Accounting for Asset Pricing Factors

Stephen Penman*
Columbia Business School, Columbia University
and Bocconi University
shp38@columbia.edu

Xiao-Jun Zhang
Haas School of Business, University of California, Berkeley
xzhang@haas.berkeley.edu

December, 2021

*Corresponding author. Our thanks go to Ray Ball, Matthias Breuer, Atif Ellahie, Hyung Il Oh, Martin Kapons, Scott Richardson, and Lu Zhang for comments.

Accounting for Asset Pricing Factors

Abstract. Many accounting numbers appear in standard factor models, including book value, investment, return on equity (ROE), and other profitability measures. However, they do so without a clear explanation of why they indicate investment risk. The papers shows that these numbers are generated under accounting principles that deal with risk, providing an explanation but also a critique of how the numbers enter extant models. That leads to a revision in factor construction. Accounting numbers are codetermined in a double-entry system, a feature that is exploited in packaging the revised factors into a factor model. Rather than entering as the separate, additive factors in extant models, adding to the "factor zoo," the numbers are combined parsimoniously to capture the information they jointly convey about risk and return in the double-entry system. Empirical tests against extant models confirm.

Accounting for Asset Pricing Factors

1. Introduction

Many factors in empirical asset pricing models are formed from accounting numbers, including book-to-price (B/P), investment, return on equity (ROE), and other profitability measures. Many have been elicited by data mining, selected because they perform relatively well in explaining returns empirically. That hints of explanatory power, but without explanation: Why do the accounting measures capture the risk to which the investor is exposed? Without that explanation, investors presumably add model risk to fundamental risk. Indeed, an investor is not sure whether the factors capture exposure to risk or abnormal returns to market mispricing. Plausible conjectures are offered, typically equating accounting measures with economic constructs. Some factors are identified formally from economic theory, as in investment-based models, with accounting numbers then embraced as representing the economic constructs.

However, accounting numbers do not necessarily capture the constructs in mind; they are accounting phenomena, generated by accounting rules, and these rules depart from economic ideal. A seemingly reasonable reaction is to adjust the measures for their so-called deficiencies, as several recent papers have done. However, this paper recognizes there is method in the madness: Accounting principles that produce the so-called accounting distortions convey risk and return, and this information can be exploited in building factor models.

The paper focuses on the B/P factor (called the value factor or HML in the Fama and French (1993) model and its subsequent extensions), the investment factor (the CMA factor in those extensions and a factor in the Investment CAPM of Hou, Xue, and Zhang, 2015 and its q^5 extension), and ROE and other profitability factors in these models. Each of these measures is determined by accounting principles that disturb them from the phenomena they are said to capture, but accounting that conveys information about the risk to investing. The paper reports that the standard factors do not fully incorporate this information, prompting an alternative factor construction. It also shows that recent attempts to adjust accounting numbers for suspected deficiencies (in Lev and Srivastava 2020, Arnott, Harvey, Kalesnik, and Linnainmaa 2021, Eisfeldt, Kim, and Papanikolaou 2020, and Bongaerts, Kang, and van Dijk 2021, for example) lose information about risk and return.

The investigation leads to an alternative factor model that exploits features of the accounting system. Extant models typically package factors together as separate, additive factors, raising the "factor zoo" complaint as another factor is "discovered" in the data. The paper recognizes that accounting numbers are produced in a coordinating system, the double-entry system, under which the accounting for one measure affects another: A "debit" always requires a matched "credit." Accordingly, accounting numbers are co-determined, conveying information jointly, and that can be exploited in building factor models. The double-entry accounting system is a packaging of those numbers, implying an alternative packaging of factors in a factor model, and parsimoniously so. The double-entry system that has often been marveled at for its properties can also be utilized in constructing factor models, and the paper does so.

The paper is offered in the spirit of placing empirical asset pricing and the accounting on which it draws on the same platform. That is done by recognizing accounting numbers that are governed by accounting principles that convey risk and return in a double-entry system.

1.2 Framing the Identification of the Relevant Accounting Information

The analysis is developed under theory that ties accounting numbers to risk. Given the no-arbitrage dividend discount model (with a constant discount rate, r) and clean surplus accounting under which earnings add to book value (by double entry) and dividends are paid out of book value,

$$P_t^T = \frac{\sum_{\tau=t+1}^{T}\left[Earnings_\tau + Dividend_\tau\left(\left(1+r^f\right)^{T-\tau}-1\right)\right]}{(1+r)^T-1} \rightarrow P_t, \quad (1)$$

the price in the infinite-horizon dividend discount model, as T → ∞ (with the usual "transversality condition" that dividends are ultimately paid out). That is, price is expected cumulative future earnings with dividends reinvested (at the risk-free rate, r^f), capitalized at the required return. (As at all points in the paper, variables subscripted τ > t are expected values.) The expression is equivalent to an infinite-horizon residual income formula; see Ohlson (1995). The constant discount rate and risk-free rate are for simplicity; the accounting involved in this paper can be tied to a risk discount represented by a varying covariance term in a general asset pricing model, as in Penman and Zhang (2020).

Transposition renders the required return:

$$(1+r)^T - 1 = \frac{\sum_{\tau=t+1}^{T}\left[Earnings_\tau + Dividend_\tau\left(\left(1+r^f\right)^{T-\tau}-1\right)\right]}{P_t^T} \quad (2)$$

for T → ∞: The required return is given by expected future cum-dividend earnings relative to the current price that is discounted for the risk to the expected earnings. This is the expected earnings yield that Ball (1978) suggested is related to risk and return, but now formally so. The theoretical and empirical question is what accounting numbers (and other information) forecast this yield. Information that forecasts expected cum-divided earnings in the numerator is pertinent, but also information that conveys the risk discount in the denominator. Given clean surplus accounting, equation (2) is equivalent to

$$(1+r)^T - 1 = \frac{P_T + \sum_{\tau=t+1}^{T} Dividends_\tau((1+r^f)^{T-\tau} - 1) - 1}{P_t^T} \quad (3)$$

for T → ∞: The expected earnings yield is equivalent to the expected return to the current price.

