overview of commonly used quality definitions and test their predictive power for stock returns. We show that quality measures predict stock returns if and only if they forecast earnings growth, and that this information is not contained in other characteristics that have been shown to drive expected stock returns. At the same time, we find that the quality premium is unrelated to different measures of distress risk, and therefore inconsistent with a risk-based interpretation. Finally, our results are robust across different regions and carry over to the corporate bond market.

1. Introduction

Size, value and momentum factors have been dominating the empirical asset pricing literature over the past few decades.[1] More recently, however, a range of accounting-based factors (see, e.g., Haugen and Baker, 1996; Sloan, 1996; Cooper et al., 2008; Novy-Marx, 2013) have been documented to have predictive power for future stock returns above and beyond their expected returns based on betas, market sizes, valuations, and momentum.[2] Because these accounting-based variables are often seen as important determinants for investors' perception of firm quality (see, e.g., McGuire et al., 1990; Asness et al., 2019; Trammell, 2014), they are also referred to as quality variables.[3]

A notable observation regarding these accounting-based (quality) factors is the lack of a common element (apart from all of them being derived from accounting statements). While different definitions are also used to measure value (e.g., book-to-price and earnings-to-price), momentum (e.g., 6-minus-1-month return and 12-minus-1-month return), and low-risk (e.g., 36-month volatility and 52-week market beta), the dispersion in definitions is substantially larger for quality. Examples of variables that are seen as quality indicators are (derivations of) return-on-equity (Haugen and Baker, 1996), low accruals (Sloan, 1996), low investments (Cooper et al., 2008), low leverage (George and Hwang, 2010), and gross profitability (Novy-Marx, 2013). While there seems to be a consensus that quality measures have predictive power for the cross-section of future stock returns, there

☆ The views expressed in this paper are solely those of the authors and not necessarily shared by Robeco or its subsidiaries. The authors thank Jeroen van Zundert, David Blitz, Viorel Roscovan, and Milan Vidojevic for valuable comments.

* Corresponding author.
E-mail addresses: kyosev@rsm.nl (G. Kyosev), matthias.hanauer@tum.de (M.X. Hanauer), jhuij@rsm.nl (J. Huij), s.lansdorp@robeco.com (S. Lansdorp).

[1] Cf. Basu (1977) for value, Banz (1981) for size, and Jegadeesh and Titman (1993) for momentum.

[2] Two of these factors, namely profitability and investments, are also included in the q-factor model of Hou et al. (2015) and the five-factor model of Fama and French (2015).

[3] Throughout the paper, we use the terms "quality" and "accounting-based" interchangeably.

https://doi.org/10.1016/j.jbankfin.2020.105785
0378-4266/

is no study which explains what drives the return differences between the measures; why some quality measures systematically work better than others; why quality measures predict future stock returns in the first place; and if the quality premium can be observed in other markets than the U.S. stock market as well.

In this study we intend to fill these gaps in the literature. To do so we employ the adjusted dividend discount model of Fama and French (2015) as a starting point. Fama and French (2015) derive a theoretical relation between expected stock returns, profitability, and valuation based on a rewritten dividend-discount model as in Eq. (1).

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}}{B_t} \tag{1}$$

We investigate the relation between quality measures and future profitability ($Y_{t+\tau}$ in the numerator of Eq. (1)) and expected returns (discount rate r in the denominator of Eq. (1)) by testing the predictive power of quality measures for future one-, three-, and five-year earnings growth. Although some authors argue that quality-related variables predict stock returns because they measure true economic profitability and have predictive power for future earnings (see, e.g., Sloan, 1996; Novy-Marx, 2013; Akbas et al., 2017), we are not aware of any empirical evidence that supports this notion and rejects a risk-based interpretation. While Sloan (1996), Novy-Marx (2013), and Akbas et al. (2017) show that quality measured by accruals, gross profitability, or trend in gross profitability, respectively, predict future earnings growth and stock returns, they do not provide evidence supporting a causal relation nor do they test for a causal relation between the quality premium and distress risk.

The main contribution of our study is that we provide empirical evidence that the predictive power of quality variables for the cross-section of stock returns originates from the variables being good proxies for future earnings growth. We find that quality measures can predict future stock returns if and only if they are good proxies for future earnings growth. Quality variables which have no predictive power for future earnings growth also have no predictive power for the cross-section of future stock returns. Hence, the predictive power of quality measures for stock returns can be attributed to their predictive power for future earnings growth. We also show that there is no causal relation between the quality premium and distress risk. As such this study contributes to the literature that provides empirical evidence supporting the conventional wisdom that quality is a measure of true economic profitability (Akbas et al., 2017) and that mispricing is a more likely cause for the existence of the quality premium than distress risk is.[4]

Another contribution of our study is that we analyze the robustness of the predictive power of accounting-based factors in an international and multi-asset setting. Existing studies investigating quality factors have mainly been performed using broad U.S. equity data. We find robust results for the predictive power of quality measures for future stock returns across the U.S., Europe, Japan, global developed markets, and emerging markets. For additional robustness, we expand our analyses to the corporate bond markets and find consistent results – bonds issued by high-quality companies outperform those issued by low-quality companies if the quality measures used are good proxies for future earnings growth.

The remainder of the paper is organized as follows: Section 2 describes the data, quality definitions, and methodology. Section 3 presents our empirical results, Section 4 applies robustness tests, and Section 5 concludes.

[4] Our results are also consistent with the recent findings of Franke et al. (2017) that challenge a risk-based explanation for the profitability and investment factors.

2. Quality definitions, methodology, and data

In this section, we describe the quality variable definitions, methodology, and data used throughout this paper.

2.1. Quality definitions

In contrast to size, value, momentum, or low-risk factors, accounting-based factors show a considerable dispersion in definitions.[5] Therefore, this section provides an overview of selected quality definitions applied throughout the paper and motivates our variable choices.

Following the documentation of size, value, and momentum patterns in average stock returns, the Fama and French (1993) three-factor and Carhart (1997) four-factor models have been the "industry standard" (Subrahmanyam, 2010, p. 35) in the empirical asset pricing literature for many years. However, Sloan (1996) already shows that accruals are negatively related to future earnings and that higher accruals predict lower stock returns. Sloan (1996) posits that investors fixate too much on book earnings and do not understand that accruals are less persistent than cash flows. As accruals tend to reverse in following years, investors are negatively (positively) surprised by lower (higher) than expected future earnings and stock returns for firms with high (low) accruals. Haugen and Baker (1996) assume that currently profitable firms have greater potential for future growth and document that ROE is positively correlated with future returns. Cooper et al. (2008) find that a general measure of investments (total asset growth) is negatively correlated with future returns.[6] Furthermore, the authors document that returns around earnings announcements are positive (negative) for low (high)-growth firms, consistent with the errors-in-expectations hypothesis in La Porta (1996).

While Fama and French (2015) also use asset growth as a proxy for investments in their five-factor model, the findings of Novy-Marx (2013) could explain why they do not use ROE as a proxy for future profitability. Novy-Marx (2013) finds that gross profitability as a top-line profitability measure is superior in predicting future stock returns to bottom-line earnings. The author argues that gross profitability performs better than ROE because it is the better proxy for true economic (expected) profitability and supports his hypothesis by showing that gross profitability has power to predict long term growth profitability.

As the average high book-to-market firm is financially distressed, Piotroski (2000) also investigates financial performance signals such as margins and profitability growth to discriminate between stocks with strong and weak prospects. Therefore, we also include the twelve months growth in return on equity (ROE growth) and earnings to sales (margins) to our list of quality variables. Finally, other studies show that proxies for the safety of a company such as debt to common equity (leverage, George and Hwang, 2010) and volatility of earnings growth (earnings variability, cf. Huang, 2009), are negatively correlated with future stock returns.

Based on the overview provided above, we select the following quality definitions for our main analysis: ROE, margins, ROE

[5] Value strategies have generally in common that they invest in stocks with a low price to their fundamentals, such as book value of equity, earnings, or dividends, while momentum strategies usually buy stocks that have had high returns over the past three to twelve months. Low-risk strategies typically invest in stocks with low beta or low volatility estimated over certain time periods and frequencies.

[6] Other studies also document a negative correlation of components of firm's investment or financing activities. See, for example, Fairfield et al. (2003, growth in net operating assets), Richardson and Sloan (2003, external financing), Titman et al. (2004, abnormal capital investments), and Pontiff and Woodgate (2008, net share issuance).

growth, leverage, earnings variability, gross profitability, accruals, and investments (total asset growth). We aim to select variables that are seen of particular importance by key academic studies or practitioners and are rather pure and at the extremes of the potential definition spectrum. E.g., gross profitability is based on the highest income statement item gross profits, while ROE is based on bottom-line net income. Similarly, accruals can be seen as the difference between operating book profits and operating cash flow. Nevertheless, we understand that our choice can be seen as somewhat arbitrary and that also other variables or combinations of existing variables could be considered.[7] We therefore have added the empirical analysis for variables that are rather derivations or combinations of the variables mentioned above in Section 2.1. The detailed variable definitions can be found in the Appendix. For a comprehensive review of the accounting anomalies literature, we refer interested readers to Richardson et al. (2010).

Most of the mentioned studies above link the predictive power of the discussed variables to their earnings predictive power either by directly predicting future earnings (e.g., Sloan, 1996 or Novy-Marx, 2013) or by investigating returns around earnings announcements (e.g., Cooper, 2008 or Piotroski, 2000). As quarterly earnings announcement is (historically) not common practice in international markets, we choose to follow the former option. The next subsection explains our methodology in more detail.

2.2. Methodology

In this paper, we use two commonly accepted approaches to explain both future earnings growth and stock returns: (i) cross-sectional Fama and MacBeth (1973) regressions, and (ii) sorting stocks into portfolios based on quality variables.

To measure the predictive power of quality variables for future earnings growth, we follow Novy-Marx (2013) and use the cross-sectional regression approach of Fama and MacBeth (1973). More specifically, we run quarterly regressions on the one-, three-, and five-year change in earnings growth scaled by book equity on individual quality characteristics. First, we conduct univariate regressions to estimate the direct relation for each quality variable on future earnings growth. Furthermore, to estimate marginal effects, we include all quality variables and the control variables capturing common risk and return factors, namely market beta, volatility, distress risk, size (market cap), book-to-price, and momentum. Furthermore, we add region and sector dummies. This analysis determines which quality variables are statistically distinct and which have no marginal power to explain future earnings growth. All independent variables (firm characteristics) are winsorized at the top and bottom percentiles and t-statistics are Newey-West adjusted using four lags. Specifically, we run the following equations:

$$\left(\frac{Earnings_{i,\ t+r} - \ Earnings_{i,t}}{BE_{i,t}}\right) = \ a_{i,t} + b_{i,t}Q_{i,t} + \varepsilon_{i,t} \tag{2}$$

$$\left(\frac{Earnings_{i,t+r} - \ Earnings_{i,t}}{BE_{i,\ t}}\right) = \alpha_{i,t} + b_{1,\ i,t}Q_{1,i,t} + \cdots + b_{N,\ i,t}Q_{N,\ i,t} + c_{\ i,t}\theta_{i,t} + s_{i,t}RS_{i,t} + \varepsilon_{i,t} \tag{3}$$

Where, $\left(\frac{Earnings_{i,\ t+r} - \ Earnings_{i,t}}{BE_{i,t}}\right)$ is the change in earnings over the next 12, 36, or 60 months for stock i, in month t, $Q_{i,\ t}$ is the relevant quality measure for stock i in month t, and $\theta_{\ i,\ t}$ is a set of control variables represented by three-year market beta, three-year volatility, distance to default, natural logarithm of market capitalization, natural logarithm of book-to-price, and 12-1 month momentum, and $RS_{\ i,\ t}$ is region, sector dummies.

