Firm Fundamentals,One-Period-Ahead Earnings Expectations and Expected Stock Returns

Pengguo Wang
University of Exeter

Zihang Peng
University of Sydney Business School

Demetris Christodoulou
University of Sydney Business School

December 30, 2017

Firm Fundamentals, One-Period-Ahead Earnings Expectations and Expected Returns

Abstract

In this paper, we develop a novel approach towards estimating firm level expected stock returns. Building on (Ohlson 1995; Feltham and Ohlson 1995) linear pricing rule, we show that the firm-level one-period-ahead expected stock return is a linear combination of book-to-market ratio, forward earnings yield, and a variable summarizing one-period-ahead value-relevant 'other information'. This 'other information' can be inferred by the firm's one-period-ahead earnings expectation and the current stock price. The empirical evidence shows that the expected return estimates are significantly positively associated with future realized returns and are associated with a range of return predictive variables.

1 Introduction

Estimating expected stock returns remains a challenge central to research and practice in accounting and finance. Existing literature documents that factor models developed in asset pricing have been proven to be no significant predictive power for future realized returns (Lyle and Wang 2015). It also documents that the popular implied cost of capital estimates are poorly associated with cross-sectional realized returns (Easton and Monahan 2005, Lee, So and Wang 2015). Motivated by (e.g. Ohlson 1995; Feltham and Ohlson 1995; Ashton and Wang 2013), this paper derives a novel approach towards estimating firm-level one-period-ahead expected returns. It is based on a linear pricing rule generalized to a setting that allows for time variation in expected returns. We show that the firm-level one-period-ahead expected return is a linear combination of forward earnings yield, book-to-market ratio, and a variable summarizing one-period-ahead 'other information'. This 'other information' is implied by the firm's one-period-ahead earnings expectation and the current stock price.

The construction of our firm-level expected return measure can be described in three steps. First, we estimate a set of valuation parameters for a portfolio of homogeneous firms, including portfolio-level fundamental valuation multiples, average costs of capital and growth rates based on the forecasts of one-period ahead earnings and no-arbitrage pricing. Consistent with long-standing industry practice, we use industry portfolios to group homogeneous firms in our empirical implementation. Second, we use these portfolio-level parameters as approximate firm-specific valuation parameters to infer the value-relevant information contents of the implied 'other information' variable. Finally, we incorporate the portfolio-level parameters to firm-specific forward earnings yield, book-to-market ratio and the implied 'other information' to estimate expected one-period-ahead stock returns.

We validate our new measure of one-period-ahead expected returns in a US sample of up to 85,385 firm-year observations from 1985 to 2014. We first show that the new measure generates an *ex ante* monotonic decile ranking of one-period-ahead realized returns. An investor who follows a long-short portfolio strategy based on this decile ranking earns a 9% hedge returns per annum on average. The expected return estimates also show theoretically predictable or empirically consistent relations with a set of firm-level return predictive variables, including firm size, financial leverage, CAPM beta, net operating assets growth, accruals, sales growth and investment.

In multivariate regressions of realized returns, we find that, unlike many ICC estimates, the new expected return measure receives significantly positive predictive coefficients even after controlling for contemporaneous cash flow news, expected return news, and other return predictive variables. In addition, we find that the average cross-sectional return predictive coefficient on the new measure is statistically indistinguishable from one after controlling for contemporaneous cash flow news and expected return news, consistent with theoretical restriction regarding the validity of expected return measures (Easton and Monahan 2005).

This paper proceeds as follows. Section 2 discusses relations of the new measure of expected returns with existing models. Section 3 details the theoretical foundations and derives the new measure of expected return. Section 4 outlines the data management procedures and obtains empirical estimates of the expected returns. Section 5 presents portfolio-based validation tests of the association of the expected return measure with realized returns and firm characteristics. Section 6 tests the association of the new measure of expected return with future realized returns through pooled and cross-sectional regressions. Finally, Section 7 concludes the paper.

2 Existing models for measuring expected returns

In this section, we provide a detailed review of existing methods for measuring expected returns and compare them with the new approach developed in this study.

2.1 Factor models

Factor models are deeply rooted in asset pricing theory and are regarded as the 'orthodox' approach to measuring expected returns, although the main use of factor models is to *explain* realized returns through a small number of common risk factors, rather than to construct *ex ante* measures of expected returns.[1] Given a factor model, the construction of an expected return measure can be described in three steps. First, use time-series regressions of firms' past stock returns on past realizations of risk factors to obtain the firms' factor betas (i.e. the firm's exposure to the risk factors). Second, estimate the risk premiums associated with the risk factors using either extreme sample mean factor realizations or cross-sectional regressions.[2] Finally, multiply the estimated betas and risk premiums to obtain expected return estimates.[3]

However, it has been shown that primary factor models yield no meaningful predictive power for future realized returns. One common explanation for the

[1] It is well known that the CAPM as an one-factor model fails to describe cross-sectional returns and expected returns (Fama and French 1992, 1996). Although Merton (1973) intertemporal CAPM (ICAPM) and Ross (1976) arbitrage pricing theory (APT) provide theoretical motivations for multi-factor models, they do not supply a clear guidance for selecting factors. Thus, most multi-factor models are empirically motivated by evidence that some variables appear to predict returns, without clear economic interpretations (e.g. Fama and French 1992; Carhart 1997; Chen et al. 2011; Hou et al. 2015).

[2] If the risk factors are returns or excess returns, then the risk premiums are their expected values, which may be estimated by taking the sample means. Alternatively, cross-sectional regressions of sample average returns on the factor betas estimated from the first step can be used, so that the slope coefficients are regarded as the risk premium estimates. The latter approach does not restrict the risk premiums to equate the sample means of factor realizations.

[3] See Lyle and Wang (2015); Lee et al. (2015); Dittmar and Lundblad (2017) for applications of this procedure.

failure of factor models is the inevitable use of noisy *ex post* realized returns. Average realized returns generate poor measures of forward-looking risk premiums and factor betas (Fama and French 1997; Elton 1999). Average returns may deviate from expected returns substantially. For instance, a stock with a recent history of low returns may have incurred an upward revision in its expected return (i.e. positive discount rate news). Thus low average realized returns may be paired with high expected returns (Pástor et al. 2008). In addition, Penman and Zhu (2017) also find that beta estimates from lagged samples are considerably different from those estimated from forward samples, suggesting that past betas are not reliable proxies for future betas. The expected return is an *ex ante* concept. It is formed within investors' forward expectations conditional on the real-time information they observe. As the conditioning information set changes, expected returns and their determinants also change. While asset pricing theory admits time-varying conditioning information, empirical implementations of factor models typically assume that the factor risk premiums and firm risk exposure (the 'betas') estimated from past information sets carry forward to the future.

2.2 Implied cost of capital models

Proposed as a potential remedy, accounting-based implied cost of capital (ICC) models back out the expected returns from forward-looking valuation models such as the residual income valuation (RIV) model and Ohlson and Juettner-Nauroth (2005) abnormal earnings growth (AEG) model (Claus and Thomas 2001; Gebhardt et al. 2001; Easton et al. 2002; Gode and Mohanram 2003; Easton 2004; Pástor et al. 2008; Nekrasov and Ogneva 2011). ICC models bypass the difficulty in specifying the factor structure of returns and estimating the factor betas and premiums by exploiting the present value relation and solve for expected returns as a primitive valuation input.

While ICC models use forward-looking data and have the merit of being virtually free of assumptions about the processes generating accounting data,[4] they all implicitly assume a constant rate of expected return for all future horizons.[5] Therefore, ICC models are more likely to capture long-run discount rates rather than one-period-ahead expected returns. Given that expected returns predictably vary over time Cochrane (2009, 2011), the relation between expected return proxies under the constant expected return assumption and the true conditional expectations of next-period realized returns is necessarily distorted (Hughes et al. 2009). Pástor et al. (2008) also elaborate on this issue and suggest more forcefully that realized returns are likely to be negatively correlated with temporal changes in long-run expected returns, leaving the relation between ICC estimates and one-period realized returns to be weak. Indeed, Easton and Monahan (2005) and Lee et al. (2015) find that popular ICC estimates contain substantial measurement errors and thus are poorly associated with cross-sectional realized returns. The new approach proposed in this study does not assume constancy of expected returns over time. This important feature lends us confidence that our expected return estimate is designed specifically to capture expected one-period-ahead stock returns. Another important distinguishing feature of the new approach is that it does not rely on arbitrary assumptions about the terminal growth rates or dividend policies. Naturally, RIV or AEG models involve discounting an infinite series of future payoffs to equate the current stock price, but implementation of these models necessarily requires truncation at a given point in the future. Researchers must then supply some explicit assumptions about the growth in the payoff measures (e.g. abnormal earnings or abnormal earnings growth) post the truncation point. In addition, while the valuation models have the dividend irrelevance property, practical applications must still involve arbitrary assumptions about dividend policies to enable fore-

[4]RIV requires only no arbitrage condition and clean-surplus accounting to hold, and AEG holds even without clean surplus accounting.

[5]Since most implied cost of capital identify expected returns as the roots that satisfies the assumed valuation model for a given firm, expected returns are not identified without these assumptions.

casting future series of earnings and/or book values. Our approach summarizes valuation implications of future earnings growth and book values into the 'other information' term and estimates the evolution of this 'other information' term from the data, thus it avoids making these questionable assumptions.