However, equations (1) and (2) are derived simply by substituting earnings for dividends in the dividend discount model, and mere substitution puts nothing on the table (as a matter of theory). Accordingly, the expressions are silent as to the accounting involved to convey the information. Indeed, random numbers satisfy the equations provided that earnings are ultimately calculated (in the limit) to yield a book value that goes to zero, along with P_T, as the firm pays out dividends. So accounting principles that convey r in equation (2) must be introduced and Section 2 does so.

This framing differs from that in Fama and French (2015) which guides the identification of factors for their extended factor model. While accounting numbers are also introduced via the clean-surplus relation there, Penman, Reggani, Richardson, and Tuna (2018) and the internet appendix to Penman and Zhu (2022) show how their comparative statistics are inconsistent with how accounting works. And they do not incorporate accounting principles that convey risk; the comparative statics in Fama and French (2015) work with random accounting numbers.

The tests in the paper are in two sections. In Section 3, the accounting-based factors in extant models are evaluated on the extent to which they convey risk and return information conveyed by accounting, with alternative factors based that information then identified and evaluated empirically. Then, in Section 4, relevant factors are packaged into a factor model under the governance of the double-entry system and compared empirically to extant factor models. But, first, Section 2 outlines the accounting principles that convey risk and return.

2. A Primer on Accounting Principles

2.1 Accounting 101

With apologies for being patronizing to the informed reader, this section lays out accounting principles and explains how they potentially convey risk and expected return.

Accrual accounting involves reporting an income statement and balance sheet that track earnings and book of equity and their components over time. With clean-surplus double-entry accounting, earnings add to balance-sheet book value, one-for-one; the two are codetermined.

The governing principle for booking earnings and the corresponding balance sheet involves an assessment of risk: Earnings are recognized only when uncertainty is resolved. Until that "realization" point, the recognition of earnings is deferred. Thus, while a firm's market price will (rationally) incorporate future earnings expectations, the accounting informs that those earnings are still at risk. They have not yet been realized, they should be discounted.

Earnings is revenue minus expenses, so this accounting is implemented under two sub-principles.

First, under the so-called revenue realization principle, revenue is not booked until a customer contracts, obligations under the contract are satisfied by both parties, and receipt of cash is "highly certain."[1] Only then is revenue added to the income statement and, correspondingly, to book value in the balance sheet, usually with a receivable or cash. Even then, a receivable is discounted (with an allowance for bad debts) to an expected cash equivalent. In asset pricing terms, revenue is recognized only when a firm can book a low-beta asset, cash or a near-cash

[1] See the Financial Accounting Standards Board (FASB) ACS 606 and the International Accounting Standards Board (IASB) IFRS 15 standards on revenue recognition. These new standards replace earlier standards, but the basic principle remains, indeed reinforced.

(discounted) receivable. Until that point, revenue recognition is deferred as positive beta, an expectation at risk of not being realized—the customer might not materialize.

Second, the accountant then recognizes expenses incurred to gain revenues, so-called matched expenses, to report net value added from revenues. However, these matched expenses are disturbed by so-called conservative accounting that expenses some investments. While many investments, in inventory, plant, and equipment, for example, are booked to the balance sheet, others are charged against earnings if the outcome to the investment is particularly uncertain. In essence, current earnings is discounted for higher risk with the expensing of these investments. R&D investment and advertising and promotion to build brand assets are the marquee examples, but the accounting is pervasive, applied to investment to develop supply chains and distribution systems, customer loyalty programs, employee training and retention, software development, start-up and organization costs, and more.[2] So-called conditional conservatism—impairing the assets actually booked to the balance sheet on lowered expectations of payoffs—reinforces this accounting.

These are principles that respond to risk, affecting accounting measures including investment, book value, earnings, and ROE that appear in asset pricing models. We will elaborate on the principles when we come to these measures. Of course, practice may not follow principle perfectly. The principles require estimates, with the possibility of manipulation (though subject

[2] Under the FASB and IASB Conceptual Framework, the qualification of "probable future benefits" applies for an asset to be booked to the balance sheet, otherwise the investment is expensed. For the accounting for R&D under FASB Statement No. 2, the FASB requires the investment to be expensed due to the "uncertainty of future benefits." In IAS 38, the IASB applies the criterion of "probable future economic benefits" to distinguish between "research" (which is expensed) and "development" (which is capitalized in the balance sheet and amortized). Even IAS 16 on property, plant, and equipment requires benefits to be "probable" for the asset to be booked. While these principles are in regulatory accounting standards, the practice of conservative accounting largely developed voluntarily before twentieth-century regulation, as did the revenue realization principle.

to audit). And perhaps accountants are "too conservative." So it is an empirical question whether reported numbers convey the information under accounting principles, which we answer.

2.2 Accounting Principles and Asset Pricing Theory

The idea of tying the expected return to the timing of risk resolution is not foreign to asset pricing research. See, for example, Epstein and Zin (1989), Bansal and Yaron (2004), Dechow, Sloan, and Soliman (2004), Weber (2018), and Mullins (2020).[3] There is no guarantee that the risk that accountants respond to is the priced risk of asset pricing models, of course, but Penman and Zhang (2020) connect the two principles—the revenue realization principle and conservative accounting for investment—to the discount factor in a general, no-arbitrage asset pricing model. They also show how the resulting accounting measures, convey priced risk as a matter of theory. Empirical support is in Penman and Zhang (2021), Penman and Yehuda (2019), and Penman and Reggiani (2018).

3. The Information about Risk and Return in Accounting Measures

We now evaluate whether the accounting numbers entering extant asset pricing models are those that convey risk under these accounting principles. We also identify alternative numbers that are so determined.

Compustat and CRSP are the data sources, as they are in most asset pricing papers referenced. Accounting measures are calculated for each year, 1964-2018, for all firms with the required data on Compustat in that year, except for Financial firms (SIC codes 6000-6999) and utilities. Firms classified as Membership Organization or Unknown SIC are also excluded. Forward annual

[3] Equation (2) assumes a constant discount rate but can be modified for the timing of earnings realization to estimate a term structure for the expected equity return. That is beyond the scope of this paper, though a higher discount for more distant realizations (from expensed investment) is implicit.

returns cover the twelve months beginning three months after fiscal-year end by which time the relevant accounting data must be reported, by law. This period covers the period when (realized) accounting numbers in the expectation in price at the beginning of the return period are reported (in four quarterly reports). The appendix details the calculation of variables.