We also conduct monthly Fama and MacBeth (1973) regressions to answer which quality variables have explanatory power for the cross-section of stock returns in the following way

$$R_{i,t+1} = \ \alpha_{i,t} + b_{\ i,t}Q_{\ i,t} + \ c_{\ i,t}\theta_{\ i,t} + s_{\ i,t}RS_{\ i,t} + \ \varepsilon_{i,t} \tag{4}$$

Besides the standard regression of next month returns on lagged characteristics, we also predict three-year ahead returns as the dividend discount model in Eq. (1) makes a statement on long-term rather than short-term returns.

$$R_{i,t+1:\ t+36} = \ \alpha_{i,t} + b_{\ i,t}Q_{\ i,t} + \ c_{\ i,t}\theta_{\ i,t} + s_{\ i,t}RS_{\ i,t} + \ \varepsilon_{i,t} \tag{5}$$

Finally, we investigate whether a causal relationship exists between priced quality measures from the preceding regression and future earnings growth. Therefore, we also control for the *realized* future three-year growth in earnings (not known ex-ante). If some quality measures are only priced because they are a good proxy for future profitability, one would expect that they become unpriced once controlled for future profitability. All firm characteristics are standardized into z-scores and t-statistics are Newey-West adjusted using three and 35 lags for regressions on monthly and three-year returns, respectively.

$$R_{i,\ t+1:\ t+36} = \alpha_{i,t} + b_{\ i,t}Q_{\ i,t} + \ c_{\ i,t}\theta_{\ i,t} + s_{\ i,t}RS_{\ i,t} + e_{\ i,t}E_{\ i,t+36} + \ \varepsilon_{i,t} \tag{6}$$

Where, $R_{i,t+1}$ is the return in the following month of stock i, in month $t+1$, $R_{i,t+1:t+36}$ is the return over the next 36 months, $Q_{i,\ t}$ is the relevant quality measure for stock i, in month t, and $\theta_{\ i,\ t}$ is a set of control variables represented by three-year market beta, three-year volatility, distance to default, natural logarithm of market capitalization, natural logarithm of book-to-price, and 12-1 month momentum, $RS_{\ i,\ t}$ is region, sector dummies, and $E_{\ i,t+36}$ is the change in earnings over the next 36 months measured by $\left(\frac{Earnings_{i,\ t+36} - \ Earnings_{i,t}}{BE_{i,t}}\right)$. In all cross-sectional regressions, all quality variables are multiplied by the expected sign, so the expected coefficient is positive.

Finally, we construct equally-weighted quintile portfolios by ranking stocks on all the variables described above. For accruals, investments, earnings variability, and leverage measures, stocks with the lowest values are assigned to the top quintile, while for the remaining variables stocks with the highest factor scores are the top quintiles. We also form two composites of quality measures ('Earnings non-predictive' and 'Earnings predictive') based on the outcome of the earnings prediction regressions in Section 3. The composites are constructed based on an equally-weighted combination of all individual variables' z-scores. For all variable sorts, factor scores are compared directly across all stocks, without imposing sector or country neutralities. We do control for regional effects in our global developed markets sample by presenting results for the U.S., Europe, and Japan in isolation. Portfolios are rebalanced monthly.

For the top, bottom, and top-minus-bottom (T-B) quintile portfolios, we report the annualized average returns (in USD and in excess of the risk-free rate), volatilities and Sharpe ratios. Furthermore, we estimate the Fama and French – Carhart 4-factor alphas and coefficients for the T-B portfolios by running the following regression:

$$R_{T-B,t} = \alpha_{T-B} + \beta_{T-B}\left(R_{M,t} - R_{f,t}\right) + s_{T-B}SMB_t + h_{T-B}HML_t + w_{T-B}WML_t + \varepsilon_{T-B,t} \tag{7}$$

where $R_{T-B,t}$ is the difference between the top and bottom portfolio returns in period t, $R_{f,\ t}$ is the risk-free return in period t,

[7] While we admit that there is still an ongoing discussion on whether these proxies can be further improved, the definitions used in this paper are usually the ones initially documented in the literature and therefore represent a conservative choice. Studies that further refine and/or improve variable definition are, for example, Thomas and Zhang (2002) and Richardson et al. (2005) for accruals, Pontiff and Woodgate (2008) for investments, Ball et al. (2015, 2016), Fama and French (2015), and Hanauer and Huber (2019) for profitability, or Akbas et al. (2017) for profitability growth.

α_{T-B} is the alpha of the top minus bottom portfolio, $R_{M,\,t}$ is the return of the market portfolio in period t, and β_{T-B}, s_{T-B}, h_{T-B}, and w_{T-B} are the estimated factor coefficients. Global and regional size (small-minus-big, *SMB*), value (high-minus-low, *HML*) and momentum (winner-minus-loser, *WML*) factors are calculated by ranking stocks, on their market capitalization, book-to-market ratio and past 12-minus-1 month local total return respectively, and taking the difference in return between the equally-weighted top and bottom terciles.

A consistent rank portfolio approach is used for our corporate bond analysis – we form equally-weighted quintile portfolios. Due to the systematically lower liquidity of corporate bonds compared to equities, we substitute the one-month holding period used for equities with a twelve months holding period. To do so, we use the overlapping portfolio approach of Jegadeesh and Titman (1993). We split the corporate bond universe into investment grade and high yield as they are effectively seen as two different asset classes by practitioners and academics (e.g., Ambastha et al., 2010).

2.3. Data

Our sample comprises developed and emerging market stocks starting from December 1985 and December 1992, respectively, until December 2015. At the end of every month, we identify all constituents of the FTSE World Developed Index and the S&P/IFC Investable Emerging Markets Index for that particular month. We exclude financial firms as they are subject to special accounting standards and do not exhibit comparable values for some of the variables we investigate. The resulting global developed large- and mid-cap universe consists of approximately 1,600 stocks on average; the actual number ranges between about 1,200 and 1,900 over time.[8] As many return anomalies are known to disappear or become significantly less pronounced when the universe is restricted to large- and mid-caps, our choice of universe is rather conservative.[9] For emerging markets, we make a similar conservative choice by restricting our sample each month to the 500 biggest stocks as measured by market capitalization in USD.[10]

We gather monthly stock returns taking into account dividends, stock splits and other capital adjustments. Our first data source for returns and outstanding shares is Interactive Data Exshare. In case this data is not available, we use MSCI return series instead. Alternatively, when neither of these is available, we calculate total returns using daily data from S&P/IFC. Daily returns above 500% are truncated at this level. In addition to returns, we gather free-float adjusted market capitalization data from FTSE and S&P/IFC and fundamental data from Compustat and Worldscope. As a proxy for the risk-free rate, we obtain the 1-month U.S. Treasury bill rate from the data library of Kenneth French. Table 1 reports descriptive statistics of our Global Markets sample.

Our corporate bond dataset is based on the Barclays U.S. Corporate Investment Grade index and U.S. Corporate High Yield index during the period January 1994 until December 2015. Bond returns are provided by Barclays and accounting data is downloaded from Compustat and Worldscope. We only include bonds for companies with publicly traded equity due to the availability of accounting information. In the case of multiple bonds outstanding we include only one as we prefer 1) senior bonds over subordinated ones, 2) bonds in the maturity segment of 5-15 years, 3) younger bonds, and 4) larger bonds. Our final sample consists of 403 investment grade bonds and 407 high yield bonds. We base our corporate bond analysis on returns in excess over duration matched treasuries as provided by Barclays. This allows us to focus on the default premium component of corporate bond returns and ignore the term premium, which can be gained by investing in government bonds.

Clean surplus accounting is neccesary to rewrite the dividend-discount-model into a model as represented in Eq. (1).[11] The majority of countries in our sample adopted IFRS (see, e.g., Hung et al., 2015) or their domestic accounting standards support already clean surplus accounting, like U.S. GAAP or U.K. GAAP. For the remaining countries we rely on Worldscope's data preparation that "take[s] into consideration the variety of accounting conventions and are designed to facilitate comparisons between companies and industries within and across national boundaries" (see Thomson Reuters, 2013, p. 27).

3. Empirical results

In this section, we conduct a set of empirical tests to shed more light on the common quality indicators. First, we test which of the widely used quality measures are forward-looking indicators for firm profitability, i.e., which ones have predictive power for future earnings growth. Second, we compare the performance of hypothetical global investment strategies based on the same set of quality definitions. Third, we create two competing quality strategies – one using the earnings predictive quality variables and one using the earnings non-predictive quality variables – and compare their performance across multiple settings. Finally, we test if there is a causal relation between the quality premium and distress risk. In addition, we conduct regional analyses to test if our global results are a manifestation of systematic regional allocation bets. For further robustness, we extend our analyses to emerging markets and corporate bonds.

3.1. Quality and growth in future profitability

A common feature of the quality variables is that they use accounting information measuring backwards-looking firm productivity. In the spirit of the dividend-discount model as in Fama and French (2015), however, expected future profitability is crucial and good quality variables should capture the true productivity of a company. A common indicator that financial analysts look at is innovations in earnings. This overlaps with the definition of Sloan (1996) and Novy-Marx (2013) that quality is a measure of true economic profitability and that it has strong predictive power for future earnings.

In Table 2 we show results of Fama-MacBeth (1973) regressions of one-, three-, and five-year growth in earnings on individual quality characteristics. In the text, we focus on the three-year change results. The results for the one and five year periods are, however, qualitatively similar.

[8] The countries included in the FTSE World Developed Index are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States.

[9] Existing academic studies investigating the quality-type factors have mainly been performed using broad U.S. equity data which can be dominated by micro-caps. E.g., Fama and French (2008) highlight that micro caps comprise on average only about 3% of the aggregated market cap of the NYSE-Amex-NASDAQ universe, but account for about 60% of the total number of stocks.

[10] The countries included in the S&P/IFC Investable Emerging Markets Index are Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, Israel, Jordan, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Portugal, Russia, South Africa, South Korea, Sri Lanka, Taiwan, Thailand, Turkey, and Venezuela. Greece, Israel, and Portugal are countries that were classified as both "emerging" and "developed" in different years of our sample period.

[11] We thank an anonymous reviewer for this comment.