2.3 Characteristics-based models

Recently, Lyle et al. (2013), Lewellen (2014), Lyle and Wang (2015) and Penman and Zhu (2017) propose a characteristics-based approach towards estimating expected returns. While motivated and implemented in different ways, these models estimate the predictive relations between firm characteristics and future returns from historical data, and apply the estimates to current characteristics to capture forward-looking expected returns. It appears that this approach has claimed some success in predicting the variation in realized returns (Lee et al. 2015). However, the choice of firm characteristics in the models are based almost purely on empirical exploration, without clear theoretical guidance.[6]

In addition, the implementations of firm characteristics-based models still rely on noisy realized returns. Thus, it is not a genuinely forward-looking approach.[7] Furthermore, it is often unclear *a priori* why the predictive power of firm characteristics for stock returns is due to risk rather than mispricing. In principle, it is plausible that firm characteristics proxy for firm-level exposure to common risk factors. For instance, Fama and French (1992) suggest that the book-to-market ratio predict returns because it proxies for the firm's factor betas. However, there has been no conclusive test of such conjectures. Admittedly, one may argue that expected returns are defined as returns that are *ex ante* predictable, and whether they are driven by mispricing is irrelevant. While this is a valid point if one views expected returns as a statistical concept and if the purpose is to build optimal trading strategies, this view does not speak to the economic

[6]Lyle et al. (2013) and Lyle and Wang (2015) are of the few exceptions.

[7]It appears that the characteristics-based expected returns perform better out-of-sample in comparison with factor models(Lewellen 2014).

question: *What is the market expectation?* In other words, the returns that investors expects to earn may deviate from optimal statistical forecasts. For research settings examining investors' information processing behavior, which is central to the accounting literature, the market expectation view is more important than the statistical view. The new approach introduced in the current study takes the former.

We also find similarities between the new approach introduced in this paper and some leading characteristics-based models. For instance, Penman and Zhu (2017) express expected returns as a linear function of earnings yield and book-to-market ratio, adjusted for a vector of variables that are empirically shown to predict earnings growth. The underlying argument is that variables that forecast earnings growth imply fundamental risk beyond current earnings and book values. In comparison, our model adjusts earnings yield and book-to-market for implied 'other information'. This 'other information' term is supposed to capture the net present values of business activities that are not recognized in the current accounting information but that will eventually feed back into future earnings and book values when they materialize. Thus, the implied other information term is naturally related to expected growth, consistent with the spirit of Penman and Zhu (2017).[8] In this vein, one can interpret the growth component in Penman and Zhu (2017) as a specialized parametric representation of our implied 'other information' term.

3 Theoretical development

Ohlson (1995) and Feltham and Ohlson (1995) show that the market price of equity can be expressed as a linear function of *current* firm fundamentals and

[8]This is also consistent with Ohlson (1995) original description of 'other information' in that any priced information must materialize in the form of incremental future abnormal earnings.

a term summarizing other value-relevant information. [9] Following Ashton and Wang (2013), we start with a simple version of the linear pricing rule:[10]

$$P_t = (1 + \alpha_1)b_t + \alpha_2 x_t + \alpha_1 d_t + \vartheta_t \tag{1}$$

where P_t is the market price of equity, b_t is the book value of equity, d_t is the total dividends, and x_t is the accounting earnings. The term ϑ_t denotes 'other information' that is priced by the market but not captured by the above accounting data items. ϑ_t follows the following process

$$\vartheta_{t+1} = \phi\vartheta_t + \epsilon_{\vartheta,t+1} \tag{2}$$

where the parameter ϕ is restricted by transversality condition such that it is allowed to be higher than the cost of capital for only some finite periods. With these structures in place, no-arbitrage condition allows one to relate prices at dates t and $t+1$ through the one period-ahead expected return ER_t:

$$ER_t \times P_t = \mathbb{E}_t[P_{t+1} + d_{t+1}] \tag{3}$$

After expanding the right-hand-side using the linear pricing rule and applying clean-surplus relation, one obtains

$$ER_t = (1 + \alpha_1 + \alpha_2)\frac{\mathbb{E}_t[x_{t+1}]}{P_t} + (1 + \alpha_1)\frac{b_t}{P_t} + \phi\frac{\vartheta_t}{P_t} \tag{4}$$

Equation (4) implies that one-period-ahead return is a linear function of book-to-market ratio, next-period earnings yield, and the price-deflated 'other information' term. To construct an expected return measure from this equation, one requires a set of parameter estimates (α_1, α_2 and ϕ) and a proxy for the 'other information'.

[9] In subsequent studies, this 'linear pricing rule' has been generalized to settings with stochastic interest rates (Gode and Ohlson 2004), risk-aversion (Feltham and Ohlson 1999; Ang and Liu 2001) and stochastic risk premia (Ang and Liu 2001; Lyle et al. 2013).

[10] Ang and Liu (2001) show that the linear form is preserved for a class of affine processes with any specification of the set of value-relevant firm fundamentals.

We pursue a forward-looking approach to obtaining the parameter estimates using the method developed in Ashton and Wang (2013). The estimation procedure can be summarized in three steps. First, We exploit the linear pricing rule and no arbitrage condition to estimate a set of common valuation parameters for a portfolio of homogeneous firms. Second, We approximate firm-level valuation parameters with estimates of their respective industry average to estimate the implied 'other information'. Finally, We combine the valuation parameter estimates with forward earnings yield, book-to-price ratio and the implied 'other information' to construct the proxy for one-period-ahead expected return.

While other information ϑ_t is unobservable, it can be substituted with the linear pricing rule $\vartheta_t = P_t - (1 + \alpha_1)b_t - \alpha_2 x_t - \alpha_1 d_t$, after which equation (4) can be re-expressed as an earnings forecast model

$$\begin{aligned}\mathbb{E}_t[x_{t+1}] = &\frac{ER_t - \phi}{1 + \alpha_1 + \alpha_2} P_t + \frac{\phi(\alpha_1 + \alpha_2)}{1 + \alpha_1 + \alpha_2} x_t \\ &+ \frac{\phi - \alpha_1 - 1}{1 + \alpha_1 + \alpha_2} b_t + \frac{\phi \alpha_1}{1 + \alpha_1 + \alpha_2} b_{t-1}\end{aligned} \tag{5}$$

Equation (5) captures the notion that prices lead earnings. That is, prices forecast earnings beyond information reflected in realized earnings and book values. If there exists a reasonable proxy for market earnings expectation $\mathbb{E}_t[x_{t+1}]$, equation (5) is an exactly identified model. In Hansen (1982) terms, there are four 'instruments' ($E_t[x_{t+1}]$, x_t, b_t and b_{t-1}) and four 'parameters' (α_1, α_2, ϕ and ER_t). Surprisingly, ER_t is directly identified in the model. In fact, Ashton and Wang (2013) run annual cross-sectional estimations of this model to obtain aggregate cost of capital estimates in a constant expected return setting. However, at the firm level, estimating ER_t directly from this equation is not feasible. Cross-sectional estimation will produce the same expected return for all firms in a given year; similarly, time-series estimation for each firm will produce an estimate without time variation. As a practical compromise, we estimate cross-sectional average valuation parameters α_1, α_2 and ϕ for portfolios of homogeneous firms on a yearly basis and use these portfolio-level estimates to approximate firm-level parameter values.[11] Specifically, We estimate the following annual cross-sectional regression model for each industry partition:

$$\begin{aligned}\frac{\mathbb{E}_t[x^i_{t+1}]}{P^i_t} &= \frac{\overline{ER}_t - \overline{\phi}}{1+\overline{\alpha}_1+\overline{\alpha}_2} + \frac{\overline{\phi}(\overline{\alpha}_1+\overline{\alpha}_2)}{\overline{1+\alpha}_1+\overline{\alpha}_2}\frac{x^i_t}{P^i_t} \\ &+ \frac{\overline{\phi}-\overline{\alpha}_1-1}{1+\overline{\alpha}_1+\overline{\alpha}_2}\frac{b^i_t}{P^i_t} + \frac{\overline{\phi}\overline{\alpha}_1}{\overline{\alpha}_1+\overline{\alpha}_2}\frac{b^i_{t-1}}{P^i_t} + \varepsilon^i_t \end{aligned} \tag{6}$$

Now the over-lines and superscripts i are used to denote industry-level parameters and firm-level variables respectively. Note that while equation (5) holds exactly for each firm i and date t, equation (6) involves a firm-specific error term ε^i_t to the extent that industry average parameter values deviate from their firm-level counterparts. The scaling by price P_t ensures that the variables used for estimation are ergodic stationary to allow for statistical evaluations based on asymptotics (see, for instance, Hayashi 2000). While using any proxy for $\mathbb{E}_t[x_{t+1}]$ is bound to introduce measurement errors[12], estimates obtained from estimating (6) are still consistent because measurement errors do not affect the right-hand-side variables.

With a set of parameter estimates for $\overline{\alpha}_1$, $\overline{\alpha}_2$ and $\overline{\phi}$, We next explicitly compute firm-level implied 'other information' as:

$$\widehat{\vartheta}^i_t = P^i_t - (1+\overline{\alpha}_1)b^i_t - \overline{\alpha}_2 x^i_t - \overline{\alpha}_1 d^i_t \tag{7}$$

Finally, We construct the firm-level expected return by combining industry average parameter estimates and firm-level forward earnings yield, book-to-price ratio, and the implied 'other information':

$$\widehat{ER}^i_t = (1+\overline{\alpha}_1+\overline{\alpha}_2)\frac{\mathbb{E}_t[x^i_{t+1}]}{P^i_t} + (1+\overline{\alpha}_1)\frac{b^i_t}{P^i_t} + \overline{\phi}\frac{\widehat{\vartheta}^i_t}{P^i_t} \tag{8}$$

[11]Implicitly, the estimation procedures used for most characteristics-based models such as Penman and Zhu (2017) and Lyle et al. (2013) are of the same spirit. That is, they assume the characteristic loadings are cross-sectional constants. Thus the cross-sectional variation derives from firm-specific characteristics.