Table 1 reports summary statistics for measures that enter later in the paper. Panel A summarizes their distributions and Panel B average cross-sectional correlations for selected measures. We simply post these tables at this point, with commentary deferred until the measures are introduced in the empirical analysis. The correlations in Panel B are important when asking if alternative measures are capturing the same thing. As factor portfolios are formed by ranking variables, we will be referring to the rank correlations.

The empirical analysis in each section below covers a specific accounting number appearing in extant factor models. We begin each section by describing the accounting for the number and how that indicates risk and return under accounting principles (or not). The empirical analysis tests whether the numbers are priced in accordance with the underlying accounting principles. This is in two subsections. The first involves tests with cross-sectional "characteristic regressions" that report how the measures correlate with forward stock returns. These regressions have served the data mining for factors, but here are run with predictions for coefficient estimates from accounting principles. As pointed out in Fama and French (2020), estimated slope coefficients can be interpreted as factor returns. However, research also forms "factor mimicking" portfolios from characteristics, so the second subsection constructs factors based on the findings from the characteristic regressions that are then evaluated relative to extant factors. That is an entree to the development of factor models in Section 4.

The standard caveat applies to our empirical analysis: We assume that the market prices risk efficiently. The accounting-arbitrage alternative—investors do not process accounting information efficiently—cannot be taken off the table. Note also here are many t-statistics reported, so the reader is reminded of the chance of reporting significant statistics with repeated drawings from the t-distribution even though there is no substantive relation. However, the analysis is based on accounting principles connected to risk in theory, mitigating both concerns,

3.1 Accounting for Investment

The investment number in factor models is the change in balance-sheet assets that has been shown to forecast lower returns. However, this number is not the firm's investment but rather just the part of that investment that accountants book to the balance sheet. Many investments, often referred to as investment in intangible assets, are expensed against earnings in the income statement under the conservative accounting principle of Section 2.1.

This principle distinguishes the two treatments based on risk: Investments where the outcome is particularly uncertain are not booked as assets to be aggregated with less risky investment as if they provide similar collateral. So, R&D expenditures into a possible drug where there is no product as yet, let alone customers and revenues (and there may not be), is distinguished from plant producing a product and from the inventory that is the product itself that yields revenues. And so for advertising that might fail to pull in customers and investments in employees who might leave to go to competitors. By expensing investments under this principle, the accounting informs that future revenues are relatively more uncertain—there is more doubt that revenues will be realized under the realization principle.

Correspondingly, investment booked to the balance sheet indicates a higher probability of realizing payoffs, projecting lower risk and return. In the parlance of finance, this investment is the realization of expected investment opportunities, a revision of the risk that investment opportunities will not be realized; it thus accords with intertemporal asset pricing theory in Merton (1973) under which risk is to (unrealized) future investment opportunities. In Berk, Green, and Naik (1999), risk declines as firms convert risky growth options to “assets in place,” and that is recognized in the accounting treatment.

Accounting principles thus explain the negative relation between balance-sheet investment and forward returns. However, research also indicates that the distinction between this investment and that booked to the income statement is also priced. Oh and Penman (2020) find the market discounts investment expensed to income statement relative to that booked to the balance sheet, and Oswald, Simpson, and Zarowin (2020) report that R&D that is capitalized to the balance sheet when it is demonstrated to be successful is priced at a two percent premium to R&D that is expensed as not yet successful. Further, expensed investments predict higher stock returns in Lev and Sougiannis (1996), Oh and Penman (2020), and Kapons (2020), among others.

Characteristic regressions

Panel A of Table 2 reports the results from annual cross-sectional regressions of forward returns, Ret_{it+1} over the risk-free rate, on investment measures. As in all cross-sectional regressions in the paper, reported coefficients are means over years, with t-statistics (in parentheses) calculated as those mean coefficients relative to their standard errors estimated from the time-series of coefficient estimates (Fama-and-Macbeth style). As factors are formed from a ranking of variables, the results are from decile-rank regressions, but those with continuous variables with an outlier treatment are similar.

Regression 1 in Panel A correlates forward returns with the investment measure in the Fama and French (2015 and 2018) five- and six-factor models, the Hou, Xue, and Zhang (2015) Investment CAPM, and the Hou, Mo, Xue, and Zhang (2020) q^5-factor model where $Investment_t/TA_{t-1} = \Delta Total\ Assets_t/Total\ Assets_{t-1}$. The mean coefficient on $Investment_t$ is reliably negative, consistent with previous findings, with *q*-theory, and intertemporal asset pricing theory. It is also consistent with the realization principle in accounting: Investment booked to the balance sheet indicates a higher probability of realizing earnings, thus conveying lower expected returns.

However, Regression 2 with expensed investment reports a different result. Here $ExpInv_t/TA_{t-1} = (R\&D_t + Advertising_t + (SG\&A - Advertising)_t \times 0.3)/Total\ Assets_{t-1}$ which is 15.9% of lagged total assets in Table 1, on average, but has a mean correlation with $Investment_t$ (booked to the balance sheet) of only 0.16.[4] An estimated portion of SG&A (Selling, General, and Administrative Expense) is included as expensed investments charged to that account, as in Peters and Taylor (2017). In contrast to the negative mean coefficient on $Investment_t$ in Regression 1, that on expensed investment is not significantly different from zero. It appears that this investment is priced differently: The t-statistic on the mean difference (over years) of -0.016 between the coefficients in Regressions 1 and 2 is -3.04.[5] The 30 percent of SG&A is

---

[4] In the correlation panel in Table 1, the investment numbers are denominated in price for later reference. The 0.16 here refers to the domination in lagged assets, as in the table here.