Table 1
Descriptive statistics
The table reports descriptive statistics of our Global Markets sample. Characteristics are calculated according to the Appendix and winsorized at 1% level. All characteristics are cross sectionally calculated each month and the mean value over all months is reported. Panel A shows our independent and control variables, Panel B shows our dependent variables, Panel C shows the average number of stocks of the Global sample as well as of the regional subsamples – US, Europe, Japan – and Emerging markets (EM). Earnings variability, volatility 3Y, and total returns are measured in percent. Market capitalization (Mcap) is measured in million U.S. Dollars.

	mean	st. deviation	median	25th percentile	75th percentile
Panel A: Independent variables					
ROE	0.12	0.14	0.11	0.05	0.17
Margins	0.17	0.26	0.12	0.05	0.26
ROE growth	−0.01	0.12	0.00	−0.03	0.03
Leverage	0.73	0.90	0.47	0.18	0.93
Earnings Variability	1.99	4.50	0.50	0.21	1.68
Gross profitability	0.28	0.18	0.24	0.15	0.38
Accruals	−0.04	0.05	−0.04	−0.07	−0.01
Investments	1.10	0.19	1.06	1.00	1.14
Mcap	7 521.81	12 129.32	3 417.50	1 460.64	7 801.82
Beta 3Y	0.94	0.58	0.88	0.51	1.29
Volatility 3Y	31.87	10.60	30.02	24.23	37.46
Distance-to-default	4.83	1.69	4.57	3.59	5.78
Book To Price	0.56	0.37	0.47	0.29	0.73
Momentum 12-1	0.11	0.28	0.08	−0.09	0.26
Panel B: Dependent variables					
Δ Earnings 12M	0.01	0.12	0.01	−0.02	0.04
Δ Earnings 36M	0.03	0.18	0.02	−0.04	0.09
Δ Earnings 60M	0.05	0.23	0.03	−0.04	0.13
Total return	0.92	7.85	0.67	−4.18	5.76

Panel C: Sample size	Global	US	Europe	Japan	EM
Number of Stocks	1521	468	469	380	491

The column ‘3Y change univariate’ shows the average univariate regression coefficients. Consistent with the studies of Sloan (1996) and Novy-Marx (2013), high gross profitability, low accruals, and low investments positively predict future earnings growth, corrected for the expected sign, with coefficients 2.48 (t-stat 2.06), 34.59 (t-stat 8.25), and 5.85 (t-stat 3.88) respectively. On the other hand, high ROE, high margins, high ROE growth, low leverage and low earnings variability are associated with a negative change in future earnings, corrected for the expected sign. This is an indication that profitability measures based on earnings tend to mean revert. Therefore, investors who want to capture future profitability should discount past earnings information when making inferences for true firm profitability.

In a univariate Fama-MacBeth setting, earnings-based measures and leverage all have significant and negative predictive power for one, three, and five years earnings growth. This means that quality strategies based on these measures will suffer from negative profitability changes. Looking back at the dividend discount model predictions in Eq. (1), negative expected earnings mean that higher expected returns do not immediately stem from the theoretical model. On the other hand, high gross profitability, low accruals, and low investments, all scaled by assets, positively predict earnings growth at all horizons meaning that all else equal, they should have higher expected returns.

In a multiple regression framework, we test the marginal predictive power of our set of quality measures after controlling for other firm characteristics, risk measures, and regional and sector effects, and results remain qualitatively similar. Gross profitability, accruals, and investments correctly predict earnings growth across all horizons. It is also important that they remain significant when included simultaneously in the regression, meaning that they contain different information about future profitability. ROE, margins, ROE growth, leverage, and earnings variability either predict earnings growth with an opposite to the expected sign or have no predictive power.

3.2. Quality and stock returns

In this section, we look at the discount rate side of the dividend discount model. We test whether quality measures which can predict earnings growth also predict returns, as predicted by Eq. (1). Furthermore, we group quality measures into earnings predictive (EP) and earnings-non-predictive (ENP) based on their significance in Table 2. We measure significance with a one-sided t-test of whether the coefficient is significant at the hypothesized direction of earnings prediction for each quality variable.

Table 3 shows cross-sectional regression results of short-term and long-term returns on our set of quality indicators, controlling for firm size, valuation, past returns, market beta, volatility, distress risk, and region as well as sector dummies. Panel A contains standard Fama-MacBeth analysis of next-month returns on lagged characteristics. If the dividend discount model predictions hold, only the earnings predictive characteristics should have significant coefficients. Our results confirm this theoretical prediction as the earnings predictive variables gross profitability, accruals, and investments also have predictive power for stock returns with coefficients of 0.13 (t-stat 4.49), 0.06 (t-stat 4.52), and 0.06 (t-stat 2.99), respectively. The earnings-non-predictive characteristics, except for ROE, have coefficients which are not statistically distinguishable from zero.

As the dividend discount model refers rather to long-term expected returns than short-term returns, we also investigate the predictive power of quality indicators for horizons. In Panel B we show Fama-MacBeth regressions of three-year stock returns on the

Table 2
Predictive power of quality measures for one, three, and five years future earnings growth
The table reports results of Fama-MacBeth (1973) regressions of future one, three, and five year growth in earnings scaled by book equity ($\frac{Earnings_{t+\tau} - Earnings_t}{BE_t}$) on individual firm characteristics. Characteristics are calculated according to the Appendix and winsorized at 1% level. t-statistics are Newey-West adjusted using four lags. Regressions are run on quarterly data during the period January 1986 - December 2010 for our Global markets sample. The column 'univariate' shows the average univariate regression coefficient for the respective quality measure. The column 'multiple' shows marginal predictive power of the quality measures, controlling for other quality, firm characteristics and region, sector dummies. All quality variables are multiplied by the expected sign of the coefficient ("+" for ROE, Margins, ROE Growth, Gross profitability, ENP, and EP, and "-" for Leverage, Earnings variability, Accruals, and Investments). After multiplying the quality measures with the expected sign, we conduct a one-sided t-test for significance. *** denote one-sided significance at the 1% level, ** - at the 5% level, and * - at the 10 % level

	1Y change		3Y change		5Y change	
	univariate	multiple	univariate	multiple	univariate	multiple
ROE	−25.44	−38.31	−38.77	−54.73	−36.39	−57.42
t-stat	[−10.44]	[−16.08]	[−9.05]	[−17.28]	[−8.20]	[−16.09]
Margins	−9.68	−2.85	−14.13	−1.50	−11.83	0.05
t-stat	[−10.75]	[−5.12]	[−10.35]	[−1.27]	[−6.86]	[0.03]
ROE growth	−26.16	−12.77	−37.46	−12.88	−46.98	−21.15
t-stat	[−11.64]	[−9.98]	[−10.54]	[−6.67]	[−13.89]	[−8.22]
Leverage	−1.17	−0.16	−2.51	−0.83	−3.05	−1.53
t-stat	[−5.28]	[−0.73]	[−6.34]	[−2.37]	[−6.45]	[−4.21]
Earnings variability	−0.28	−0.06	−0.46	−0.04	−0.40	−0.02
t-stat	[−5.41]	[−3.98]	[−5.12]	[−1.21]	[−3.92]	[−0.50]
Gross profitability	1.38***	2.17***	2.48***	5.14***	5.52***	5.43***
t-stat	[2.18]	[3.33]	[2.06]	[4.64]	[3.94]	[3.38]
Accruals	19.81***	2.04**	34.59***	8.30***	38.88***	11.11***
t-stat	[6.84]	[1.74]	[8.25]	[4.10]	[7.81]	[3.75]
Investments	10.18***	1.31***	5.85***	2.59***	4.46***	1.59*
t-stat	[4.53]	[2.38]	[3.88]	[3.16]	[2.43]	[1.37]
ln(mcap)		−0.04		−0.08		−0.26
t-stat		[−0.49]		[−0.76]		[−1.32]
Beta 3Y		0.53		−0.04		−0.09
t-stat		[1.59]		[−0.08]		[−0.20]
Volatility 3Y		−0.07		−0.03		−0.03
t-stat		[−4.80]		[−1.29]		[−1.29]
Distance-to-Default		0.21		0.24		−0.05
t-stat		[2.96]		[2.02]		[−0.53]
ln(Book-to-price)		−7.44		−9.74		−11.33
t-stat		[−15.95]		[−17.85]		[−12.33]
Momentum 12-1		8.28		7.47		4.97
t-stat		[15.55]		[7.74]		[4.56]
R-squared (%)		0.32		0.29		0.24

same set of quality characteristics and control variables. The coefficient on gross profitability, accruals, and investments remain positive (t-stats of 1.46, 2.95, and 2.74, respectively), while the remaining quality variables are still insignificant.

Since the dividend discount model states that, all else equal (e.g., book-to-market), higher future earnings imply higher expected stock returns, the effects described above should be explained after controlling for the relevant information. Therefore, in Panel C we augment our regression specification and regress three-year returns on quality characteristics, controlling for three-year growth in earnings and the same control variables as in Panels A and B. This adjustment makes the results exactly the opposite to the ones in Panels A and B. Accruals and investments become insignificant while only gross profitability keeps its sign and becomes significant with a t-statistic of 2.91. On the other hand, all other characteristics with the exception of leverage – ROE, margins, past ROE growth, and earnings variability – become significant after controlling for the negative earnings growth associated with them. Finally, the coefficient on change in earnings is highly significant in all regressions with t-statistics around 18. These results have important implications for causality of the relationship between quality indicators and stock returns. They show that what is driving returns is future earnings growth and different measures used to define quality are effectively different ways to predict earnings growth. It also shows that earnings information is highly relevant as all earnings-based characteristics are significantly related to stock returns after controlling for earnings mean reversion associated with them. All in all, our results indicate that a true quality definition should include measures that positively predict earnings growth and the abnormal returns will follow as a consequence of that.

When we look at the combined quality measures – earnings predictive (EP) and earnings-non-predictive (ENP) similar conclusions hold. EP is significantly associated with returns on a 1-month and 36 months horizons (t-statistics of 5.50 and 3.40, respectively) while ENP is insignificant in both cases (t-statistics of 1.73 and 0.91, respectively). After controlling for future earnings growth the coefficients flip sign. EP becomes insignificant (t-statistic of 0.80) meaning that its return predictive power is fully attributed to its predictive power for future earnings growth. On the other hand, ENP becomes highly significant after controlling for the negative earnings surprises associated with it.

3.3. Performance of quality strategies

In this section, we test the performance of our earnings non-predictive (ROE, margins, ROE growth, leverage, and earnings variability) and earnings predictive (gross profitability, accruals, and investments) quality definitions.

Table 3
Predictive power of quality measures for stock returns
The table reports results of Fama-MacBeth (1973) regressions of stock returns on individual firm characteristics. Characteristics are calculated according to the Appendix, t-statistics are Newey-West adjusted using three lags for 1 month return regressions and 35 lags for 36-month return regressions. All regressions correct for the following set of control variables (Controls): log(Mcap), log(Book-to-price), Momentum 12-1M, Beta 3Y, Volatility 3Y, Distance-to-Default, and region, sector dummies. The last row shows the average adjusted R-squared. Results are calculated on monthly data for the period January 1986 - December 2015 for 1 month returns and January 1986 – December 2012 for 3-year return for our Global markets sample. Panel A shows standard Fama-MacBeth regression with next month returns. Panel B shows results of regressions of future three year returns on the respective quality measure. Panel C controls for change in future three years earnings change ($\frac{Earnings_{t+\tau} - Earnings_t}{BE_t}$). Every column represents regressions for the respective quality measure. All right-hand side variables are standardized into z-scores and multiplied by the expected sign ("+" for ROE, Margins, ROE Growth, Gross profitability, ENP, and EP, and "-" for Leverage, Earnings variability, Accruals, and Investments). ENP stands for Earnings-non-predictive measures and EP stands for earnings-predictive measures. ENP is the equally weighted z-score of ROE, Margins, ROE Growth, Leverage, Earnings variability, and EP is the equally weighted z-score of Gross profitability, Accruals, and Investments.