[12]The measurement error comes from the fact that any proxy used may deviate from the true market expectation of the firm's next-period earnings.

Note that only information available at date t are used to construct expected return ER_t, thus the look-ahead bias is avoided. One notable advantage of our approach is that it requires only one-year-ahead earnings forecast, thus imposing less data requirement. By contrast, typical ICC models require earnings forecasts for two to four years out. When these forecasts are unavailable, prior studies typically drop the respective observations or impute the forecasts from other data. Such difficulties do not constrain the model described above. In addition, this model does not require earnings forecasts or earnings growth to be positive, whereas many ICC models based on earnings growth are inapplicable to loss firms and firms with negative expected growth.

4 Data and estimates

To estimate the new measure of one-year-ahead expected return, we obtain accounting data from Compustat Fundamentals Annual file. Stock return data are extracted from CRSP Monthly Stock File (MSF). Analyst forecasts, stock prices and numbers of shares outstanding are collected from IBES. We construct a base sample of all firms listed in NYSE, Amex and Nasdaq identified at the intersection between Compustat, CRSP and IBES from 1985 to 2014. We exclude all non-equity issues such as ADRs as indicated by CRSP share code (SHRCD). Table 1 outlines the sample construction procedure.

Estimation of expected returns require earnings (x_t), dividends (d_t) current and lagged book values of equity (b_t and b_{t-1}), and one-year-ahead analyst earnings forecasts ($FE_{t,t+1}$). All variables are translated to per-share basis as per IBES number of shares outstanding unless otherwise indicated. Consistent with Penman and Zhu (2014), earnings are calculated as earnings before extraordinary items (Compustat item IB) plus special items (SPI), minus preferred dividends (DVP), with a tax allocation to special items at the statutory income tax rate for the year. Book value of equity is calculated as Compustat item

Table 1: Construction of estimation sample

Data restrictions	Observations	Firms
Compustat data with non-missing x_t, d_t, $b_t > 0$ and $b_{t-1} > 0$	147,674	16,866
Less observations with newly acquired assets greater than 30% of total assets	(972)	(54)
Less observations with missing one-year-ahead analyst forecasts, share price and number of shares outstanding from IBES	(42,263)	(3,743)
Less observations without SIC codes from Compustat or not classified under Fama-French five-industry classification scheme	(15,902)	0
Less observations with b_t/P_t not in the range $(0.01, 100)$, x_t/P_t not in the range of $(-1, 1)$, and $P_t < 0.5$	(152)	(1,912)
Estimation sample	85,385	11,157

Notes: This table reports the effects sample selection restrictions on the estimation sample. The first column describes data restrictions applied. The second and third columns reports their effects on the sample size in terms of total number of observations and the number of unique firms. Values in parentheses indicate reductions.

CEQ, plus preferred treasury stock (TSTKP) and minus preferred dividends in arrears (DVPA). Dividends are measured as Compustat item DVT.[13] We exclude firm-year observations with negative book values or lagged book values and those heavily impacted by mergers and acquisitions activities as indicated if the total amount of acquired assets is estimated to be grater than 30% of the beginning balance of total assets.[14]

As a proxy for market expectation of one-year-ahead earnings ($\mathbb{E}_t[x_{t+1}]$), $FE_{t,t+1}$ is measured as the median analyst forecast (IBES item MEDEST) formed on the third Thursday of April following year t financial year end from the IBES Summary History file.[15] Stock price (IBES item PRICE) and the number of shares outstanding (SHOUT) are based on IBES Pricing & Ancillary file, and they are observed on the same date as $FE_{t,t+1}$.[16]

We accumulate one-year-ahead buy-and-hold stock returns R_{t+1} from May of year $t+1$ to April of year $t+2$. We follow Shumway (1997) recommendation and apply -33% adjustment for delisting returns. The estimation is performed over five industries each year, using Fama and French (1997) industry classification scheme. SIC codes (Compustat item SICH) are used to perform the classifications. Risk-free rates r_t^f are measured as the yield of 10-year treasury bonds, observed at the analyst forecast date. Following prior studies, We delete observations with book-to-price ratio (b_t/P_t) lower than 0.01 or higher than 100, earnings yield (x_t/P_t) lower than -1 or higher than 1, or stock price lower than \$0.5 to mitigate effects of extreme values on estimations. After applying these restrictions, the estimation sample consists of $85,385$ firm-year observations with $11,157$ unique firms. This sample size is significantly larger than

[13]This measurement choice avoids share-based transactions.

[14]We estimate the total amount of acquired assets as the sum of acquired inventory (ACQINVT), acquired property, plant and equipment (ACQPPE), acquired intangible assets (ACQINTAN), acquired goodwill (ACQGDWL), and acquired other assets (ACQAO). These data items are obtained from Compustat.

[15]IBES typically compiles summary files for each firm on the third Thursday of each month.

[16]Because the items SHOUT and PRICE are not well populated before 1985 in the IBES version that we have access to, the sample period starts from 1985.

those used in prior ICC studies that use analyst forecasts (e.g. Nekrasov and Ogneva 2011).

Table 2 reports the details of annual sample distributions across the five Fama-French industries. Over the 30-year sample period, the number of firms ranges from 2,211 in 1985 to 3,958 in 1997. Across the industries, We obtain a typical sample size of around 600 firms per year except for the Hlth industry, where the number of firms per year grew from only 120 in 1985 to 326 in 2014, peaking at 399 in 1997. Thus, the parameter estimates for Hlth industry may be less precise than estimates for other industries, especially in early years of the sample period. Overall, Table 2 shows that the industry-year partitions produce reasonably well populated samples for estimation.

In the following sections, we will evaluate the empirical performance of the new measure of expected return with a battery of validation tests. These tests will require additional data including future realized one-period-ahead returns and a set of firm characteristics that have been shown to predict stock returns empirically. This reduces the size of the testing sample to 80,940.

Table 3 Panel A provides sample descriptive statistics of the variables used in estimating equation (6) for constructing expected returns. The mean (median) of book-to-market ratio b_t/P_t is 0.62 (0.52), comparable to the sample means (medians) reported in prior studies (Nekrasov and Ogneva 2011). The forward earnings yield $FE_{t,t+1}/P_t$ is higher than trailing earnings yield x_t/P_t at all reported percentiles of their distributions, consistent with analysts' tendency to issue optimistic forecasts.

Table 3 Panel B reports descriptive statistics of the testing sample. Note that the empirical distributions of b_t/P_t, x_t/P_t and $FE_{t,t+1}/P_t$ remain largely unchanged compared to the estimation sample, except for a notable rise in lower tail of the distribution of x_t/P_t (with 1st percentile increased from -0.85 to

Table 2: Sample distribution by year and industry classification

Year	Cnsmr	Manuf	HiTec	Hlth	Other	Total
1985	481	646	412	120	552	2,211
1986	512	638	438	132	558	2,278
1987	504	607	448	153	602	2,314
1988	517	600	451	165	615	2,348
1989	503	601	438	173	652	2,367
1990	506	593	410	181	627	2,317
1991	487	622	413	198	606	2,326
1992	554	638	428	264	633	2,517
1993	639	665	499	297	726	2,826
1994	706	715	589	318	979	3,307
1995	730	776	648	309	1,030	3,493
1996	719	803	781	333	1,090	3,726
1997	743	798	844	399	1,174	3,958
1998	710	759	829	389	1,158	3,845
1999	654	687	786	338	1,092	3,557
2000	563	607	863	304	992	3,329
2001	478	555	800	328	863	3,024
2002	457	531	697	317	852	2,854
2003	469	509	682	344	860	2,864
2004	481	517	703	353	877	2,931
2005	471	521	689	372	918	2,971
2006	483	540	646	369	920	2,958
2007	465	556	630	379	876	2,906
2008	441	552	603	347	819	2,762
2009	451	559	590	329	872	2,801
2010	438	555	560	315	781	2,649
2011	436	525	519	299	748	2,527
2012	409	510	523	297	749	2,488
2013	403	496	538	287	739	2,463
2014	384	486	528	326	744	2,468
Total	15,794	18,167	17,985	8,735	24,704	85,385

Notes: This table reports the distribution of sample observations across Fama and French (1997) five industry portfolios over the sample period 1985-2014. Industry 'Cnsmr' includes consumer durables, non-durables, wholesale, retail, and some services (e.g. laundries, repair shops). Industry 'Manuf ' includes manufacturing, energy, and utilities. Industry 'HiTec' includes business equipment, telephone and television transmission. Industry 'Hlth' includes health care, medical equipment, and drugs. Industry definitions are obtained from Keneth French's online data library `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

-0.69), suggesting extreme loss makers fail to survive the additional data requirement. Average one-period-ahead buy-and-hold return $R_{t,t+1}$ is 12% with a large standard deviation of 48%.