[5] The investment measure increases with realized earnings reinvested in the business. So, in part, it captures current earnings realizations (and fully so with no dividend payout or net share issue). Similar results are observed when investment is confined to changes in gross property, plant, and equipment and inventory (tangible assets), as in some papers. Results are also similar with $Investment_t = \Delta NOA_t / NOA_{t-1}$ (that excludes investment in cash) and with expensed investment, $ExpInv_t$, denominated in NOA_{t-1}. NOA is Net Operating Assets = Operating Assets – Operating Liabilities (defined in the appendix). Some papers, for example, Cho, Kremens, Lee, and Polk (2021) define investment as change in book equity or net share issuance, but equity can change with a debt-equity swap that does not affect investment in the business (as with a repurchase financed by borrowing).

presumably a very noisy estimate.[6] So Regression 3 repeats the second regression without this component, that is, $ExpInv_t/TA_{t-1}$ (w/o SG&A) = (Expensed $R\&D_t$ + $Advertising_t$)/Total $Assets_{t-1}$. The difference between the two measures is quite large at all percentiles in Table 1 and at the mean. Results for Regression 3 are similar to Regression 2. The t-statistic in the mean difference (over years) of -0.015 between the coefficients in Regressions 1 and 3 is -2.37.

These differences are accentuated when $Investment_t$ and $ExpInv_t$ are denominated relative to P_{t-1} in Regression 4 rather than lagged total assets. That price incorporates expectations and the discount for risk for the period, so investments are relative to expectations and, as such, enters the denominating price in the framing equation (2). Further, the measures have the same (lagged) denomination in price as returns, so additions to price are compared to accounting numbers added to price. With this deflator, the mean coefficient on $Investment_t$ is -0.009 (t = -2.95) and that on $ExpInv_t$ is 0.012 (t = 4.95).[7] The t-statistic on the mean difference (over years) between the two coefficients is -5.00.

In both the Fama and French extended factor models and the *q*-investment models, investment enters with a profitability measure. In the latter, profitability, measured by ROE, is implied by *q*-theory as the relation between investment and the required return is conditional on expected profitability. So Regression 5 adds ROE, as in the investment models, and Regression 6 adds the

---

[6] It is doubtful whether a number can be identified with the current disclosures on SG&A. It presumably varies over firms and eliciting the asset component for a given firm is fraught with difficulties, even with disclosures. The advertising component of SG&A might produce future revenues (an asset), but also generates current revenues and thus is a current period expense. And so with other aspects of SG&A, where the asset component of transactions is inseparable. For example, a bonus for employee retention (so-called investment in human capital) is paid with current salary and supplier and customer loyalty incentives are comingled with current revenues and expenses. Investments in "organization capital" are similarly made in conjunction with current operations. How much of the expenditure is an asset? The accounting principle is the impossibility of allocation of expenditures incurred jointly. It is this impossibility that frustrates the accounting (and the researcher). Thus the 30% of SG&A estimate of Peters and Taylor (2017), the same number applied to all firms, is questionable. Enache and Srivastava (2018) make an alternative estimate, but it looks more like one to distinguish fixed versus variable costs rather than assets.

[7] With $ExpInv_t/P_{t-1}$ in the regression alone, as in Regression 2, the mean slope coefficient is 0.014 (t = 5.20).

cash-flow profitability measure of Ball, Gerakos, Linnainmaa, and Nikolaev (2018), CP_t/TA_{t-1} chosen in Fama and French (2018). These profitability measures add explanatory power (for reasons later), but mean coefficients on the investment numbers are little different; the difference in coefficients on balance-sheet $Investment_t$ and expensed investment $ExpInv_t$ remain.

The final regression in Table 2 adds characteristics underlying factors in the Fama and French six-factor model to Regressions 6. The mean coefficient on expensed investment remains significantly positive: While investment booked to the balance sheet is included in these models (with a negative association with returns, as here), expensed investment incrementally projects higher average returns. These models do not capture an aspect of risk and return relevant to asset pricing.

Investment Factor Returns

With a view to factor construction, Panel B of Table 2 reports mean returns from long minus short positions determined from ranking firms on investment measures each year. The mean differences in returns between these portfolios (High – Low) is interpreted (as usual) as a "factor mimicking" return as if it is reward for risk from exposure to a latent priced factor. To align in calendar time, the long and short positions are determined from a ranking of firms at March 31 each year, with firms entering the ranking based on the accounting measure for its most recent fiscal year. At this date, the ranking number is recent for the majority of firms (with December 31 fiscal-years), but it is staler for fiscal-years ending in other months.[8]

[8] The March 31 date means the measures are as fresh as possible for a cross-sectional ranking in calendar time. However, results are similar with the ranking at each June 30, as in the Fama and French construction.

$Investment_t/TA_{t-1}$ is the measure used to construct the investment factors in factor models, CMA in Fama and French (FF) models and IA in *q* models. For both this measure and $ExpInv_t/TA_{t-1}$, the factor returns are those implied by the regression results, -0.093 for balance-sheet investment but an insignificant 0.025 for expensed investment. $TotInv_t$ in the panel is the sum of the two types of investment. The significant negative long – short return difference for $Investment_t$ is maintained, though lower with $ExpInv_t$ included.

However, the panel reports a very different result with the two investment numbers relative to beginning-of-period equity price. The negative factor return for $Investment_t$ survives, but the 0.105 return for $ExpInv_t$ is now significantly positive. This is also so in Bongaerts, Kang, and van Dijk (2021) who calculate an intangible asset factor (based on expensed investment) and report its associated Sharpe Ratio is higher than that for established factors. The return High – Low return difference for $TotInv_t$ (relative to price) is effectively zero.

The difference in the results with the two deflators is explained by Total Assets/Price, part of Book/Price, and that depends, in part, on the amount of investment expensed rather than booked to book value. B/P is ranked correlated 0.33 with $ExpInv_t/P_{t-1}$ in Table 1 but the correlation with $ExpInv_t/TA_{t-1}$ (not reported there) is -0.26. That leads us to the next section.