	ROE	Margins	ROE growth	Leverage	Earnings variability	Gross profitability	Accruals	Investments	ENP	EP
Panel A: Regressions of next month returns on quality measures										
Intercept	1.00	0.97	1.02	1.03	1.02	1.05	1.00	1.03	1.02	1.02
t-stat	[2.91]	[2.85]	[2.99]	[3.04]	[3.02]	[3.07]	[2.92]	[3.01]	[2.98]	[3.00]
Quality measure	0.09	0.03	0.01	0.01	0.00	0.13	0.06	0.06	0.05	0.13
t-stat	[2.89]	[1.07]	[0.93]	[0.36]	[0.02]	[4.49]	[4.52]	[2.99]	[1.73]	[5.50]
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Panel B: Regressions of next three year returns on quality measures										
Intercept	49.72	48.58	49.74	49.43	49.71	49.94	49.28	49.86	49.73	49.48
t-stat	3.63	3.55	3.59	3.65	3.59	3.60	3.53	3.60	3.61	3.58
Quality measure	1.51	0.77	0.25	−0.39	−0.18	1.41	1.33	1.90	0.92	2.67
t-stat	[1.65]	[1.06]	[0.82]	[−0.36]	[−0.38]	[1.46]	[2.05]	[2.74]	[0.91]	[3.40]
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
Panel C: Regressions of next three year returns on quality measures, controlling for change in future earnings										
Intercept	45.76	38.48	47.07	47.15	48.44	47.26	46.89	46.28	47.82	46.34
t-stat	4.26	3.50	4.02	4.07	4.06	4.00	3.96	3.99	4.15	3.98
Quality measure	12.01	7.19	4.18	0.75	1.04	2.37	−0.21	−0.14	9.64	0.64
t-stat	[8.33]	[7.89]	[6.48]	[0.76]	[2.37]	[2.91]	[−0.33]	[−0.18]	[7.18]	[0.80]
ΔEarnings	22.66	21.04	20.32	19.30	19.35	19.29	19.29	19.37	21.48	19.24
t-stat	[17.19]	[18.50]	[18.92]	[19.59]	[19.54]	[19.72]	[19.35]	[19.31]	17.94	19.30
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.32	0.31	0.30	0.29	0.29	0.29	0.29	0.29	0.31	0.29

In the first part of the analysis, we present the performance of top, bottom, and the top-minus-bottom (henceforth T-B) quintile portfolios.

Panel A of Table 4 shows the performance of investment strategies based on the quality characteristics which are not associated with positive earnings growth. We also create an overall quality measure 'Combined' by constructing an investment strategy which uses an equally-weighted combination of all individual variables. Focusing on the T-B quintile portfolios, we see that all of them produce small positive returns and ROE seems to be superior to the rest with an annualized return of 3.0%. Due to short-sale constraints practitioners often focus on the top quintile portfolio. Therefore, we also present separate results for the long and the short leg of the self-financing portfolio. By looking at the top quintile portfolio we notice that, with the exception of leverage and ROE growth, all variables outperform the market portfolio. The combined quality strategy generates a T-B quintile return of 2.0%. However, none of the T-B return spreads are statistically different from zero.

Controlling for the standard risk factors such as market beta, size, value, and momentum, Panel B shows a similar picture. The strongest variable (ROE) has positive loadings on the value and momentum factors, and the alpha of 2.6% per annum is again not statistically different from zero (t-statistic of 1.49). In terms of factor loadings, the combined quality strategy is similar to ROE in terms of factor loadings but results in a marginally significant alpha of 3.1% per annum (t-stat 2.04). One variable that stands out is Earnings variability with a four-factor alpha of 4.1% per annum (t-statistic of 3.68). Its market loading of -0.27 (t-statistic of -12.14) hints that it behaves like another well-known effect, namely the low-risk effect documented by Black, Jensen, and Scholes (1972), Blitz and van Vliet (2007), and Frazzini and Pedersen (2014). Results from Panel A confirm this notion as the top portfolio has a volatility of 13.3% and the bottom – 18.3% compared to market volatility of 16.0%. Therefore, its usage as a quality indicator is questionable since it can also be seen as a low-risk measure.

In Table 5, we show similar information but now for quality characteristics that are associated with positive future earnings growth.

Panel A shows that the T-B portfolios for all three characteristics have positive returns – 4.0% for gross profitability, 2.6% for accruals, and 3.2% for investments. Furthermore, all top quintile portfolios also outperform the total market portfolio. The combined quality definition clearly benefits from diversification as it has better performance than each individual characteristic (T-B return of 5.1% with comparable volatility).

The earnings predictive definitions remain strong after correcting for other risk factors as each individual factor has a highly significant alpha. Novy-Marx (2013) has documented that stocks with high gross profitability tend to be relatively more expensive and that a good working investment approach is to combine profitability and value or the so-called 'quality at a reasonable price' strategy. Our global results point in the same direction as gross profitability has a negative (but insignificant) loading on HML. However, Table 5 indicates that an investor can also achieve a performance improvement by diversifying across multiple quality signals. The combined earnings predictive quality strategy has an alpha of 6.0% per annum (t-statistic of 5.04), which is substantially higher than gross profitability, accruals, and investments stand alone. Further, the earnings predictive quality factor is superior to the earn-

Table 4
Performance of earnings non-predictive quality measures
The table shows performance characteristics for multiple quality strategies. Panel A consists of returns, volatilities, and Sharpe ratios for Top, Bottom, and Top minus Bottom (T-B) portfolios sorted on the relevant factor. Top is the portfolio with the highest 20% ranked stocks, Bottom is the portfolio with the lowest 20% ranked stocks, and T-B is a self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). The factors are calculated as explained in the Appendix. Returns and volatilities are estimated based on monthly data and then annualized. Panel B contains regression coefficients based on Fama and French / Carhart 4-factor model. The factors used are based on our replication of original factors and are based on the investment universe used for the analysis. Alphas are annualized. The sample period is January 1986 - December 2015.

		ROE	Margins	ROE growth	Leverage	Earnings variability	Combined (ENP)	Universe
Panel A: Performance of Top, Bottom, and Top-minus-Bottom portfolios								
Top	Return	9.7%	8.3%	8.2%	7.8%	9.1%	9.5%	8.2%
	Volatility	15.8%	15.3%	17.0%	16.0%	13.3%	14.9%	16.0%
	Sharpe ratio	0.61	0.54	0.48	0.49	0.69	0.64	0.52
Bottom	Return	6.6%	7.4%	7.9%	7.6%	8.0%	7.5%	
	Volatility	20.0%	20.6%	18.8%	16.1%	18.3%	19.7%	
	Sharpe ratio	0.33	0.36	0.42	0.47	0.44	0.38	
T-B	Return	3.0%	0.9%	0.4%	0.2%	1.0%	2.0%	
		[1.41]	[0.35]	[0.27]	[0.15]	[0.69]	[0.99]	
	Volatility	11.5%	13.3%	6.8%	7.6%	8.1%	10.9%	
	Sharpe ratio	0.26	0.07	0.05	0.03	0.13	0.19	
Panel B: Fama and French 4 factor regression coefficients								
	alpha	2.6%	2.0%	0.0%	0.5%	4.1%	3.1%	
	t-stat	[1.49]	[1.03]	[−0.03]	[0.38]	[3.68]	[2.04]	
	Mkt-RF	−0.06	−0.14	−0.05	−0.05	−0.27	−0.17	
	t-stat	[−1.86]	[−3.71]	[−2.09]	[−1.70]	[−12.14]	[−5.47]	
	SMB	−0.43	−0.82	0.11	0.26	−0.19	−0.45	
	t-stat	[−4.80]	[−8.31]	[1.84]	[3.64]	[−3.31]	[−5.65]	
	HML	0.16	0.47	−0.16	−0.35	−0.03	0.16	
	t-stat	[2.19]	[5.79]	[−3.20]	[−6.07]	[−0.62]	[2.43]	
	WML	0.31	0.21	0.19	0.08	0.07	0.25	
	t-stat	[7.87]	[4.70]	[7.15]	[2.52]	[2.70]	[6.94]	
	R-squared (%)	0.38	0.43	0.19	0.10	0.47	0.46	

Table 5
Performance of earnings predictive quality measures
The table shows performance characteristics for multiple quality strategies. Panel A consists of returns, volatilities, and Sharpe ratios for Top, Bottom, and Top minus Bottom (T-B) portfolios sorted on the relevant factor. Top is the portfolio with the highest 20% ranked stocks, Bottom is the portfolio with the lowest 20% ranked stocks, and T-B is a self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). The factors are calculated as explained in the Appendix. Returns and volatilities are estimated based on monthly data and then annualized. Panel B contains regression coefficients based on Fama and French / Carhart 4-factor model. The factors used are based on our replication of original factors and are based on the investment universe used for the analysis. Alphas are annualized. The sample period is January 1986 - December 2015. Returns of the top and bottom portfolios are in excess of the risk free rate.

		Gross profitability	Accruals	Investments	Combined (EP)	Universe
Panel A: Performance of Top, Bottom, and Top-minus-Bottom portfolios						
Top	Return	10.1%	9.2%	9.6%	10.4%	8.2%
	Volatility	14.7%	17.1%	17.1%	15.8%	16.0%
	Sharpe ratio	0.68	0.54	0.56	0.66	51.5%
Bottom	Return	6.1%	6.6%	6.4%	5.2%	
	Volatility	16.9%	17.3%	18.7%	18.4%	
	Sharpe ratio	0.36	0.38	0.34	0.29	
T-B	Return	4.0%	2.6%	3.2%	5.1%	
		[2.79]	[2.44]	[1.93]	[3.75]	
	Volatility	7.7%	5.6%	8.9%	7.3%	
	Sharpe ratio	0.52	0.46	0.36	0.71	
Panel B: Fama and French 4 factor regression coefficients						
	alpha	5.3%	2.8%	3.2%	6.0%	
	t-stat	[4.10]	[2.66]	[2.38]	[5.04]	
	Mkt-RF	−0.11	0.01	−0.08	−0.11	
	t-stat	[−4.35]	[0.35]	[−2.81]	[−4.65]	
	SMB	−0.18	−0.24	0.01	−0.25	
	t-stat	[−2.73]	[−4.40]	[0.19]	[−3.99]	
	HML	−0.07	0.14	0.52	0.38	
	t-stat	[−1.32]	[3.07]	[8.97]	[7.54]	
	WML	0.08	−0.02	−0.05	0.00	
	t-stat	[2.62]	[−0.90]	[−1.75]	[−0.03]	
	R-squared (%)	0.22	0.05	0.36	0.25	

Table 6
International performance of earnings predictive and earnings non-predictive quality factors
The table shows returns and alphas of the combined earnings non-predictive and earnings predictive quality definitions for multiple regions. Panel A shows returns of Top minus Bottom (T-B) quality portfolios. T-B is a self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). Returns are estimated based on monthly data and then annualized. Panel B contains annualized 4-factor Fama and French / Carhart alphas per region. The factors used are based on our replication of original factors using the same investment universe as used for the analysis. The universe definitions of United States, Europe, and Japan are based on carveouts of these regions from our Global markets universe. Emerging markets universe is based on the biggest 500 stocks measured by market capitalization. The sample period is January 1986 - December 2015 for Global markets, United States, Europe, and Japan and January 1993 - December 2015 for Emerging markets.