Panel B also provides summary statistics of some commonly used risk proxies. Financial leverage Lev_t, calculated as total financial debt divided by the book value of equity, is highly right-skewed with its mean (0.635) far exceeding the median (0.215).[17] Firm size $Mcap_t$ is measured as the natural log of market capitalization on the analyst forecast date.[18] The log transformation of $Mcap_t$ smooths out the high skewness in firm size, with the mean and median both take values just above 6. The mean CAPM beta $Beta_t$ is 1.14, slightly higher than the market average beta of 1, probably due to the fact that firms receiving analyst coverage exhibit stronger co-movement with the market (Lee and So 2017).[19] Momentum Mom_t is defined as the the buy-and-hold stock return over the 12-month period prior to the forecast date.

Prior studies also show that growth is related to risk and expected returns(Penman and Zhu 2014, 2017), so Panel B also summarizes some leading growth-related variables. Operating accrual Acc_t receives negative mean and median, consistent with prior findings that accruals tend to bear bad earnings news (Sloan 1996; Givoly and Hayn 2000). Net operating asset growth ΔNOA_t, percentage sales growth rate $Saleg_t$, investments in property, plant & equipment Inv_t and net external financing Exf_t all have positive mean and medians, indicating that firms typically experience positive growth.[20] Overall, the summary statistics in

[17]Total financial debt is estimated by the sum of debt in current liabilities (Compustat item DLC) and debt in long-term liabilities (item DLT). Note that the extremely large 99th percentile is not due to a few extreme cases, as the the distribution of Lev_t between its 75th and 99th percentiles are indeed quite wide.

[18]Market capitalization is calculated as share price (PRICE) multiplied by number of shares outstanding (SHOUT) obtained from IBES.

[19]We estimate CAPM beta $Beta_t$ using at least 18 and up to 60 monthly stock returns prior to the earnings forecast date. Note that the market return used for estimating $Beta_t$ is CRSP value-weighted return, thus the mean beta in our sample is not necessarily one.

[20]Acc_t is measured as the sum of changes in receivables (Compustat item RECT), inventory (INVT) and other current assets (ACO), less depreciation and amortization charges (DP) and changes in other current liabilities (LCO). ΔNOA_t is calculated as the total changes in receivables (RECT), inventory (INVT), other current assets (ACO), property, plant and equipment (PPENB), intangible assets (INTAN) and other long-term assets (AO), less changes

Table 3 show that our sample is largely comparable with prior studies.

Model parameters in equation (6) are obtained from cross-sectional estimations. [21] Standard errors of parameter estimates can be computed by applying the delta method. Slutsky's theorem ensures that consistent estimation of the linear coefficients will lead to consistent estimation of the model parameters (Hayashi 2000). However, to avoid reparameterization and complicated delta method computations, We use the numerically equivalent equal-weighting one-step generalized method of moments (GMM) to estimate (6), which estimates the non-linear model parameters and provides their statistical properties directly.[22]

Table 4 presents annual GMM estimates of parameters in equation (6). The second, fourth and fifth columns report the time series of cross-sectional average expected returns $\overline{ER}_t$, expected risk premia $\overline{ERP}_t$ and risk-free rates r_t^f. Figure 1 plots these data over time. The time-series average expected return is 8.57% and the average risk premium is just under 3%, which is quite similar to the those reported in Claus and Thomas (2001), Gebhardt et al. (2001), Easton et al. (2002), Nekrasov and Ogneva (2011) and Ashton and Wang (2013). All estimates of expected returns are highly significant with annual t-statistics averaging at 19.50. Expected return $\overline{ER}_t$ is generally declining over the sample period from around 10% to 7-8%, mainly due to the apparent decline in the 10-year treasury

in payables (AP), other current and long-term liabilities (LCO + LO). Inv_t is the sum of increments in the gross costs of property, plant and equipment and in inventory. Accr_t, ΔNOA_t and Invest_t are scaled by average assets of year t. Sgr_t is simply year-t change in sales (SALE) divided by sales of prior year. Exf_t equals the cash proceeds from long term debt issues (item DLTIS) and equity issues (item SSTK) plus the net changes in current debt (item DLCCH) less cash payments for retiring long term debts (item DLTR), for equity share repurchases (item PRSKC) and for dividends (item CDVC).

[21] To estimate equation (6), one can reparameterize it into a linear regression and estimate its linear coefficients with ordinary least squares (OLS), and then reverse the reparameterization to uncover the underlying parameters $\overline{ER}_t$, $\overline{\phi}$, $\overline{\alpha}_1$ and $\overline{\alpha}_2$ (as in Ashton and Wang 2013).

[22] All subsequent analyses are replicated with OLS, and the results are changed only due to negligible rounding errors. Using the standard two-step spectral density weighted GMM also produces almost identical results, and it is asymptotically more efficient (Hayashi 2000). However, we do not proceed with this two-step estimation because the size of a typical industry-year sub-sample is in the order of hundreds, meaning that the spectral density matrix estimates are likely to be unreliable (see, for example Ferson and Foerster 1994; Hansen et al. 1996; Cochrane 1996). Hence, one-step estimates are likely more robust in this case.

Table 3: Summary statistics

Variable	Mean	SD	P_1	P_{25}	Median	P_{75}	P_{99}
Panel A: Estimation sample $N = 85{,}385$							
b_t/P_t	0.620	0.490	0.071	0.318	0.520	0.781	2.357
x_t/P_t	0.002	0.267	−0.848	0.011	0.048	0.073	0.188
$FE_{t,t+1}/P_t$	0.051	0.098	−0.311	0.040	0.063	0.086	0.182
b_{t-1}/P_t	0.627	0.659	0.054	0.284	0.488	0.766	2.930
Panel B: Testing sample $N = 80{,}940$							
b_t/P_t	0.619	0.467	0.076	0.324	0.525	0.782	2.260
x_t/P_t	0.012	0.222	−0.689	0.016	0.050	0.075	0.187
$FE_{t,t+1}/P_t$	0.056	0.081	−0.247	0.043	0.065	0.087	0.182
$R_{t,t+1}$	0.121	0.480	−0.750	−0.179	0.074	0.339	1.788
Risk proxies							
Lev_t	0.612	1.680	0.000	0.032	0.215	0.635	6.134
$Mcap_t$	6.160	1.853	2.566	4.808	6.021	7.361	10.968
$Beta_t$	1.144	0.771	−0.219	0.637	1.042	1.502	3.702
Mom_t	0.187	0.672	−0.750	−0.158	0.096	0.381	2.480
Growth proxies							
Acc_t	−0.027	0.094	−0.275	−0.068	−0.030	0.009	0.259
ΔNOA_t	0.032	0.142	−0.306	−0.027	0.014	0.070	0.564
$Saleg_t$	0.117	0.334	−0.688	0.004	0.091	0.207	1.102
Inv_t	0.079	0.212	−0.224	0.001	0.038	0.106	0.763
Exf_t	0.055	0.167	−0.161	−0.006	0.008	0.060	0.795

Notes: Panel A and Panel B report descriptive statistics for the estimation sample and testing sample respectively. Sample means, standard deviations, 1th, 25th, 50th, 75th and 99th percentiles are reported in columns 2-8 respectively. b_t/P_t is the book-to-price ratio. x_t/P_t and $FE_{t,t+1}/P_t$ are trailing and forward earnings yields respectively. b_{t-1}/P_t is the ratio of lagged book value per share over current stock price. Lev_t is financial leverage; $Mcap_t$ is the natural log of market capitalization; Acc_t is operating accruals; $Beta_t$ is CAPM beta; ΔNOA_t is net operating asset growth; $Saleg_t$ is realized (percentage) sales growth rate; Inv_t is firm investments in property, plant & equipment and inventory; Exf_t is total external financing; Mom_t is momentum (past 12-month buy-and-hold return); and $R_{t,t+1}$ is one-period-ahead realized return.

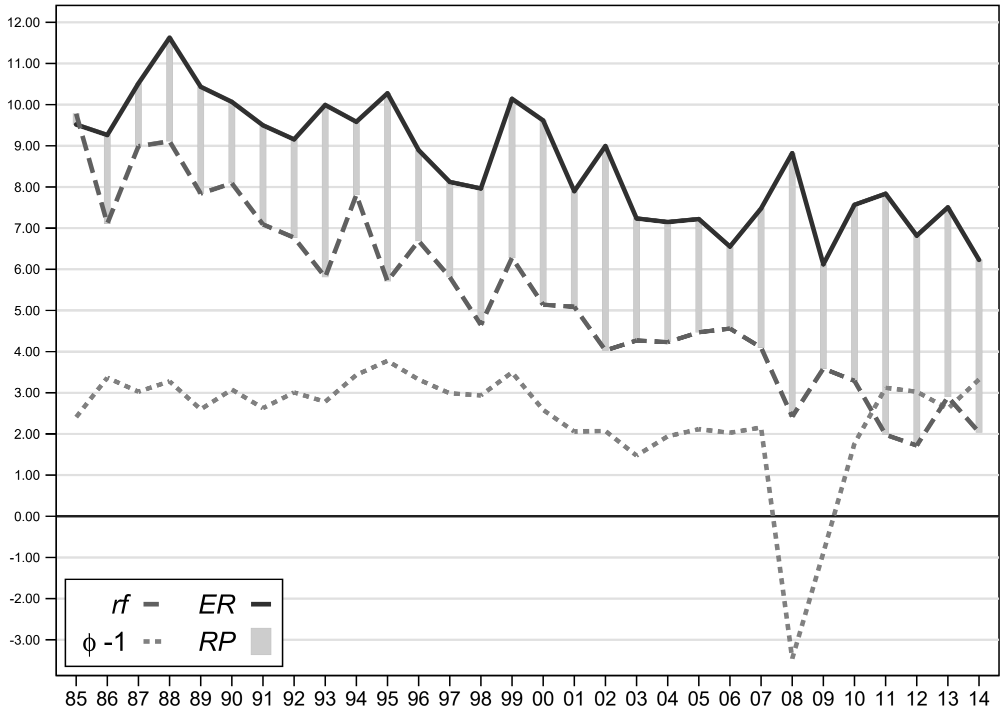
Table 4: Annual cross-sectional estimates of average valuation parameters

$$\frac{\mathbb{E}_t[x^i_{t+1}]}{P^i_t} = \frac{\overline{ER}_t - \overline{\phi}}{\overline{\alpha}_1+\overline{\alpha}_2} + \frac{\overline{\phi}(\overline{\alpha}_1+\overline{\alpha}_2-1)}{\overline{\alpha}_1+\overline{\alpha}_2}\frac{x^i_t}{P^i_t} + \frac{\phi-\overline{\alpha}_1}{\overline{\alpha}_1+\overline{\alpha}_2}\frac{b^i_t}{P^i_t} + \frac{\overline{\phi}\overline{\alpha}_1-1}{\overline{\alpha}_1+\overline{\alpha}_2}\frac{b^i_{t-1}}{P^i_t} + \varepsilon^i_t$$