3.2 Accounting for Book Value

Book-to-price is ubiquitous in asset pricing, with higher B/P said to indicate higher risk and thus the basis for a B/P factor. Characteristic regressions support the inference. At first glance, that appears to be at odds with accounting principles. Earnings add to book value, so P/B is future earnings expected in price relative to that not yet recognized by the accounting. Those unrecognized earnings are earnings yet to be realized, earnings still at risk, so a high P/B, rather than a low P/B, should indicate higher risk. Book value is often interpreted as "assets in place," with B/P indicating the value (and risk) of "growth opportunities over assets in place." However, book value (net assets) does not include expensed investment. That might suggest that book value should be adjusted to include these assets. Several papers have done so, but the previous section indicates that this loses information about assets with different risks. There is much to be sorted out here.

Characteristic regressions

Table 3 has a similar layout to Table 2, but with measures of book value rather than investment. Regression 1 in Panel A reports the standard finding: B/P is positively related to returns. The measure in Regression 2, $CapExpInv_t/P_t = \frac{Capitalizec\ Expensed\ Investment_t}{P_t}$, capitalizes the expensed investment in Table 2 with ex ante amortization rates in the variable description in the appendix.[9] It is denominated in the same price as B/P. Panel A of Table 1 informs that capitalizing expensed investment adds significantly to balance sheet assets, 31.6 cents per dollar of price to B/P at the mean and 13.9 cents at the median with wide cross-sectional variation. Panel B there reports a mean rank correlation with B/P of 0.46, and the mean coefficient on this as-if book value in Regression 2 is similar to that on B/P in Regression 1. Regression 3 has the

[9] An ex ante amortization schedule for an asset where outcomes are highly uncertain and inevitably contains error. Amortization rates are often assumed or estimated from data that involves only surviving firms where the investment was successful, as in Lev and Srivastava (2020), for example. With this ex post bias, these rates are not applicable ex ante where less successful outcomes are also entertained; amortization rates based on successful outcomes would result in anticipated impairments. Peters and Taylor (2017) use BEA industry-specific amortization rates for R&D and a constant 20% for the capitalized SG&A component. Arnott, Harvey, Kalesnik, and Linnainmaa (2021) follow the Peter's and Taylor (2017) scheme. We make no claim that our assumed rates are any better. Barker, Lennard, Penman, and Teixeira (2021) discuss the accounting issues in recognizing these "intangible" assets. Note that we ignore deferred taxes which likely would be booked if an accountant capitalized expensed investments in practice. (These presumably would also be discounted in price as uncertain tax provisions.)

calculation without capitalizing and amortizing the SG&A component. The sizeable difference between the two measures is clear from Table 1, but results here are similar.

Lev and Srivastava (2020), Arnott, Harvey, Kalesnik, and Linnainmaa (2021), and Eisfeldt, Kim, and Papanikolaou (2020) report that the "value premium" improves with the addition of a similar measure to book value, and Eisfeldt and Papanikolaou (2013) show that capitalized SG&A is related to expected stock returns. However, Rizova and Saito (2020) claim there is little improvement from adjusting book values for capitalized assets. In Regressions 4 and 5, the improvement is explicit: Given B/P, the capitalized expensed investment adds to the explanation of returns. These regressions include the profitability measures, not only because these measures are included with B/P in factor models, but because, under double entry, book value cannot be considered without the complementary earnings which the accounting for book value affects. Regressions 6 and 7 show that this holds with further characteristics that enter existing factor models. With the exception of $Size_t$, these characteristics do not add explanatory power.

Regression 8 includes capitalized expensed investment in an adjusted book value measure, $TotB_t/P_t = (B_t + CapExpInv_t)/P_t$. The measure loads, but B/P, retained in the regression, now has no additional explanatory power. This agrees with the finding in Kapons (2020) where a book-to-price ratio adjusted for capitalized R&D has a higher association with future returns than the unadjusted B/P. And it agrees with Li (2020) where book-to-price with added intangible assets subsumes the standard B/P in explaining returns. It is explained by the capitalized (expensed) investment conveying higher risk and return than balance-sheet assets, as indicated in Table 2, and again cautions against aggregating the two types of investment. Interestingly, the mean of this measure is above 1.0 in Table 1 (though somewhat lower at the median), indicating that the market discounts the capitalized expense component in the denominating price.

B/P Factor Returns

Panel B of Table 3 reports returns for portfolios formed from B/P and capitalized expensed investment relative to price, constructed in the same way as in the corresponding panel in Table 2. The High – Low (factor) annual return for capitalized expensed investment (0.116) is higher than that for B/P (0.091), though the t-statistic on mean difference between the two is only1.08. For the sum of the two book value measures in $TotB_t/P_t$, the High - Low return of 0.106 is only a little higher than that for B/P (the t-statistic on the mean difference is 1.78). There is little difference between the $TotB_t/P_t$ return and that for capitalized expensed investment (t-statistic = 0.29).

With $TotB_t/P_t$ subsuming B/P in Regression 8 in this table, one might think of replacing B/P with $TotB_t/P_t$ in factor models. Some papers have done, with the argument that tangible and intangible assets should be accounted for in the same way. Indeed, those papers report that doing so improves the "value premium" based on B/P and partially explains the decline in that premium in recent years. This is the issue of how accounting numbers are to be packaged into a factor number, which we take up in Section 4. Note, however, that the results here indicate that adjusting book value with expensed investment losses information about the different risk associated with the two types of investments.

Another accounting feature bears on book value in factor models: With double entry accounting, the calculation of book value also affects the calculation of earnings. That is clear from the accounting that omits investments from the balance sheet—it requires them to be expensed to the income statement. It is also clear that, if the expensed investments are added to book value, earnings must be correspondingly changed. Curiously, many factor models include B/P but not E/P with which B/P is codetermined and which, as a realization of outcomes at risk, potentially conveys risk. And E/P forecasts future earnings-to-price in the framing equation (2). Earnings do enter extant factor models via ROE and that leads to the next section which points to a mistake in factor design. And it leads to the resolution of the first point in this subsection: Under accounting principles, it is a high P/B rather than a P/B that indicates risk and return, seemingly in conflict with the evidence.