	Earnings non-predictive	Earnings predictive
Panel A: Top-minus-Bottom return differential		
United States	1.1%	6.5%
t-stat	[0.47]	[3.79]
Europe	2.8%	5.2%
t-stat	[1.45]	[4.05]
Japan	−2.9%	2.8%
t-stat	[−1.06]	[1.75]
Global markets	2.0%	5.1%
t-stat	[1.02]	[3.86]
Emerging markets	0.9%	6.2%
t-stat	[0.37]	[2.66]
Panel B: Fama and French 4 factor alphas		
United States	3.3%	6.7%
t-stat	[2.15]	[4.06]
Europe	4.7%	5.2%
t-stat	[3.20]	[4.01]
Japan	2.6%	2.7%
t-stat	[1.11]	[1.66]
Global markets	3.1%	6.0%
t-stat	[2.04]	[5.04]
Emerging markets	5.6%	8.7%
t-stat	[2.87]	[4.09]

ings non-predictive one for both T-B raw returns and after correcting for risk factors.

4. Robustness tests

4.1. Regional and emerging markets results

In this section, we extend the scope of the study as well as check for robustness of our results across regions. Section 3 presents results on global large- and mid-capitalization stocks which are commonly used as an investment universe by practitioners. Our findings confirm previously documented U.S. results on profitability, accruals, and investments. However, what we find could potentially be driven by a strong systematic U.S. bias in the data which results in us effectively comparing the performance of U.S. to non-U.S. stocks. As such, we aim to provide evidence that the global results are not just the result of some systematic regional allocation bets. We therefore further split the Global universe into three main regions – United States, Europe, and Japan as well as add Emerging markets for additional out of sample robustness tests.

Table 6 summarizes the performance for the two combined quality strategies – Earnings non-predictive and Earnings predictive. The main takeaway is that the combined 'earnings predictive' strategy consistently outperforms 'earnings non-predictive' one based on both T-B returns as well as alphas.

Panel A compares the long-short return of the two strategies. Focusing on the combined 'earnings predictive' definition, we see that the T-B returns for the United States are highest within global developed markets. Furthermore, the composite 'earnings predictive' quality factor yields positive returns in all regions (significant with the exception of Japan). On the other hand, the 'earnings non-predictive' quality definition does not exhibit returns which are statistically distinguishable from zero. Finally, the emerging market results reinforce the superiority of the 'earnings predictive' definition over the 'earnings non-predictive' one. These results can serve as a true out-of-sample test as this universe is much less looked at in academic studies. Correcting for the 4-factor model of Fama and French/Carhart in Panel B yields similar conclusions meaning that the results cannot be attributed to the well-known size, value, momentum, and market beta. Finally, also factor spanning tests (not reported in Table 6 for brevity) in which we regress the earnings predictive T-B returns on earnings-non-predictive T-B returns (with and without adding the Fama and French/Carhart factors) yield highly significant alphas in all of our regions with the exception of Japan.

The results for individual variables reinforce our conclusions. Fig. 1 shows that within every region, earnings predictive measures (gross profitability, accruals, and investments) are stronger than the earnings non-predictive ones. Furthermore, all 'earnings predictive' variables have positive T-B quintile returns in all regions (though returns for gross profitability is weak in Japan and investments – in Emerging markets). On the other hand, for the 'earnings non-predictive' definitions, we find mixed results across regions.

A further examination of the two strategies is shown in Fig. 2, which plots their regional performance in the volatility-return space. There we see consistently high Sharpe ratios for 'earnings predictive' quality definitions across regions compared to its 'earnings non-predictive' counterpart.

4.2. Regional cross-sectional regressions

After documenting the standalone portfolio returns and four-factor model alphas in the previous two sections, we are now interested which quality variables carry unique information and whether this holds in an international setup. Therefore, we employ the Fama-MacBeth (1973) methodology to estimate the marginal effects of the single quality variables after controlling for each other.

In Table 7, we estimate the marginal effects of the single quality variables after controlling for each other, all controlled for the standard factors size, beta, value, and momentum, as well as risk proxies such as market beta, volatility, and distress risk, and region as well as sector dummies.

Starting with our Global sample, we see that the marginal predictive power of the earnings predictive variables – gross profitability, accruals, and investments – have significant predictive power for future stocks returns while with the exception of ROE, the non-earnings predictive variables also have no marginal predictive power for stock returns. When we split the sample into sub-regions – United States, Europe, and Japan - results remain qualitatively similar and earnings predictive measures have systematically stronger predictive power compared to earnings non-predictive ones. In Emerging Markets, the same relationship generally hold with the exception that ROE has positive marginal predictive power for stock returns and Investments is insignificant.

4.3. Quality and distress-risk

Our results so far indicate that the predictive power of quality variables for stock returns originates from the variables being good proxies for future earnings growth. If this earnings growth

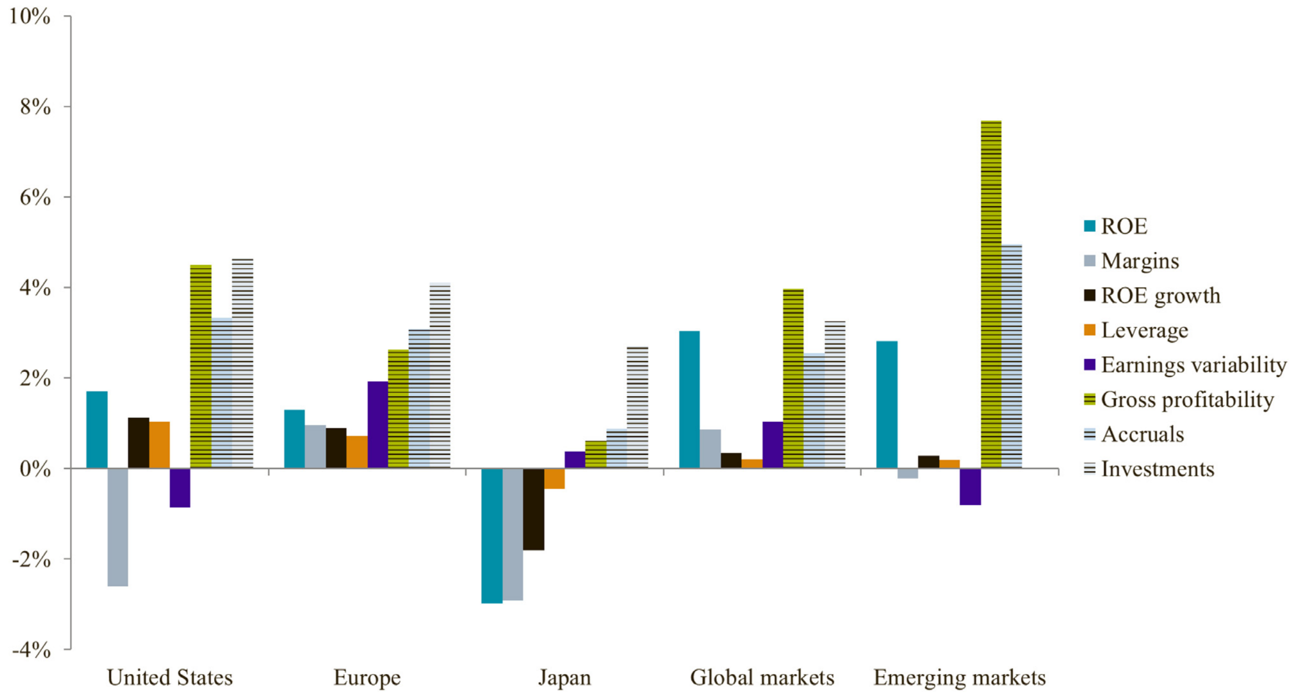

Fig. 1. International performance of different quality characteristics
In Fig. 1 we show Top-minus-Bottom (T-B) returns of the alternative quality definitions for multiple regions. T-B is a self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). Returns are estimated based on monthly data and then annualized. The universe definitions of United States, Europe, and Japan are based on carve-outs of these regions from our Global markets universe. Emerging markets universe is based on the biggest 500 stocks measured by market capitalization. The sample period is January 1986 - December 2015 for Global markets, United States, Europe, and Japan and January 1993 - December 2015 for Emerging markets.

Fig. 2. International performance of different quality characteristics
In Fig. 2 we show volatility-return scatter plots of the Earnings non-predictive and Earnings predictive Quality definitions per region. Results apply for a Top-minus-Bottom (T-B) self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). Returns and volatilities are estimated based on monthly data and then annualized. The universe definitions of United States, Europe, and Japan are based on carve-outs of these regions from our Global markets universe. Emerging markets (EM) universe is based on the biggest 500 stocks measured by market capitalization. The sample period is January 1986 - December 2015 for Global markets, United States, Europe, and Japan and January 1993 - December 2015 for Emerging markets.

Table 7
Regional Fama-MacBeth (1973) regressions
The table reports results of Fama-MacBeth (1973) regressions of monthly stock returns on individual firm characteristics. Characteristics are calculated according to the Appendix, *t*-statistics are Newey-West adjusted using three lags. The regressions include all quality variables simultaneously and correct for the following set of control variables: log(Mcap), log(Book-to-price), Momentum 12-1M, Beta 3Y, Volatility 3Y, Distance-to-Default, and region, sector dummies. All right-hand side variables are standardized into z-scores and multiplied by the expected sign ("+" for ROE, Margins, ROE Growth, Gross profitability, ENP, and EP, and "-" for Leverage, Earnings variability, Accruals, and Investments). The last row shows the average adjusted R-squared. The sample period is January 1986–December 2015 for Global markets, United States, Europe, and Japan and January 1993–December 2015 for Emerging markets. The universe definitions of United States, Europe, and Japan are based on carveouts of these regions from our Global markets universe. Emerging markets (EM) universe is based on the biggest 500 stocks measured by market capitalization.