Year	ER	$t(ER)$	RP	r^f_t	$\phi-1$	$t(\phi-1)$	α_1-1	$t(\alpha_1)$	α_2	$t(\alpha_2)$
1985	9.51	28.61	0.26	9.26	2.41	2.07	0.01	0.59	0.16	3.23
1986	9.26	21.35	2.15	7.11	3.37	5.36	0.02	0.97	0.21	3.40
1987	10.51	24.54	1.53	8.99	3.03	3.73	0.05	1.98	0.22	3.92
1988	11.62	37.97	2.52	9.11	3.28	3.85	0.03	1.54	0.40	4.84
1989	10.43	41.68	2.60	7.84	2.60	3.30	0.06	3.50	0.15	2.71
1990	10.06	37.31	1.35	8.08	3.08	2.99	0.03	2.27	0.22	4.48
1991	9.49	30.51	2.29	7.09	2.63	4.62	0.04	2.69	0.19	3.94
1992	9.15	24.83	2.39	6.77	3.01	4.16	0.03	1.52	0.18	3.03
1993	9.99	29.76	3.71	5.77	2.78	4.54	0.06	2.44	0.28	5.12
1994	9.58	35.28	2.11	7.81	3.44	6.00	0.01	0.67	0.28	4.84
1995	10.27	36.02	4.11	5.71	3.78	5.03	0.02	2.02	0.28	5.70
1996	8.89	29.14	2.44	6.30	3.32	7.28	0.04	2.31	0.18	3.77
1997	8.12	19.71	2.31	5.81	2.97	3.68	0.02	1.30	0.31	4.62
1998	7.96	16.61	3.31	4.65	2.94	4.07	0.03	0.69	0.25	4.34
1999	10.14	28.50	3.86	6.28	3.50	3.70	0.09	3.26	0.37	5.28
2000	9.61	12.49	4.03	5.24	2.58	2.61	0.00	-0.01	0.11	2.87
2001	7.89	15.58	2.76	5.09	2.06	2.05	0.06	1.06	0.18	4.22
2002	8.99	10.39	4.97	4.03	1.39	1.50	0.01	0.65	0.27	5.80
2003	7.23	13.94	2.97	4.27	1.47	2.43	0.02	0.46	0.34	5.13
2004	7.14	8.77	2.92	4.23	1.94	1.30	0.01	0.49	0.56	4.83
2005	7.22	14.35	3.04	4.47	2.12	1.56	0.04	1.80	0.53	5.02
2006	6.55	12.19	1.99	4.56	1.76	1.29	0.02	0.57	0.52	4.28
2007	7.47	9.37	3.38	4.10	2.16	1.44	0.01	0.41	0.30	4.39
2008	8.82	5.17	6.40	2.42	-3.46	-1.46	0.00	-0.01	0.15	3.75
2009	6.12	8.78	2.53	3.59	-1.66	-1.71	0.01	0.33	0.33	3.60
2010	7.46	16.94	4.18	3.29	1.77	1.56	-0.01	-0.34	0.43	4.97
2011	7.83	16.28	5.86	1.98	3.12	3.15	0.01	0.23	0.49	4.74
2012	6.81	9.44	5.10	1.72	3.03	1.84	0.04	1.45	0.33	3.87
2013	6.89	9.90	4.42	2.90	2.61	1.96	0.02	0.53	0.57	4.27
2014	7.61	7.70	5.08	2.21	3.32	2.04	0.15	2.70	0.70	4.61
Mean	8.57	19.50	2.96	5.24	2.63	2.99	0.03	1.18	0.27	4.28

Notes: This table reports cross-sectional average estimates from annual cross-sectional GMM estimations of equation (6). Column ER reports annual average expected return estimate $\overline{ER}_t$; column $t(ER)$ is the associated t-statistic of $\overline{ER}_t$. RP and r^f_t indicate expected risk premium and risk-free rate respectively, where RP is the difference between ER and r^f_t and r^f_t is the yield of 10-year treasury bond. Columns $\phi-1$ and $t(\phi)$ report estimates and t-statistics of $\overline{\phi}-1$. Columns α_1 and α_2 ($t(\alpha_1)$ and $t(\alpha_2)$) report estimates (t statistics) of $\overline{\alpha}_1$ and $\overline{\alpha}_2$ respectively Values in columns ER, RP, r^f_t and $\phi-1$ are reported in percentage scale, and values in remaining columns are reported in raw scales.

Figure 1: Time-series plot of expected returns and implied other information growth



yield r_t^f, leaving the risk premium mostly stationary. Notably, the risk premium $\overline{ERP}_t$ hiked after the 1987 black Monday, 1990 recession, 2001 'tech bubble' burst and the 2007-2008 sub-prime debt crisis. This seems to be consistent with the notion that risk premiums tend to be high when the overall market and economy suffer.

Table 4 and Figure 1 also track the estimate $\overline{\phi} - 1$ over time. We find that all estimates of $\overline{\phi} - 1$ are positive and around 3% except for the years 2008 and 2009, when it sharply dropped to -3.5% in the middle of the most recent global financial crisis.[23] Ashton and Wang (2013) interpret the quantity $\overline{\phi} - 1$ as the growth rate in 'other information', which is likely to capture the effect of macroeconomic growth on asset prices. Indeed, this growth rate averages around 2.6%, which is in the order of typical US GDP growth over the sample period. It bottomed up in 2008 but rebounded back to around 4% after the

[23]Note that this finding shows that other information in stock prices may not converge to zero in the long run, perhaps due to the existence of new investment opportunities and technology advances. However, it can been seen that $\overline{\phi}$ is smaller than the risk free rates except for 3 years after GFC, consistent with the cointegration requirement

GFC, signaling strong signs of economic recovery. These observations seem to validate the Ashton and Wang (2013) interpretation.

The last four columns in Table 4 provide point estimates and t-statistics of the linear pricing parameters. Most estimates of $\overline{\alpha}_1$ are very close to one, with $\overline{\alpha}_1 - 1$ mostly positive but insignificantly different from zero in all but seven years in the sample. All estimates of $\overline{\alpha}_2$ are significantly positive, consistent with earnings positively contributes to market price after controlling for book values, dividends and other information.

In the next two sections, We perform external validation tests to evaluate the quality of the new expected return measure at the portfolio and firm level.

5 Portfolio returns and associations with firm characteristics

Our first validation test is based on portfolio-level associations of the new expected return measure with future realized returns and commonly used risk and growth proxies.

An important question in evaluating the usefulness of an expected return measure is whether an investor can expect to earn economically meaningful gains if she forms portfolios based on the measure. Thus, it is important to test whether a long-short strategy based on the expected return measure is associated with sizable out-of-sample gains. This approach is also useful in mitigating the measurement errors in the expected return measure as firm-level measurement errors are 'smoothed out' at the portfolio level (Easton and Monahan 2005). Table 5 reports time-series averages of future realized returns on cross-sectional expected-return-sorted decile portfolios over various investment horizons. Over one-year horizon, the new expected return measure ER_t monotonically sorts realized returns except for the highest (10th) decile. An investor who follows a zero-cost strategy going long on the 10th and short on the 1st decile portfolios is expected to earn a statistically significant and economically large 9.87% (t-statistic 11.61) return over the next year. Thus, the new measure significantly predicts one-period-ahead realized return in the cross-section.

Table 5: Future realized returns on cross-sectional expected-return-sorted portfolios

Deciles	ER_t	$R_{t,t+1}$	$R_{t,t+2}$	$R_{t,t+3}$	$R_{t,t+4}$	$R_{t,t+5}$	$R_{t,t+10}$
1 (low)	.0153	.0517	.2096	.3593	.4860	.6534	2.190
2	.0508	.1115	.2264	.4096	.5257	.7015	2.276
3	.0653	.1121	.2333	.3989	.5417	.7421	2.165
4	.0758	.1135	.2534	.4208	.5941	.7866	2.328
5	.0846	.1283	.2762	.4424	.6370	.8161	2.203
6	.0929	.1425	.2905	.4493	.6257	.8422	2.170
7	.1021	.1459	.3003	.4746	.6904	.9381	2.372
8	.1144	.1539	.3025	.4867	.6971	.9138	2.391
9	.1356	.1566	.3332	.5473	.8168	1.116	3.170
10 (high)	.2059	.1504	.3429	.5799	.8723	1.144	3.028
High-low return		**.0986**	**.1333**	**.2205**	**.3863**	**.4905**	**.8380**
t-statistics		11.61	8.04	8.16	10.86	10.06	3.58

Notes: This table reports time-series average future realized returns on cross-sectional expected-return-sorted decile portfolios over 1-, 2-, 3-, 4-, 5- and 10-year horizons. All portfolios are formed on equally weighted basis at date t. $R_{t,t+j}, j = 1, 2, 3, 4, 5, 10$ is the realized return from date t to date $t + j$. The second last row reports returns to buying the 10th and selling the 1st decile portfolios over the respective horizons, and the last row reports their respective t-statistics.