3.3 Accounting for ROE

In existing factor models, ROE is interpreted as a measure of return on investment, with *q* theory then justifying its inclusion on those models. But the accounting for ROE gives it a different interpretation. An economic measure of return on investment would compare the earnings from investment to an investment base. But, with conservative accounting for investment, earnings in the numerator of ROE is reduced by expensed investment, mixing investment with return on investment, a seeming perversion. Consequently, this investment is missing from the denominator. See Feltham and Ohlson (1995) and Zhang (2000) for these deterministic accounting features. Thus, with these seeming distortions, conjectures that higher accounting ROE results from a higher hurdle rate are off base, as are investment CAPM models that embrace ROE as measuring the productivity of investment.

However, the conservative accounting for investment conveys risk. A firm with low earnings (and thus a low ROE) due to expensed investment is potentially a high-risk firm—it is making investments that may not pay off. If earnings on the investment are realized, ROE is high because, due to the initial expensing, the investment is not in the book value base. Thus, a high ROE indicates risk realized and a lower expected return for risk. This contrasts with existing factor models where ROE is said to be positively related to risk.

The connection to the discount rate in theory is in Penman and Zhang (2020), and the empirical documentation that low (high) ROE under conservative accounting projects higher (lower) risk to the earnings outcomes in equation (2) is in Penman and Zhang (2021). The comparison of an early-stage bio-tech with a mature pharmaceutical is illustrative. The start-up with little (or no) revenue as yet but high expensed R&D and consequently low (or negative) earnings and low (or negative) ROE is risky—the R&D may not pay off. The mature pharmaceutical is realizing revenue and earnings from R&D, reporting a high ROE due to the R&D investment missing from the denominator—it is less risky *ceteris paribus.* Penman and Reggiani (2018) provide further examples.

However, there are subtleties involved.

First, under the deterministic leveraging equation,

$$ROE_t = RNOA_t + Leverage_{t-1}[RNOA_t - \ Net\ Borrowing\ Rate_t]$$

where $RNOA_t$ is the book return on net operating assets (the return on business operations, the enterprise return) and $Leverage_{t\text{-}1} = Net\ Debt_{t\text{-}1}/Equity_{t\text{-}1.}$ So, as leverage adds to risk and the expected equity return in theory, higher ROE indicates a higher expected return to the extent it is affected by leverage. The conservative accounting effect enters through RNOA in the opposite direction, with expensed investment reducing RNOA. So it is a question of which dominates.

Second, in many asset pricing models, ROE enters along with B/P, with ROE_t taken as an expectation of forward ROE at risk. However,

$$ROE_{t+1} \times \frac{B_t}{P_t} = \frac{Earnings_{t+1}}{B_t} \times \frac{B_t}{P_t} = \frac{Earnings_{t+1}}{P_t}.$$

That is, adding ROE to B/P recovers the forward E/P ratio (multiplicatively), and does so additively with logged variables in most characteristic regressions in asset pricing research. The forward E/P ratio gains its legitimacy as an indicator of the expected return from equation (2), consistent with the yield interpretation of E/P in Ball (1978). In support, E/P is positively associated with returns in many empirical tests as in Basu (1977, 1983). In most asset pricing models, this joint implication of B/P and ROE is not recognized; rather, the two numbers enter as separate contributors to explaining returns and the basis for separate, additive factors. Instead B/P implicitly drops out in the construction, with the focus effectively on E/P: Under the accounting principle whereby earnings and book values articulate in a double-entry system, the product, $ROE_{t+1} = \frac{Earnings_{t+1}}{B_t} = \frac{Earnings_{t+1}}{P_t} \times \frac{P_t}{B_t}$, is insensitive to how book value is measured: lower book value implies higher $\frac{Earnings_{t+1}}{B_t}$ by accounting construction without an effect on E/P.

Third, in tying ROE to expected returns, there are differences between positive and negative ROE. By the leveraging equation above, leverage decreases ROE for negative RNOA providing the borrowing rate is positive. Further, while ROE is typically positively correlated, negative current ROE_t has been documented to be less informative about future ROE—it is affected by

Characteristic regressions

ROE enters regressions in Table 4, Panel A with positive ROE firms, Panel B with negative ROE. In contrast to the positive association of ROE with returns typically inferred, Regression 1 reports a slightly negative correlation in Panel A (unconditionally). For negative ROE in Panel B, the mean coefficient is positive, indicating that less negative ROE projects higher expected future earnings in equation (2), distinguishing temporary losses from those expected to be more permanent. Regression 2 adds B/P which loads with a positive coefficient, and ROE is now positively associated with returns in both panels (conditionally). That has prompted the addition of an ROE factor to a B/P factor in factor models.

The positive correlation of ROE with returns has always been observed with B/P in the regression, as here. However, by the ROE x B/P = E/P calculation, ROE with B/P implies E/P, an indicator of the (future) earnings yield on current price. Regression 3 thus has E/P alone, with a significant positive correlation for profit firms: E/P supplies information in B/P and ROE. That fits with the framing equation (2): Current earnings denominated in price indicates expected future earnings denominated in the same price (discount), suggesting that E/P conveys r. With earnings highly serially correlated that is likely. And it suggests an E/P factor rather than the two separate, additive B/P and ROE that jointly imply E/P. The mean negative coefficient for loss firms also projects higher returns (with the negative E/P).

Regression 4 adds ROE to E/P rather than to B/P. If ROE is depressed by conservative accounting for investment, it conveys risk and return under accounting principles and thus is a commentary on the earnings in E/P. For profit firms, the mean coefficient on ROE is negative; a lower (higher) ROE for a given E/P indicates higher (lower) returns. As ROE and E/P have the same numerator, the only number added in Regression 4 over Regression 3 is book value. But, rather than entering in B/P, book value enters as the denominator of earnings in ROE, and that indicates risk under conservative accounting: As that accounting results in lower book value when investment is expensed, realized earnings on low book value indicates lower risk—the investment omitted from book value as particularly risky has paid off. For loss firms, the mean coefficient on earnings is negative, projecting higher returns for higher losses, but that on ROE is positive, indicating that, given negative E/P, higher (less negative) ROE project higher future earnings in equation (2) with lower risk.