	US		EU		JAP		Global		EM	
	mean	*t*-stat	mean	*t*-stat	mean	*t*-stat	mean	*t*-stat	mean	*t*-stat
Intercept	1.05	2.90	1.08	3.14	0.62	1.46	1.00	2.98	1.36	2.36
ROE	0.12	2.78	0.12	3.74	0.06	1.34	0.12	3.58	0.15	2.09
Margins	−0.06	−1.74	−0.02	−0.61	0.02	0.55	−0.02	−0.83	−0.02	−0.40
ROE growth	−0.02	−0.98	−0.03	−1.40	−0.04	−1.46	−0.02	−1.61	−0.06	−1.44
Leverage	0.02	0.47	0.03	1.12	−0.14	−2.97	−0.02	−0.94	−0.02	−0.36
Earnings Variability	0.00	0.19	0.01	0.58	0.01	0.51	−0.01	−0.36	0.02	0.69
Gross profitability	0.06	1.21	0.06	1.91	0.09	2.13	0.10	3.46	0.23	4.00
Accruals	0.07	3.30	0.05	2.48	0.06	2.27	0.06	4.53	0.12	2.94
Investments	0.03	0.95	0.09	4.05	0.03	1.00	0.06	3.13	0.00	0.03
ln(mcap)	−0.12	−2.71	−0.08	−1.79	−0.09	−1.39	−0.09	−2.51	−0.05	−0.67
Beta 3Y	−0.02	−0.28	0.07	1.06	0.03	0.56	0.05	0.78	0.09	0.98
Volatility 3Y	−0.01	−0.16	−0.03	−0.54	−0.07	−1.23	−0.04	−1.01	−0.12	−1.55
Distance-to-default	0.06	1.27	0.05	1.22	0.01	0.16	0.05	1.31	0.01	0.16
ln(Book-to-price)	0.11	1.51	0.22	3.86	0.41	6.03	0.24	4.56	0.43	4.40
Momentum 12-1	0.08	1.29	0.30	4.39	0.07	0.99	0.21	3.67	0.45	5.21
R-squared (%)	0.18		0.13		0.19		0.17		0.16	

Table 8
Distance-to-default and earnings predictive vs. earnings non-predictive quality factors
The table shows returns and alphas of the combined earnings non-predictive and earnings predictive quality definitions for five distance-to-default groups. Panel A shows returns of Top minus Bottom (T-B) quality portfolios. T-B is a self-financing portfolio which is long the top 20% stocks (Top) and short the bottom 20% stocks (Bottom). Returns are estimated based on monthly data and then annualized. Panel B contains annualized 4-factor Fama and French / Carhart alphas per distance-to-default group. The factors used are based on our replication of original factors and are based on the investment universe used for the analysis. Alphas are annualized. The sample period is January 1986 - December 2015.

Distance-to-default group	Earnings non-predictive	Earnings predictive
Panel A: Top-minus-Bottom return differential		
Low	0.9%	9.4%
t-stat	[0.32]	[3.49]
2	2.3%	5.0%
t-stat	[1.14]	[2.97]
3	3.1%	4.3%
t-stat	[1.64]	[3.09]
4	2.2%	3.1%
t-stat	[1.39]	[2.44]
High	4.1%	2.0%
t-stat	[3.01]	[1.78]
Panel B: Fama and French 4 factor alphas		
Low	0.6%	9.4%
t-stat	[0.23]	[3.67]
2	2.2%	5.8%
t-stat	[1.13]	[3.54]
3	3.7%	4.5%
t-stat	[2.10]	[3.24]
4	2.4%	3.8%
t-stat	[1.63]	[3.09]
High	4.5%	2.3%
t-stat	[3.40]	[2.05]

is unexpected by the market (e.g. due to limited attention of investors) then high-quality firms will be undervalued relative to low-quality firms. On average, such mispricing will be corrected in the long run. However, also a risk-based explanation could be possible although it seems counterintuitive that high-quality firms (i.e., firms with high gross profitability, low accruals, and low investments) are more risky than low-quality firms. Furthermore, the relation between quality and stock returns remains after controlling for well-known factors in time-series and cross-sectional regressions.[12]

Although these standard tests do not indicate a risk-based explanation, it could be possible that the profitability of the quality premium is concentrated in the stocks with the highest distress risk (see also Avramov et al., 2013). To investigate this hypothesis, we apply a dependent double-sort in which we first sort all stocks into quintiles by distress risk, measured by distance-to-default (DtD). Then, within each DtD quintile, we sort stocks into quintiles by the firm's earnings predictive (EP) or earnings-non-predictive (ENP) composite score. In Table 8, we present returns and alphas for the top-minus-bottom quality portfolio within each DtD group.

The double sorts show that the top-minus-bottom returns and alphas of the earnings-predictive measure are positive and significant across distress risk groups suggesting that the return premium is not concentrated in the stocks with the highest levels of distress risk. In contrast, the EP premiums are higher in the lowest distress risk group. For the earnings-non-predictive top-minus-bottom returns, we find only significant returns for the highest distance-to-default group. The remaining distance-to-default group exhibit lower and statistically insignificant returns. Based on these results we conclude that mispricing is a more likely cause for the existence of the quality premium than distress risk is.

If investors with limited attention are surprised by the (unexpected) earnings growth, the predictive information contained in the earnings-predictive signal should only gradually diffuse in market prices (consistent with investor underreaction). In contrast, in the case of investor overreaction, positive returns associated with the predictive information would be followed by a subsequent re-

[12] In the cross-sectional regressions, we additionally control for volatility and distance-to-default.

Table 9
Predictive power of alternative quality measures for stock returns
The table reports results of Fama-MacBeth (1973) regressions of stock returns on individual firm characteristics. Characteristics are calculated according to the Appendix, t-statistics are Newey-West adjusted using three lags for 1 month return regressions and 35 lags for 36-month return regressions. All regressions correct for the following set of control variables (Controls): log(Mcap), log(Book-to-price), Momentum 12-1M, Beta 3Y, Volatility 3Y, Distance-to-Default, and region, sector dummies. The last row shows the average adjusted R-squared. Results are calculated on monthly data for the period January 1986 - December 2015 for 1 month returns and January 1986 – December 2012 for 3-year return for our Global markets sample. Panel A shows standard Fama-MacBeth regression with next month returns. Panel B shows results of regressions of future three year returns on the respective quality measure. Panel C controls for change in future three years earnings change ($\frac{Earnings_{t+\tau} - Earnings_t}{BE_t}$). Every column represents regressions for the respective quality measure. All right-hand side variables are standardized into z-scores and multiplied by the expected sign ("+" for CFO, RNOA, and "-" for NOAA).

	CFO	NOAA	RNOA
Panel A: Regressions of next month returns on quality measures			
Intercept	0.93	1.09	1.01
t-stat	[2.74]	[3.25]	[2.94]
Quality measure	0.18	0.13	0.10
t-stat	[6.12]	[3.84]	[3.79]
Controls	yes	yes	yes
R-squared (%)	0.16	0.16	0.16
Panel B: Regressions of next three year returns on quality measures			
Intercept	48.38	50.95	49.87
t-stat	[3.52]	[3.66]	[3.61]
Quality measure	4.23	2.72	1.69
t-stat	[3.73]	[2.50]	[2.22]
Controls	yes	yes	yes
R-squared (%)	0.18	0.18	0.18
Panel C: Regressions of next three year returns on quality measures, controlling for change in future earnings			
Intercept	44.66	47.54	46.55
t-stat	[3.85]	[4.03]	[4.03]
Quality measure	5.97	1.85	4.49
t-stat	[6.04]	[1.60]	[6.41]
ΔEarnings	19.51	19.29	19.70
t-stat	[19.69]	[19.77]	[19.72]
Controls	yes	yes	yes
R-squared (%)	0.30	0.29	0.30

turn reversal. Our results in Table 2 indicate that earnings predictive quality measures have predictive power for returns which extend to 36 months in the future. Therefore, we conclude that underreaction to the predictive signal is more likely than overreaction as we do not observe a return reversal over the 36 months after portfolio formation.

4.4. Alternative quality measures

As mentioned above, accounting-based factors show a considerable dispersion in definitions. We aim to select variables that are seen of particular importance by key academic studies or practitioners and are rather pure and at the extremes of potential definition spectrum. E.g., gross profitability is based on the highest income statement item gross profits, while ROE is based on bottom-line net income. Similarly, accruals can be seen as the difference between operating book profits and operating cash flows. Nevertheless, we understand that our choice can be seen as somewhat arbitrary and also other variables or combinations of existing variables could be considered. For instance, Sloan (1996) investigates also cash profitability (CFO, i.e., removing the accruals part from net income); Hirshleifer et al. (2004) introduce net operating assets to lagged asset (NOAA), with net operating assets defined as the sum of cumulative operating accruals and cumulative investments; and Penman and Zhang (2002) use return on net operating assets (RNOA), defined as operating income after tax (between gross profits and net income) to net operating assets.[13] We investigate the performance of these variables in Table 9.

Removing the accruals part from net income (CFO), combining cumulative accruals and cumulative investments (NOAA) or scaling operating income with net operating assets (RNOA), leads to variables that can predict returns, both on the one-month and on the 36-month horizon as shown in Panels A and B.[14] When controlling for future earnings growth in Panel C, we observe a similar pattern as in Table 3. Variables that do not represent a level of profitability (NOAA) become insignificant while CFO and RNOA remain significant.

4.5. Quality in small caps

In our main analysis we restrict our sample to investable stocks by using a large- and mid-cap universe. In this section, we expand our sample by also including small caps in order to show how they affect our main results. Specifically, we use the S&P Developed BMI index as an investment universe which contains around 6,000 non-financial stocks as of ultimo 2015, compared to around 1,900 in FTSE Developed Index.

In Table 10 we present the same analysis as in Table 3 but on the broader universe. The main conclusion is that result remain economically similar to the ones in Table 3 but with higher significance levels.

On a one-month return horizon Earnings predictive measures (EP) have around three times higher predictive power for future returns compared to Earnings-non-predictive (ENP), coefficients of

[13] We thank an anonymous reviewer for suggesting these alternative quality variables.

[14] The higher statistical significance for RNOA compared to ROE (see Table 3) is mainly driven by the numerator, i.e., operating income after tax is a stronger predictor than net income. Similarly, also operating profitability as defined in Ball (2015) shows strong (unreported) results.