The ability of ER_t to rank future realized returns remains highly significant up to 10 years out, and the statistical significance of the long-short strategy is not seriously impaired up to five years out from the date of portfolio formation. These results shows that while ER_t is designed to measure one-period-ahead expected returns, it is persistent enough to forecast economically large cross-sectional spreads in realized return up to at least 10 years out-of-sample.

We next examine the relation between the new expected return measure and a range of firm characteristics Botosan and Plumlee (2005) and Botosan et al. (2011) suggest that a reasonable expected return measure should be well associated with firm characteristics that serve as 'risk proxies'. While this argument is intuitive, the identity of 'risk proxies' is controversial, which leads Easton and

Monahan (2016) to caution against this approach to assessing the validity of expected return measures. Table 6 examines the relation between expected returns and firm characteristics on the portfolio basis. The first five reported risk proxies include book-to-market ratio (b_t/P_t), firm size ($Mcap_t$), leverage (Lev_t), CAPM beta ($Beta_t$) and 'momentum' (Mom_t, proxied by the buy-and-hold return over the past 12 months).[24] The results show that firms with higher expected returns tend to have higher book-to-market ratios, smaller market capitalizations, and higher financial leverage, consistent with their univariate associations of these variables with realized returns, although the ranking is not monotonic. Momentum appears to be decreasing in expected return from the 2nd to the 9th deciles, roughly consistent with reversal patterns of stock returns over 12-month or longer horizons (Jegadeesh and Titman 1993). CAPM beta, however, appears to exhibit a U-shaped relation with expected return and the 1st decile portfolio has higher betas than the 10th portfolio, which is inconsistent its theoretical relation with expected returns.

Table 6 also reports the means of four 'growth proxies' across expected return deciles. These variables have been shown to negatively predict future returns (Lakonishok et al. 1994; Sloan 1996; Fairfield et al. 2003; Zhang 2007; Penman and Zhu 2014, 2017). The patterns again appear to be mixed and puzzling. While net operating asset growth ΔNOA_t is decreasing in expected returns from the 2nd to 10th deciles, accruals Acc_t, investments Inv_t and external financing Exf_t are not associated with expected returns in the manner consistent with their empirical predictive relation with realized returns.

One potential explanation for the mixed findings reported in Table 6 is the lack of power of univariate tests. Thus, we perform multivariate analysis by esti-

[24] While these variables are commonly used in empirical asset pricing applications, their identity as risk proxies is nonetheless far from robust. b_t/P_t has been shown to correlate with analyst forecast errors (Piotroski and So 2012); $Mcap_t$ has been shown to have lost its power to predict returns (Van Dijk 2011); Lev_t has been shown to negatively predict stock returns in several studies (e.g. Chava and Purnanandam 2010); $Beta_t$ is notorious for its poor association with stock returns despite its deep theoretical appeal (e.g. Frazzini and Pedersen 2014); and Mom_t predicts future stock return reversals, inconsistent with any risk-based explanation (Jegadeesh and Titman 1993).

Table 6: Firm characteristics for cross-sectional expected-return-sorted portfolios

Deciles	b_t/P_t	$Mcap_t$	Lev_t	$Beta_t$	Mom_t	ΔNOA_t	Acc_t	Inv_t	Exf_t
1	.757	5.31	.629	1.52	.080	.0207	-.0491	.0777	.142
2	.577	6.09	.407	1.29	.262	.0412	-.0287	.0975	.0726
3	.534	6.37	.374	1.17	.241	.0388	-.0289	.0952	.056
4	.544	6.44	.427	1.08	.245	.0323	-.0292	.0841	.0416
5	.566	6.43	.470	1.03	.219	.0326	-.0242	.0787	.0365
6	.592	6.41	.559	.997	.218	.0285	-.0226	.0706	.0333
7	.623	6.22	.631	1.01	.198	.0292	-.0192	.0714	.0343
8	.660	6.15	.767	1.05	.195	.0343	-.0173	.0772	.0399
9	.705	5.89	.934	1.13	.160	.0356	-.0145	.0842	.0524
10	.732	5.55	1.11	1.20	.079	.0294	-.023	.0856	.0797

Notes: This table reports time-series average future realized returns on cross-sectional expected-return-sorted decile portfolios over 1-, 2-, 3-, 4-, 5- and 10-year horizons. All portfolios are formed on equally weighted basis at date t. $R_{t,t+j}, j = 1, 2, 3, 4, 5, 10$ is the realized return from date t to date $t + j$. The second last row reports returns to buying the 10th and selling the 1st decile portfolios over the respective horizons, and the last row reports their respective t-statistics.

mating the cross-sectional relationship between expected returns and the firm characteristics. We control for forward earnings yield in these regressions. The results are reported in Table 7. The most striking finding in Table 7 is that ER_t is significantly negatively associated with b_t/P_t after controlling the forward earnings yield. This is consistent with Ohlson (2005), who analytically shows that, after controlling for forward earnings yield, the expected return is negatively related to book-to-market ratio in presence of growth. The theoretical construction of ER_t also makes the source of negative relation clear: implied other information ϑ_t is negatively related to book-to-market ratio because low valuation (i.e. high b_t/P_t) is associated with lower future earnings expectations. Turning to the other 'risk proxies', note that $Beta_t$ receives positive and marginally significant coefficients; $Mcap_t$ receives highly significant and negative coefficients; Lev_t loads significantly positively on ER_t; and Mom_t continues to imply a return reversal in expected return.

Columns (1)-(3) in Table 7 include only one of the 'growth proxies' to account for potential high correlations between these variables by construction.[25] Now all these variables start to receive negative coefficients, consistent with their empir-

[25] Refer to variable definitions provided in Section 4.

ical relation to realized returns. ΔNOA_t is highly significant with t-statistics of -7.38 and Inv_t is only significant at 10% level, but Acc_t marginally fails to meet any conventional significance level. In the last column, when all the 'growth proxies' are included, only ΔNOA_t continues to be significant, and the sign of the coefficient on Acc_t turns negative yet insignificant.

Overall, the results in Table 7 shows that the associations of ER_t with firm characteristics are mostly consistent with prior literature and our expectations. However, caution must be applied when interpreting these findings because there is little *a priori* guidance on how expected returns *should* be associated with these firm characteristics.

6 Regression-based validation tests

To further validate the predictive ability of ER_t for future realized returns, We perform several regression-based validation tests. The validation methodology is based on a tautological decomposition of realized stock returns (Campbell and Shiller 1988; Campbell 1991; Vuolteenaho 2002). To start with, the realized return $R_{t,t+1}$ can be decomposed into 'true' expected return ER_t^* plus 'news' ε_{t+1}^*, such that ER_t^* is the conditional expectation and ε_{t+1}^* is of zero mean and not *ex ante* forecastable:

$$R_{t,t+1} = ER_t^* + \varepsilon_{t+1}^*$$

This implies that the true expected return ER_t^* predicts future realized return one-to-one. Of course, ER_t^* is latent and any empirical measure of it $\widehat{ER^*}_t$ is bound to involve measurement errors. Consider the following regression model

Table 7: Fama-Macbeth regressions of expected returns on firm characteristics

	Dependent variable: ER_t			
	(1)	(2)	(3)	(4)
$FE_{t,t+1}/P_t$	0.487***	0.488***	0.486***	0.488***
	(11.18)	(11.05)	(11.14)	(11.11)
b_t/P_t	-0.0216***	-0.0212***	-0.0214***	-0.0219***
	(-7.60)	(-7.48)	(-7.49)	(-7.62)
$Beta_t$	0.00151	0.00148*	0.00143	0.00165*
	(2.03)	(2.05)	(1.98)	(2.29)
$Mcap_t$	-0.00293***	-0.00294***	-0.00292***	-0.00295***
	(-8.83)	(-8.81)	(-8.61)	(-8.70)
Lev_t	0.00229***	0.00226***	0.00225***	0.00224***
	(9.98)	(10.16)	(9.72)	(10.20)
Mom_t	-0.0108**	-0.0107**	-0.0109**	-0.0108**
	(-3.57)	(-3.51)	(-3.56)	(-3.57)
ΔNOA_t	-0.0209***			-0.0228***
	(-7.38)			(-4.87)
Acc_t		-0.00875		0.0116
		(-1.93)		(1.65)
Inv_t			-0.00635*	-0.00174
			(-2.63)	(-0.74)
constant	0.0865***	0.0856***	0.0866***	0.0872***
	(18.59)	(18.34)	(18.32)	(18.25)
R^2	0.197	0.195	0.195	0.197

Notes: This table reports time-series average future realized returns on cross-sectional expected-return-sorted decile portfolios over 1-, 2-, 3-, 4-, 5- and 10-year horizons. All portfolios are formed on equally weighted basis at date t. $R_{t,t+j}, j = 1, 2, 3, 4, 5, 10$ is the realized return from date t to date $t + j$. The second last row reports returns to buying the 10th and selling the 1st decile portfolios over the respective horizons, and the last row reports their respective t-statistics.

$$R_{t,t+1} = \delta_0 + \delta_1 \widehat{ER^*}_t + \varepsilon_{t+1} \tag{9}$$

An unbiased expected return measure should measure ER_t^* with a zero-mean measurement error, thus $\mathbb{E}_t[R_{t,t+1}|\widehat{ER^*}_t] = \widehat{ER^*}_t$. This implies the slope coefficient δ_1 should be close to one and the intercept δ_0 should be close to zero.