Regression 5 adds B/P to Regression 4, to assess whether it now adds to explaining returns. B/P loads for profit firms but ROE now adds little explanatory power. This is explained by B/P = E/P x 1/ROE. That is, given E/P, adding B/P is equivalent to adding the inverse of ROE. In short, book value coveys information about risk and return in conjunction earnings, not as a separate feature relative to price, but by information in the denominator of ROE. For loss firms, B/P also takes from the coefficient on ROE. This not only revises the standard interpretation of ROE in extant models, but also explains why B/P enters.

With current ROE_t in the regression, the relation ROE x B/P = E/P is an approximation. Precisely, $\frac{Earnings_t}{B_{t-1}} \times \frac{B_t}{P_t} = \frac{Earnings_t}{P_t} \times \frac{B_t}{B_{t-1}}$. So Regression 6 adds $\frac{\Delta B_t}{B_t}$ to Regression 4. That takes away from the coefficient on ROE_t. But $\frac{\Delta B_t}{B_t} = \frac{Earnings_t}{B_{t-1}} - \frac{Net\ Dividends_t}{B_{t-1}}$ by the clean-surplus relation, so the negative mean coefficient on $\frac{\Delta B_t}{B_t}$ reflects ROE_t. ΔB_t is also reduced by net dividends (dividends plus share repurchases less share issues) which increase leverage *ceteris paribus*, further contributing to the negative coefficient on $\frac{\Delta B_t}{B_t}$.

A low ROE in the cross-section can be due to conservative accounting that expenses investment, and that is the attribution given to coefficients on ROE in the regressions here. However, a low ROE can also be due to poor ex post performance. So, Regression 7 adds the expensed investment which is the conservative accounting effect on the numerator of ROE. The two are negatively correlated, -0.23 on average. The $ExpInv_t/P_t$ measure loads with a positive mean coefficient for both profit and loss firms, as expected of a variable that conveys risk. The expensed investment variable takes away some of the explanatory power of ROE for profit firms, reinforcing the attribution to the conservative accounting effect on ROE. For loss firms, the ROE coefficient remains positive, supporting the conjecture that ROE in this case conveys information of the extent to which losses are transitory.

Regression 8 adds Beta, Size, and momentum (MOM) to Regression 7. With the exception of Size, these variables do not load, although R^2 increases. The mean coefficients on ROE remain similar to those in Regression 7. The final two regressions involve ROE, Investment, and Size which together enter *q*-investment models. In Regression 9, the mean coefficient on ROE is positive, consistent with *q*-theory that predicts that, controlling for investment, ROE indicates higher risk and return. However, with the addition of E/P and expensed investment in Regression 10, the coefficient on ROE goes to zero for profit firms. The positive coefficient on ROE survives only with loss firms (negative ROE) in Panel B where Investment is no longer significant.

Factor Returns

Panel C of Table 4 reports potential factor returns based on the pertinent variables in the regressions in Panels A and B. The long minus short return for positive E/P is positive, consistent with the regressions, though not so for negative E/P. In contrast, the corresponding return for ROE is not significantly different from zero: The positive ROE factor return in many extant models is conditional on B/P, as in Regression 2 in the table. However, both earnings and ROE are reduced by expensed investment which conveys expected returns in the regressions, and here the mean zero-net-investment return for $ExpInv_t/P_t$ is a significant 11.7%. So, while a higher E/P indicates a higher expected return on average, a lower E/P reduced by expensed investment indicates a higher expected return.

The accounting informs that these are not stand-alone factors. The message from Regression 4 in Panel A is that E/P in combination with ROE_t conveys expected return. But, further, expensed investment combines with that information in Regression 7. The issue, then, is how these variables are to be combined into a factor model, the issue for Section 4.

3.4 Accounting for Alternative Profitability Measures

Fama and French models involve alternative profitability measures to ROE. Table 5 adds the operating profitability measure, OP_t/B_{t-1}, and the cash-flow profitability measure, CP_t/TA_{t-1}, to selected regressions in Table 4 to see if they add information about expected returns to that in ROE and the other accounting measures there. The comparison with ROE_t is intriguing because the measures exclude expensed R&D, though include SG&A which may involve expensed investment. The mean cross-sectional rank correlation between CP_t/TA_{t-1} and ROE_t in Table 1 is 0.53, and 0.86 for OP_t/B_{t-1}.

With this positive correlation, the operating profitability measure in Regression 1 in Table 5 returns similar coefficients to ROE in Table 4 in both panels. Further, OP_t/B_{t-1} adds little explanatory power to E/P and ROE in Regression 4, and the positive coefficient on E/P and negative coefficient on ROE are little changed from that in the corresponding regression in Table 4. Moving across the panels to Regression 7 (that includes expensed investment) and Regression 10 (with other Fama and French features), OP has little additional explanatory power for profit firms.

Not so for the cash-flow profitability measure. In Regression 2, 5, and 8, it adds positively to returns in both panels, with and without ROE and expensed investment in the regressions. Further, in Regression 11 it adds to the other features in FF models, supporting it as the choice of profitability measure in Fama and French (2018) as fitting best to returns. We also investigated the Novy-Marx (2013) gross profitability measure. Like CP_t, that loaded with significant positive coefficients in regressions similar to Regressions 5 and 8, but not in Regression 11.