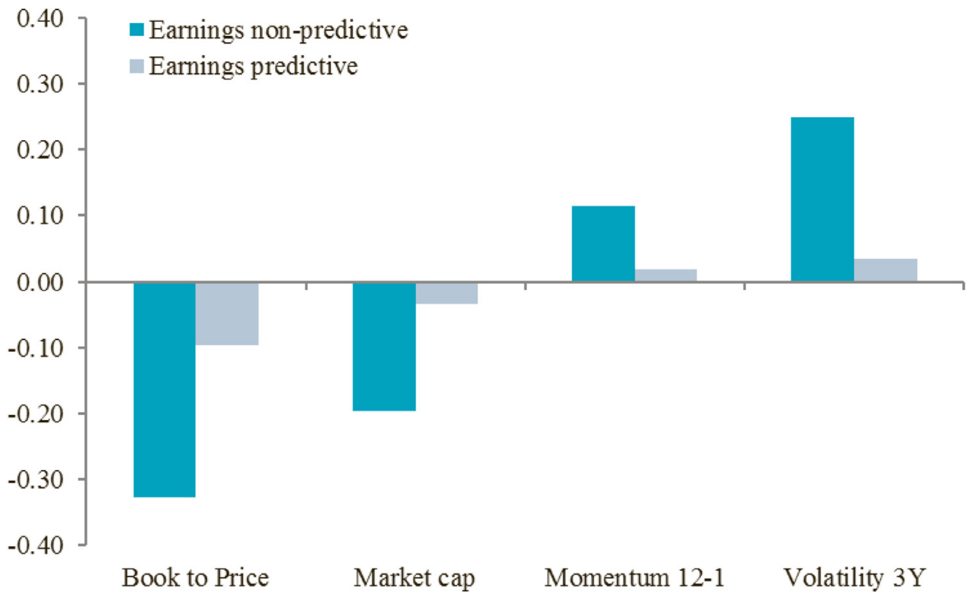

Fig. 3. Rank correlation between quality and other factors
In Fig. 3 we show average rank correlation of the Earnings predictive and Earnings non-predictive Quality definitions with Book to Price, the negative of Market capitalization in USD (Market cap), past 12 minus 1 month return (Momentum 12-1), and the negative of past 3 years monthly volatility (Volatility 3Y). Each month the rank correlation is calculated and then averaged over the full sample. Results are estimated based on our Global universe and the sample period is January 1986 - December 2015.

0.21 (t-statistics 7.68) and 0.07 (t-statistic 2.38), respectively. This relationship remains at the 36 month-horizon as EP is still highly significant (t-statistic 7.87) and ENP is at the border of significance with a t-statistic of 1.96. Just like in our main results, this relation inverts once we control for future earnings growth. The coefficient of EP goes down from 4.74 to 1.41 and the coefficient on ENP goes up from 1.73 to 10.78. This effect reinforces our main results as it shows that even when small cap stocks are considered, future earnings growth is the main underlying factor for the relation between future stock returns and quality measures.

4.6. Corporate bonds

With this section, we aim at two main objectives. First, gather strong evidence for the robustness of quality as a factor by testing it in a fundamentally different setting than previously done in the literature. Second, stimulate future research on the existence of similar underlying return drivers across asset classes (e.g. Bhojraj and Swaminathan, 2009; Correia et al., 2012; Jostova et al., 2013; Haesen et al., 2017; Houweling and Van Zundert, 2017).

To do so, we directly apply our ‘earnings non-predictive’ and ‘earnings predictive’ combined quality definitions from the previous section on corporate bonds. Corporate bonds fundamentally differ from equities with features such as maturity date, duration, and interest rate risk. The latter one has no impact on our results due to using excess returns over duration matched securities, focusing on the default premium. Nevertheless, we acknowledge that for a proper quality definition, further adjustments to the variables could be made. Using simple equity quality definitions makes our results conservative.

Table 11 shows performance statistics for both investment grade and high yield bonds. The top portfolio investment grade bonds based on both quality definitions outperforms the market in terms of excess return as well as on a risk-adjusted basis (Sharpe ratios of 0.15 and 0.22 compared to 0.08 for the market) showing evidence for a quality premium in the corporate bond space. The ‘earnings predictive’ definition stands out in terms of identifying ‘low quality’ bonds as the bottom portfolio performs worse than the bottom industry portfolio and the market portfolio. These results in a significant top-minus-bottom premium 0.6% (t-stat 1.99) for the ‘earnings predictive’ definition compared to 0.0% for the ‘earnings non-predictive’ one.

The results for high yield bonds show strong evidence that an investment strategy based on quality can also be profitably applied in corporate bond markets. Furthermore, the superiority of the ‘earnings predictive’ definition proves robust once again with a top-minus-bottom premium of 4.3% (t-stat of 3.57) compared to 1.3% for the ‘earnings non-predictive’ definition. The better performance of quality among high yield bonds relative to the performance in investment grade bonds can be partially attributed to the relative riskiness of both segments. In corporate bonds, the downside risk, heavily influenced by defaults, is generally much higher than the upside potential. A closer examination of the risk and return profiles of the top and bottom quality portfolios hints that investing in high-quality bonds effectively lowers the risk of default, as well as earns a return premium.

4.7. Quality and other factor premiums

Finally, we discuss the relation between quality-related and other factor premiums to address the question of whether it is a separate factor or just a reframing of already documented effects. The results of the previous sections show that the earnings predictive quality definition seems to be a robust and also sizeable new factor as the premiums exist within several regions and based on a large- and mid-cap investable sample. For the ‘earnings non-predictive’ quality definition, however, we observe overall weaker results. Furthermore, some observations, such as the low beta of the earnings variability variable raise the question if there is some overlap between factors. Naturally, the answer to this question depends on the exact definition of the anomaly which we aim to clarify with this paper. Apart from the single factor academic definitions of among others Sloan (1996), Novy-Marx (2013), and Fama and French (2015), the studies of Piotroski (2000) and Asness et al. (2019) propose more complex quality factor composition consisting of multiple characteristics separated in thematic groups. One of these groups – namely stability - is also related to the low-risk anomaly documented by Black, Jensen, and Scholes (1972), and Blitz and van Vliet (2007).

Table 10
Predictive power of quality measures for stock returns – small cap sample
The table reports results of Fama-MacBeth (1973) regressions of stock returns on individual firm characteristics. Characteristics are calculated according to the Appendix, *t*-statistics are Newey-West adjusted using three lags for 1 month return regressions and 35 lags for 36-month return regressions. All regressions correct for the following set of control variables (Controls): log(Mcap), log(Book-to-price), Momentum 12-1M, Beta 3Y, Volatility 3Y, and Distance-to-Default. The last row shows the average adjusted R-squared. Results are calculated on monthly data for the period January 1986 - December 2015 for 1 month returns and January 1986 – December 2012 for 3-year return for Broad Market Global markets sample where number of stocks varies between ~1500 in 1986 and 6000 in 2015. Panel A shows standard Fama-MacBeth regressions with next month returns. Panel B shows results of regressions of future three year returns on the respective quality measure. Panel C controls for change in future three years earnings change ($\frac{Earnings_{t+\tau} - Earnings_t}{BE_t}$). Every column represents regressions for the respective quality measure. All right-hand side variables are standardized into z-scores and multiplied by the expected sign ("+" for ROE, Margins, ROE Growth, Gross profitability, ENP, and EP, and "-" for Leverage, Earnings variability, Accruals, and Investments). ENP stands for Earnings-non-predictive measures and EP stands for earnings-predictive measures. ENP is the equally weighted z-score of ROE, Margins, ROE Growth, Leverage, Earnings variability, and EP is the equally weighted z-score of Gross profitability, Accruals, and Investments.

	ROE	Margins	ROE growth	Leverage	Earnings variability	Gross profitability	Accruals	Investments	ENP	EP
Panel A: Regressions of next month returns on quality measures										
Intercept	0.88	0.82	0.89	0.87	0.85	0.92	0.85	0.93	0.88	0.91
t-stat	[2.11]	[1.96]	[2.13]	[2.13]	[2.05]	[2.23]	[2.03]	[2.26]	[2.13]	[2.20]
Quality measure	0.13	0.06	0.00	−0.01	−0.03	0.19	0.08	0.13	0.07	0.21
t-stat	[3.77]	[2.38]	[−0.10]	[−0.22]	[−2.19]	[5.48]	[5.52]	[5.28]	[2.38]	[7.68]
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Panel B: Regressions of next three year returns on quality measures										
Intercept	49.29	47.60	49.09	47.98	48.15	49.95	48.08	49.71	49.36	49.41
t-stat	3.39	3.30	3.35	3.42	3.32	3.39	3.27	3.40	3.38	3.36
Quality measure	3.24	1.72	0.27	−1.06	−0.70	4.07	2.02	2.71	1.73	4.74
t-stat	[3.46]	[2.67]	[0.99]	[−0.81]	[−1.91]	[3.76]	[4.38]	[4.13]	[1.96]	[7.87]
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.14	0.13	0.13	0.13	0.13	0.14	0.13	0.13	0.13	0.14
Panel C: Regressions of next three year returns on quality measures, controlling for change in future earnings										
Intercept	45.92	38.55	45.91	45.40	46.10	46.47	45.43	44.26	47.30	45.08
t-stat	3.98	3.36	3.74	3.81	3.72	3.74	3.67	3.70	3.87	3.70
Quality measure	12.28	7.66	4.54	0.33	0.75	4.60	−0.48	−0.44	10.78	1.41
t-stat	[6.92]	[7.30]	[6.82]	[0.29]	[2.21]	[4.15]	[−0.95]	[−0.43]	[6.84]	[2.01]
ΔEarnings	26.34	25.04	24.28	23.28	23.37	23.37	23.36	23.51	25.49	23.23
t-stat	[16.24]	[17.53]	[18.15]	[18.61]	[18.69]	[18.78]	[18.63]	[18.29]	[16.85]	[18.44]
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared (%)	0.27	0.26	0.25	0.25	0.25	0.25	0.25	0.25	0.26	0.25

Table 11
Corporate bonds analysis
The table shows performance characteristics for the Earnings predictive and Earnings non-predictive quality strategies for U.S. Investment Grade and U.S. High Yield corporate bonds. To calculate the return in month *t* we take the average return of the portfolios constructed from month *t-11* to *t*. Each month the universe is split in 5 portfolios sorted on the relevant factor as Top means the 20% highest quality bonds, and Bottom means 20% lowest quality bonds. T-B is the difference between the return of the Top portfolio and the return of the Bottom portfolio. Investment grade is defined as stocks with credit ratings AAA, AA, A, BBB; High Yield – BB, B, CCC, CC, C. Returns and volatilities are calculated based on monthly data and then annualized. The sample period is January 1994 - December 2015.

		Investment Grade			High Yield		
		Earnings non-predictive	Earnings predictive	Market	Earnings non-predictive	Earnings predictive	Market
Top	Return	0.5%	0.8%	0.3%	2.7%	4.7%	1.7%
	Volatility	3.60%	3.78%	4.27%	8.59%	10.10%	9.42%
	Sharpe ratio	0.15	0.22	0.08	0.32	0.47	0.18
Bottom	Return	0.5%	0.2%		1.4%	0.4%	
	Volatility	4.71%	4.62%		13.09%	11.97%	
	Sharpe ratio	0.11	0.04		0.11	0.03	
T-B	Return	0.0%	0.6%		1.3%	4.3%	
		[0.00]	[1.99]		[0.86]	[3.57]	
	Volatility	1.64%	1.50%		6.94%	5.70%	
	Sharpe ratio	0.00	0.42		0.18	0.76	

To give some new insights to this discussion, we aim to elaborate on how the 'earnings non-predictive' and 'earnings predictive' definitions overlap with other factors. However, unlike in Section 3.1, we do not focus on returns but rather on the underlying stocks that are favored by the two approaches.