However, as is discussed in Lee et al. (2015), in most research settings, the focus is on the *relative* levels of expected returns, and the bias in the measurement error is often neutralized and thus irrelevant for almost all applications.[26] Thus, the central issue regarding assessing the quality of an expected return measure is how closely the *variation* in $\widehat{ER^*}_t$ 'tracks' the *variation* in $ER^*{}_t$. In this vein, a good measure of ER_t^* should produce δ_1 significantly positive and close to one, and the intercept is allowed to differ from zero to absorb the measurement bias. Of course, due to the attenuation effect of measurement errors in the independent variable, the coefficient is expected to be smaller than one.

While this approach is theoretically sound, Easton and Monahan (2005) and Easton and Monahan (2016) point out that the news component ε_{t+1}^* is likely to have a non-zero mean and be correlated with expected returns in the post-war US data.[27] Thus, they develop a method that explicitly controls for proxies of the news component ε_{t+1}^*. Specifically, Easton and Monahan (2005) specify the following regression model based on the Vuolteenaho (2002) return decomposition framework:

$$r_{t,t+1} = \delta_0' + \delta_1' \hat{\mu}_t + \delta_2' cfn_{t+1} - \delta_3' drn_{t+1} + \varepsilon_{t+1}' \tag{10}$$

[26]Lee et al. (2015) surveyed more than 80 papers published in top accounting and finance journals that use some measures of expected returns. Only three of these papers require unbiasedness of the expected return measures, while the others are concerned with cross-sectional or time-series variations in expected returns.

[27]Easton and Monahan (2016) attribute this to the extraordinary success of the US economy and capital market development in the post-war period.

where $r_{t,t+1}$ is the log of realized return, $\hat{\mu}_t$ is the natural log of $\widehat{ER^*}_t$, and cfn_{t+1} and drn_{t+1} are contemporaneously measured cash flow news and discount rate news respectively. An expected return measure that is perfectly correlated with the 'true' expected return should lead to approximately $\delta_1' = \delta_2' = \delta_3' = 1$.[28] In this study, we measure cash flow news as

$$\begin{aligned} cfn_{t+1} &= \frac{roe_{t+1} - froe_{t,t+1}}{1 - 0.96\kappa_c} \\ &= \frac{log(1 + x_{t+1}/b_t) - log(1 + FE_{t,t+1}/b_t)}{1 - 0.96\kappa_c} \end{aligned}$$

where κ_c is the expected persistence of roe_{t+1} estimated from a two-year cross-sectional hold-out sample for each industry.[29] The second equality defines roe_{t+1} and $froe_{t,t+1}$. The number 0.96 is approximately one minus historical average dividend yield.[30] This measure of cfn_{t+1} is consistent with Easton and Monahan (2005) except that it only requires one-year-ahead earnings forecasts to avoid further data loss. One must evaluate the 'closeness' statistically.

Discount rate news is measured as the following:

$$drn_{t+1} = \frac{\hat{\mu}_{t+1} - \kappa_r \hat{\mu}_t - \bar{r}}{1 - 0.96\kappa_r}$$

where $\bar{r}$ and κ_r are the constant and persistence parameters of the expected return process, and $\hat{\mu}_{t+1} = \bar{r} + \kappa_r \hat{\mu}_t + \varepsilon^{\mu}_{t+1}$. Because the new measure of expected return is time-varying, this measure of drn_{t+1} differs considerably from that used in Easton and Monahan (2005), which applies only to models assuming constancy of expected returns. Implicitly in this measurement, we assume expected returns follow a first-order affine process, which is a common assumption in expected return models that admit time-varying expected returns.

[28] Note that this relation is not exact because the Vuolteenaho (2002) return decomposition is based on first-order Taylor approximation.

[29] All estimates of κ_c are significantly below 1 rejecting unit roots.

[30] Vuolteenaho (2002) estimates this value to be 0.97, Easton and Monahan (2005) estimate this value to range from 0.95 to 0.985. Empirically, as long as it is slightly smaller than one, the exact value chosen has no impact on almost all applications, so We do not estimate it. Rather, We replicate all subsequent analyses after replacing 0.96 with 0.95, 0.97 and 0.98, and the results are unaffected.

Table 8: Pooled regression test of the association of expected returns and realized returns

Models	(1)	(2)	(3)	(4)
Dependent variable: $r_{t,t+1}$ (test sample N=62,882)				
$\hat{\mu}_t$	0.894***	1.131***		1.027***
	(20.55)	(26.56)		(24.34)
cfn_{t+1}		0.854***		0.854***
		(46.37)		(45.59)
drn_{t+1}		-0.922***		-0.882***
		(-28.73)		(-27.53)
b_t/P_t			0.103***	0.0846***
			(20.28)	(17.42)
$Mcap_t$			0.00512***	-0.00654***
			(5.67)	(-7.64)
Lev_t			0.00608***	0.00565***
			(4.34)	(4.27)
$Beta_t$			-0.0308***	-0.0131***
			(-12.20)	(-5.44)
ΔNOA_t			-0.170***	-0.101***
			(-9.56)	(-5.95)
Acc_t			-0.0238	-0.0773***
			(-0.96)	(-3.34)
Inv_t			-0.0462***	-0.0382***
			(-3.80)	(-3.35)
Constant	0.00265	0.0173***	0.0280**	0.0365***
	(0.65)	(4.33)	(3.25)	(4.06)
adj. R^2	0.032	0.145	0.050	0.164

Pooled regressions

We first run pooled regression tests to evaluate the quality of the new expected return measure ER_t. We consider four model specifications for the regression tests. In model (1), We run the univariate test as in equation (9) after log-transformations of realized and expected returns. In model (2), We estimate equation (10) by augmenting the model (1) with estimated cfn_{t+1} and drn_{t+1}. Model (3) regresses (log) realized returns on a vector of firm characteristics examined in Table 6 and 7, and model (4) includes all variables in estimation to test if $\hat{\mu}_t$'s explanatory power for realized returns is incremental to these variables.

Table 8 reports the regression coefficients, their two-way robust t-statistics and adjusted R^2. Note that the requirement for non-missing values for cfn_{t+1} and drn_{t+1} further restricts the testing sample to 62,882 firm-year observations. In the univariate specification, model (1) produces a slope coefficient of 0.894, which is highly significant (t-statistic 20.55) and numerically close to the theoretical value of one. In comparison, leading ICC measures typically receive coefficients between 0.2 and 0.5 (see, for example, Table 3 of Lee et al. 2015). In addition, the regression intercept is statistically insignificant, suggesting the bias in $\hat{\mu}_t$ is small. Nonetheless, the t-test rejects the null that the slope coefficient is one at 95% level (p-value 0.0152).

Model (2) reports a 1.131 coefficient on $\hat{\mu}_t$, 0.854 on cfn_{t+1} and -0.922 on drn_{t+1} —all numerically close one. Notably, the adjusted R^2 is considerably improved to 14.5% compared to that of model (1) (3.2%). It appears, however, adding cash flow news and discount rate news proxies has introduced additional measurement bias, giving rise to a significant intercept. F-test rejects the joint hypothesis that coefficients on er_t, 0.854 and cfn_{t+1} are equal to one and the coefficient on drn_{t+1} is -1.

Model (3) estimates show that the set of firm characteristics explain as much as

5% of total variation in one-period-ahead realized returns. The results for model (4) show that the predictive ability of $\hat{\mu}_t$ for $r_{t,t+1}$ is robust to controlling for cash flow news, discount rate news and firm characteristics simultaneously.

Cross-sectional regressions

While the pooled regression results seem to support that the new measure of expected return has sensible predictive power for one-period-ahead realized returns, most applications of expected return measures concern whether investors can improve their cross-sectional portfolio strategies using these measures. Hence, We next evaluate the performance of the measure in the cross-section.

Table 9 presents the cross-sectional regression results based on Fama and MacBeth (1973) coefficient aggregation method. The second column (labeled 'univariate') shows that in a univariate regression, the expected return measure $\hat{\mu}_t$ claims a positive coefficient of 0.570, which is significantly positive (t-statistic 4.05) yet statistically different from one at 5% level (p-value 0.0047). Nonetheless, compared to existing measures of expected returns, it appears that $\hat{\mu}_t$ outperforms the other mesaures.[31]

In the third column labeled 'EM', We report estimates based on Easton and Monahan (2005) regression. EM regression has proved a hard test in the literature, as most leading ICC measures are found to be negatively associated with realized returns after controlling for cfn_{t+1} and drn_{t+1} (Easton and Monahan 2005; Nekrasov and Ogneva 2011).[32] The results suggest that the new measure $\hat{\mu}_t$ significantly outperforms leading ICC measures. The coefficient on $\hat{\mu}_t$ is 0.823,

[31] Guay et al. (2011) reports that most ICC measures receive coefficients between -0.33 and 0.43 in univariate cross-sectional regressions, and these coefficients are not statistically distinguishable from zero at conventional levels.