Fig. 3 shows the average rank correlation of quality and value, size, momentum, and low-volatility factor portfolios. Indeed, we see that the 'earnings non-predictive' definition of quality is relatively highly correlated with low-volatility due to explicitly including characteristics that focus on stability. At the same time, these stocks tend to be relatively more expensive as the rank correlation with book-to-price is negative. The higher price of 'quality' is not a new insight as it has been documented by Novy-Marx (2013) and Asness et al. (2019). Both quality definitions show similar correlations with the other. However, the 'earnings predictive' quality is correlated to a much more limited extent making it a more independent factor. Its low rank-correlation of 0.03 with low-volatility shows that the defensive features of quality come indirectly as a result of the strong underlying fundamentals and not by directly targeting low-volatile companies.

5. Concluding comments

In this paper, we investigate a common set of accounting-based variables commonly referred to as measuring the quality of a firm, and test their predictive power for future earnings growth and stock returns. We find that the predictive power of quality factors originates from its measures being good proxies for future earnings growth. Quality measures can predict future stock returns if and only if they are good proxies for future earnings growth. Quality variables that are shown to be no good proxies for future earnings growth have no to little predictive power for stock returns. The potential predictive power of quality measures for stock returns can be fully attributed to their predictive power for future earnings growth. We find no relationship between the quality premium and distress risk. As such our results support the conventional wisdom that quality is a measure of true economic profitability and that mispricing is a more likely cause for the existence of the quality premium than distress risk is.

We also analyze the robustness of the predictive power of quality for stock returns in an international and multi-asset setting: we investigate the predictive power of quality measures for future stock returns in both the U.S., Europe, Japan, emerging markets, and corporate bond markets. Our results are consistent across regions, and asset classes – stocks and bonds issued by high-quality companies outperform those issued by low-quality companies if the quality measures used are good proxies for future earnings growth.

Appendix

Table A1.

Table A1
Variable definitions
The table shows detailed definition of each anomaly variable. We obtain the fundamental data, in order of preference, Compustat quarterly, Compustat annual, Worldscope quarterly, Worldscope semi-annual, Worldscope annual. To avoid a forward looking bias we lag Compustat data by three and Worldscope data by six months. The sign after each variable name shows the expected sign: (+) indicates that higher values are attractive and (-) indicates that lower values are attractive.

Name (sign), Reference	Formula	Description
ROE (+), Haugen and Baker (1996)	$\frac{NI_t}{BE_t}$	Income before extraordinary items (NI) divided by book equity (BE).
Margins (+)	$\frac{NI_t}{SALES_t}$	Income before extraordinary items (NI) divided by sales (SALES).
ROE growth (+)	$ROE_t - ROE_{t-12}$	12-months difference in ROE as defined above.
Earnings variability (-)	$\sqrt{\frac{1}{4}\sum_{y=0}^{-4}(\Delta ROE_y - \overline{\Delta ROE})^2}$	The standard deviation of y-o-y ROE growth over the last five years.
Leverage (-), George and Hwang (2010)	$\frac{Debt_t}{BE_t}$	Total debt (Debt) to book equity (BE). Lower values are attractive.
Accruals (-), Sloan (1996)	$\frac{\Delta((CA_t - Cash_t) - (CL_t - SD_t - TP_t))_t - Depr_t}{TA_t}$	The change in operating working capital (ΔWC) minus depreciation, depletion and amortization (Depr), all deflated by total assets (TA). Thereby, operating working capital is current asset (CA) minus cash and short term investments (Cash) minus changes in current liabilities (CL) plus short-term debt (SD) and taxes payable (TP) (both if available).
Investment (-), Cooper, Gulen, and Schill (2008)	$\frac{TA_t}{TA_{t-12}}$	The ratio of total assets (TA) in month t to total assets in month t−12.
Gross profitability (+), Novy-Marx (2013)	$\frac{SALES_t - COGS_t}{TA_t}$	Sales (Sales) minus cost of goods (COGS) sold both divided by total assets (TA)
Cash profitability (+), Sloan (1996)	$\frac{CFO_t}{TA_t}$	Net cash flow from operating activities (CFO) divided by total assets (TA)
Net operating assets to assets (NOAA) (-), Hirshleifer et al. (2004)	$\frac{(TA_t - CASH_t) - (TA_t - STD_t - LTD_t - MI - PSTK_t - CE_t)}{TA_{t-1}}$	Net operating assets (NOA) divided by lagged total assets. Net operating assets are calculated as the difference between operating assets and operating liabilities. Operating assets are total assets (TA) minus cash and short-term investment (CASH). Operating liabilities are total assets minus short-term (STD) and long-term debt (LTD), minus minority interest (MI), minus preferred stock (PSTK) and common equity (CE).
Return on net operating assets (RNOA) (+), Penman and Zhang (2002)	$\frac{OI_t * (1 - \text{ statutory tax rate})}{(NOA_{t-1} + NOA_t)/2}$	Operating income (OI) after taxes divided by average net operating assets (NOA). We measure after-tax operating income without allocating taxes between operating and financing activities and assume that any difference between effective tax rates on operating income and statutory rates (calculated as the Federal rate plus 2 percent) is temporary. For the Federal rate we assume 35 percent.
Distance-to-default (DtD) (+), Merton (1974)	$\frac{\log(\frac{V_t}{X_t})}{\sigma_E(1 - \frac{V_t}{X_t})}$	Distance-to-default (DtD) originates from the Merton (1974) structural model, in which equity is modeled as a call option on the firm's assets, with the strike price being the value of the debt. For the empirical estimation, we follow Haesen et al. (2017). We proxy the value of the firm (V) with the sum of book value of total liabilities and market value of equity. The strike (X) is set equal to the book value of total liabilities and is the past 52-week equity return volatility.
Earnings growth, Novy-Marx (2013)	$\frac{Earnings_{t+\tau} - \ Earnings_t}{BE_t}$	Earnings growth is calculated as future earnings minus current earnings, scaled by current book equity (BE)

CRediT authorship contribution statement

Georgi Kyosev: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Matthias X. Hanauer:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Joop Huij:** Supervision, Conceptualization, Writing - original draft, Writing - review & editing. **Simon Lansdorp:** Supervision, Conceptualization, Writing - original draft, Writing - review & editing.

References

Akbas, F., Jiang, C., Koch, P.D., 2017. The trend in firm profitability and the cross-section of stock returns. Account. Rev. 92 (5), 1–32.
Ambastha, M., Dor, A.B., Dynkin, L., Hyman, J., Konstantinovsky, V., 2010. Empirical duration of corporate bonds and credit market segmentation. J. Fixed Income 20 (1), 5–27.
Asness, C.S., Frazzini, A., Pedersen, L.H., 2019. Quality minus junk. Rev. Account. Stud. 24 (1), 34–112.
Avramov, D., Chordia, T., Jostova, G., Philipov, A., 2013. Anomalies and financial distress. J. Financ. Econ. 108 (1), 139–159.
Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V.V., 2015. Deflating profitability. J. Financ. Econ. 117 (2), 225–248.
Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V.V., 2016. Accruals, cash flows, and operating profitability in the cross section of stock returns. J. Financ. Econ. 121 (1), 28–45.
Banz, R.W., 1981. The relationship between return and market value of common stocks. J. Financ. Econ. 9 (1), 3–18.
Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. J. Finance 32 (3), 663–682.
Bhojraj, S., Swaminathan, B., 2009. How does the corporate bond market value capital investments and accruals? Rev. Account. Stud. 14 (1), 31–62.
Black, F., Jensen, M.C., Scholes, M., 1972. The capital asset pricing model: Some empirical tests. Stud. Theor. Cap. Mark 81 (3), 79–121.
Blitz, D., Van Vliet, P., 2007. The volatility effect: lower risk without lower return. J. Portfolio Manag. 102–113.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Finance 52 (1), 57–82.
Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. J. Finance 63 (4), 1609–1651.
Correia, M., Richardson, S., Tuna, İ, 2012. Value investing in credit markets. Rev. Account. Stud. 17 (3), 572–609.
Fairfield, P.M., Whisenant, J.S., Yohn, T.L., 2003. Accrued earnings and growth: implications for future profitability and market mispricing. Account. Rev. 78 (1), 353–371.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33 (1), 3–56.
Fama, E.F., French, K.R., 2008. Dissecting anomalies. J. Finance 63 (4), 1653–1678.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116 (1), 1–22.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. J. Polit. Econ. 607–636.
Franke, B., Müller, S., Müller, S., 2017. The q-factors and expected bond returns. J. Bank. Finance 83, 19–35.
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. J. Financ. Econ. 111 (1), 1–25.
George, T.J., Hwang, C.Y., 2010. A resolution of the distress risk and leverage puzzles in the cross section of stock returns. J. Financ. Econ. 96 (1), 56–79.
Haesen, D., Houweling, P., van Zundert, J., 2017. Momentum spillover from stocks to corporate bonds. J. Bank. Finance 79, 28–41.
Hanauer, M.X. and Huber, D., 2019. Constructing a powerful profitability factor: international evidence. Available at SSRN 3234436.
Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. J. Financ. Econ. 41 (3), 401–439.
Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? J. Account. Econ. 38, 297–331.
Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Rev. Financ. Stud. 28 (3), 650–705.
Houweling, P., Van Zundert, J., 2017. Factor investing in the corporate bond market. Financ. Anal. J. 73 (2), 100–115.
Huang, A.G., 2009. The cross section of cashflow volatility and expected stock returns. J. Empir. Finance 16 (3), 409–429.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Finance 48 (1), 65–91.
Jostova, G., Nikolova, S., Philipov, A., Stahel, C.W., 2013. Momentum in corporate bond returns. Rev. Financ. Stud. 26 (7), 1649–1693.
La Porta, R., 1996. Expectations and the cross-section of stock returns. J. Finance 51 (5), 1715–1742.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108 (1), 1–28.
McGuire, J.B., Schneeweis, T., Branch, B., 1990. Perceptions of firm quality: a cause or result of firm performance. J. Manag. 16 (1), 167–180.
Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. J. Finance 29 (2), 449–470.
Penman, S.H., Zhang, X.J., 2002. Accounting conservatism, the quality of earnings, and stock returns. Account. Rev. 77 (2), 237–264.
Piotroski, J.D., 2000. Value investing: the use of historical financial statement information to separate winners from losers. J. Account. Res. 1–41.
Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. J. Finance 63 (2), 921–945.
Richardson, S.A., Sloan, R.G., Soliman, M.T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. J. Account. Econ. 39 (3), 437–485.
Richardson, S., Tuna, I., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. J. Account. Econ. 50 (2–3), 410–454.
Sloan, R., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? (Digest summary). Account. Rev. 71 (3), 289–315.
Subrahmanyam, A., 2010. The cross-section of expected stock returns: what have we learnt from the past twenty-five years of research? Eur. Financ. Manag. 16 (1), 27–42.
Thomas, J.K., Zhang, H., 2002. Inventory changes and future returns. Rev. Account. Stud. 7 (2–3), 163–187.
Thomson Reuters, 2013. Worldscope database - data definitions guide. Issue 14.2 https://blogs.cul.columbia.edu/business/files/2014/02/Worldscope-Data-Definition-Guide-Issue-14.2.pdf.
Trammell, S., 2014. Quality control. CFA Instit. Mag. 25 (2), 29–33.