[32] To our knowledge, no existing ICC measure based on analyst forecasts receives a coefficient close to one and significantly above zero in EM regressions. The measure proposed by Nekrasov and Ogneva (2011) marginally meets this criterion only after it is adjusted for predictable analyst forecast errors.

Table 9: Cross-sectional regression test of the association of expected returns and realized returns

	Univariate	EM	Extended EM
Dependent variable: $r_{t,t+1}$ (test sample N=62,882, T=30)			
$\hat{\mu}_t$	0.570***	0.823***	0.736***
	(4.05)	(5.89)	(6.44)
cfn_{t+1}		0.845***	0.830***
		(18.36)	(21.47)
drn_{t+1}		-0.626***	-0.591***
		(-6.13)	(-5.59)
b_t/P_t			0.0315*
			(2.74)
$Mcap_t$			-0.00373
			(-1.16)
Lev_t			0.00388
			(1.45)
$Beta_t$			-0.0105
			(-0.76)
ΔNOA_t			-0.0245
			(-0.96)
Acc_t			-0.109*
			(-2.64)
Inv_t			-0.0522
			(-2.02)
Constant	0.00924	0.0236	0.0474
	(0.28)	(0.72)	(1.37)
Average R^2	0.009	0.121	0.137

which is highly significant (t-statistic 5.89) and statistically indistinguishable from one (p-value 0.216). The intercept also remains statistically insignificant. In addition, the coefficients on cash flow news and especially discount rate news proxies are also more reasonable than those reported in prior studies. The coefficient on cfn_{t+1} is 0.823, which is suggestively close to one, although it is statistically different from one (p-value 0.002) due to its small standard error. The coefficient on drn_{t+1} amounts to -0.626, which is much closer to -1 than discount rate proxies computed from existing ICC measures. This reflects a much stronger negative association between discount rate news and contemporaneous realized returns. In comparison, even discount rate news proxies based on the best-performing ICC measures receive coefficients lower than 0.20, suggesting that stock returns appear insensitive to discount rate changes. (Easton and Monahan 2005; Nekrasov and Ogneva 2011; Mohanram and Gode 2013). Our new measure $\hat{\mu}_t$ significantly improves in this respect.

The last column of Table 9 reports estimates from an 'extended' version of EM regression. Specifically, We regress $r_{t,t+1}$ on $\hat{\mu}_t$, cfn_{t+1}, drn_{t+1} and a vector of firm characteristics that have been shown to predict stock returns in the cross-section. The coefficient estimates on $\hat{\mu}_t$, cfn_{t+1} and drn_{t+1} are quantitatively similar to those in EM regression, although the coefficient on $\hat{\mu}_t$ becomes statistically distinguishable from one at 99% level (p-value 0.028) due to its correlation with firm characteristics. Importantly, it seems that only b_t/P_t and Acc_t remain marginally significant in explaining the cross-section of realized returns, once expected returns and news proxies are controlled for. In other words, the return predictive role of most of these characteristics are 'driven out'. These results suggest the new measure of expected return exhibits strong association with one-period-ahead realized returns in the cross-section.

7 Summary

In this paper, We develop and evaluate a new method for constructing a measure of one-period-ahead expected returns from firm fundamentals. The measure is forward looking and admits time-varying expected risk premiums through time-varying volatility in the firm fundamentals. It also avoids making arbitrary assumptions about firms' terminal payoff growth rates and dividend policies. The expected return estimates show a strong association with future realized returns and consistent associations with a range of return predictive variables.

References

Ang, A. and Liu, J. (2001). A general affine earnings valuation model. *Review of Accounting Studies*, 6(4):397–425.

Ashton, D. and Wang, P. (2013). Terminal valuations, growth rates and the implied cost of capital. *Review of Accounting Studies*, 18(1):261–290.

Botosan, C. A. and Plumlee, M. A. (2005). Assessing alternative proxies for the expected risk premium. *The accounting review*, 80(1):21–53.

Botosan, C. A., Plumlee, M. A., and Wen, H. (2011). The relation between expected returns, realized returns, and firm risk characteristics. *Contemporary Accounting Research*, 28(4):1085–1122.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal*, 101(405):157–179.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Chava, S. and Purnanandam, A. (2010). Is default risk negatively related to stock returns? *The Review of Financial Studies*, 23(6):2523–2559.

Chen, L., Novy-Marx, R., and Zhang, L. (2011). An alternative three-factor model.

Claus, J. and Thomas, J. (2001). Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets. *The Journal of Finance*, 56(5):1629–1666.

Cochrane, J. H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy*, 104(3):572–621.

Cochrane, J. H. (2009). *Asset Pricing:(Revised Edition)*. Princeton university press.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance*, 66(4):1047–1108.

Dittmar, R. F. and Lundblad, C. T. (2017). Firm characteristics, consumption risk, and firm-level risk exposures. *Journal of Financial Economics*.

Easton, P., Taylor, G., Shroff, P., and Sougiannis, T. (2002). Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research*, 40(3):657–676.

Easton, P. D. (2004). Pe ratios, peg ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review*, 79(1):73–95.

Easton, P. D. and Monahan, S. J. (2005). An evaluation of accounting-based measures of expected returns. *The Accounting Review*, 80(2):501–538.

Easton, P. D. and Monahan, S. J. (2016). Review of recent research on improving earnings forecasts and evaluating accounting-based estimates of the expected rate of return on equity capital. *Abacus*, 52(1):35–58.

Elton, E. J. (1999). Presidential address: expected return, realized return, and asset pricing tests. *The Journal of Finance*, 54(4):1199–1220.

Fairfield, P. M., Whisenant, J. S., and Yohn, T. L. (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The accounting review*, 78(1):353–371.

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

Fama, E. F. and French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance*, 51(1):55–84.

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of financial economics*, 43(2):153–193.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636.

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

Feltham, G. A. and Ohlson, J. A. (1999). Residual earnings valuation with risk and stochastic interest rates. *The Accounting Review*, 74(2):165–183.

Ferson, W. E. and Foerster, S. R. (1994). Finite sample properties of the generalized method of moments in tests of conditional asset pricing models. *Journal of Financial Economics*, 36(1):29–55.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

Gebhardt, W. R., Lee, C., and Swaminathan, B. (2001). Toward an implied cost of capital. *Journal of Accounting Research*, 39(1):135–176.

Givoly, D. and Hayn, C. (2000). The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of accounting and economics*, 29(3):287–320.

Gode, D. and Mohanram, P. (2003). Inferring the cost of capital using the ohlson–juettner model. *Review of Accounting Studies*, 8(4):399–431.

Gode, D. and Ohlson, J. (2004). Accounting-based valuation with changing interest rates. *Review of Accounting Studies*, 9(4):419–441.

Guay, W., Kothari, S., and Shu, S. (2011). Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management*, 36(2):125–149.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.

Hansen, L. P., Heaton, J., and Yaron, A. (1996). Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, 14(3):262–280.

Hayashi, F. (2000). Econometrics.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Hughes, J., Liu, J., and Liu, J. (2009). On the relation between expected returns and implied cost of capital. *Review of Accounting Studies*, 14(2-3):246–259.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, 49(5):1541–1578.

Lee, C., So, E. C., and Wang, C. C. (2015). Evaluating firm-level expected-return proxies.

Lee, C. M. and So, E. C. (2017). Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics*, 124(2):331–348.

Lewellen, J. (2014). The cross section of expected stock returns. *Critical Financial Review*, forthcoming.

Lyle, M. R., Callen, J. L., and Elliott, R. J. (2013). Dynamic risk, accounting-based valuation and firm fundamentals. *Review of Accounting Studies*, 18(4):899–929.

Lyle, M. R. and Wang, C. C. (2015). The cross section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics*, 116(3):505–525.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, pages 867–887.

Mohanram, P. and Gode, D. (2013). Removing predictable analyst forecast errors to improve implied cost of equity estimates. *Review of Accounting Studies*, 18(2):443–478.

Nekrasov, A. and Ogneva, M. (2011). Using earnings forecasts to simultaneously estimate firm-specific cost of equity and long-term growth. *Review of Accounting Studies*, 16(3):414–457.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research*, 11(2):661–687.

Ohlson, J. A. (2005). A simple model relating the expected return (risk) to the book-to-market and the forward earnings-to-price ratios.

Ohlson, J. A. and Juettner-Nauroth, B. E. (2005). Expected eps and eps growth as determinantsof value. *Review of accounting studies*, 10(2):349–365.

Pástor, L., Sinha, M., and Swaminathan, B. (2008). Estimating the intertemporal risk–return tradeoff using the implied cost of capital. *The Journal of Finance*, 63(6):2859–2897.

Penman, S. H. and Zhu, J. L. (2014). Accounting anomalies, risk, and return. *The Accounting Review*, 89(5):1835–1866.

Penman, S. H. and Zhu, J. L. (2017). Accounting-based estimates of the cost of capital: A third way. *Browser Download This Paper*.

Piotroski, J. D. and So, E. C. (2012). Identifying expectation errors in value/glamour strategies: A fundamental analysis approach. *The Review of Financial Studies*, 25(9):2841–2875.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory*, 13(3):341–360.

Shumway, T. (1997). The delisting bias in crsp data. *The Journal of Finance*, 52(1):327–340.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3):289–315.

Van Dijk, M. A. (2011). Is size dead? a review of the size effect in equity returns. *Journal of Banking & Finance*, 35(12):3263–3274.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance*, 57(1):233–264.

Zhang, X. F. (2007). Accruals, investment, and the accrual anomaly. *The Accounting Review*, 82(5):1333–1363.