As these profitability measures exclude R&D, we suspect they might refine E/P as an expectation of the future earnings yield in equation (2).[10] So, Regression 12 drops E/P from Regression 5, leaving CP_t profitability and ROE_t. Cash-flow profitability takes up the positive coefficient on E/P for profit firms while that on ROE remains negative, though not for loss firms. This survives with $ExpInv_t/P_t$ added in Regression 13. Regression 14 denominates CP_t in P_t rather than TA_{t-1}, a measure of yield to price, and the coefficient on CP_t/P_t is 0.015 with a t-statistic of 5.99 for profit firms and also significantly positive for loss firms. With expensed investment still in the regression, that is telling. Perhaps CP_t/P_t is a yield that is closer to a yield from investment than E/P and accordingly is a better indicator of the expected earnings yield in equation (2). Yet why the exclusion of short-term accruals (but not long-term accruals) would contribute is unclear.[11]

[10] To be frank, we do not quite understand what OP_t and CP_t are capturing, except to report that they are not capturing the risk and return conveyed by conservative accounting. CP_t is nominated as a cash flow measure. It adjusts operating earnings for short-term accruals but not depreciation and amortization (accruals) and long-term accruals such as those for pensions, long-term receivables, and compensation. (Long-term deferred revenues are taken out). By excluding accruals it omits aspects of earnings realizations—sales on credit, expenses incurred but not yet paid for, and deferred revenues not yet realized, for example. Gross profitability is clearer, the immediate profit from revenues from trading products and services over their costs. But leaving out other expenses to generate this revenue is at odds with a full accounting for earnings from revenues; these are not clean-surplus earnings. OP_t includes interest, very much like ROE. Neither of the measures include taxes but effective tax rates vary considerably with tax planning in the cross-section…..and we have to pay our taxes!

[11] Ball, Gerakos, Linnainmaa, and Nikoleav (2020) introduce retained earnings relative to price, showing that it forecasts returns better than B/P. This measure of accumulated (average) past earnings could be a better predictor of the future earnings yield than a one period of earnings yield, but it did not perform well in regressions relative to the profitability metrics in Table 5. This might be expected given the metric reduces past earnings for payout whereas the target is equation (2) is cum dividend.

Regression 3 introduces another profitability measure,

$$\text{ROE}_t^{\text{Adj}} = \frac{(\text{Earnings}+\text{Expensed Investment}-\text{Amortized CapExpInv})_t}{\text{TotB}_{t-1}}.$$

This measure reverses the conservative accounting effect on both earnings and book value to estimate a book rate of return unperturbed by the accounting. The denominator is adjusted book value that includes capitalized expensed investment and the numerator replaces expensed investment in earnings with the amortization of their capitalized value. For profit firms, the correlation with returns is negative, zero for loss firms. It adds little when added in the other regressions, indicating that "correcting" the accounting does not produce a profitability measure that better explain returns.

The result for this adjusted-ROE measure adds commentary on substituting TotB_t/P_t for B/P as the basis for a factor. In the discussion of the TotB_t/P_t measure in Table 2, it was pointed out that aggregating expensed investment with reported book value losses information about the different risk associated with the two types of investment that conservative accounting identifies. Here, the corresponding profitability measure under double-entry accounting does not add information to compensate. The adjustment for conservative accounting loses information about risk and return, both in the balance sheet and the income statement.

Factor Returns

Panel C of Table 5 repeats the portfolio construction in the corresponding panel in Table 4 but now with the alternative earnings measures relative to price, cash flow profit yield, CP_t/P_t, and operating profit yield, OP_t/P_t. These are the measures of cash flow profitability and operating profitability in Fama and French models except the denomination is in price rather that lagged balance sheet numbers. Both measures report significant long minus short returns in the panel, higher that those with E/P in Table 4, though the difference is negative for negative profits.

3.5 Size?

In Tables 2 - 5, firm size adds explanatory power. The "size effect in returns" is not well explained in the literature, at least in terms of fundamentals. Size is not an accounting number, but could it be associated with information about risk conveyed by the accounting? That makes sense: Small firms are those investing intensively in R&D, advertising, software, and other expensed intangibles which $ExpInv_t$ measures. But $ExpInv_t$ is an imperfect measure to which size might add to proxy for this investment. Further, small firms are also those where significant revenues from these investments are yet unrealized, still at risk. In contrast, large firms—a mature pharmaceutical rather than an early-stage biotech, for example—are those realizing earnings and thus less risky. This raises the specter that the size premium in returns is explained by the extent of conservative accounting, a hypothesis investigated in the affirmative in Penman, Reggiani, Richardson, and Tuna (2018) and Penman and Reggiani (2018).

We partitioned positive ROE firms each year into large (highest 1/3 in a size ranking) and small (lowest 1/3) and found that, for Regression 4 in Table 4, the mean coefficient on ROE for small firms is -0.021 (t = -3.93) while that for large firms is only -0.003 (t = -0.90).[12]

[12] This points to a resolution of a puzzle in Fama and French (2015), p. 16: "The biggest problem for the five-factor model in Table 11 is the portfolio of small stocks in the lowest profitability and highest investment quartiles."

Many papers exclude micro-caps when constructing factors. So we re-estimated Table 4 excluding firms below the 20th percentile of NYSE firms ranked on market capitalization. For loss firms, results were similar. For profit firms, the mean coefficient on ROE in Regression 4 was only -0.003 (t = -1.06) but all other results were similar. This reinforces that the accounting effect is stronger in smaller firms. The next section carries the investigation further.

4. Packaging Accounting Measures into a Factor Model

By recognizing accounting principles that pertain to risk, the preceding analysis has identified potential factors that imbed those principles. How should a factor model be built from these observations? In many extant models, each factor enters as a separate, additive feature said to add to the explanation of priced risk.[13] Is this the way to construct factor models?

The principles of accounting say: No. Accounting numbers are produced jointly in the double-entry system, not as independent features, and they jointly convey risk and return under accounting principles. So, book value does not enter as a stand-alone measure potentially conveying risk, but with the earnings with which it is co-determined (as above). And ROE enters, not as a basis for an additional independent factor but as a measure involving both earnings and book value that are jointly determined, conveying risk and return together (as the characteristic regressions confirm). This calls for a different packaging of the relevant accounting numbers into a factor model, to which we proceed. There is no claim that we are producing the definitive model; we are just bringing evidence of how accounting principles applied in a double-entry system can be exploited to build factor models.

[13] Factor interactions are involved in Duan, Gong, and Qi (2021). In *q*-investment models, the investment and profitability factors are conditional on each other.

The starting point is equation (2) where the expected return, *r*, is given by the expected cum-dividend earnings yield; investors buy expected earnings but those earnings, yet unrealized, are at risk, requiring a discount to the price in the denominator of the equation. Accordingly, information that conveys the expected earnings yield indicates the expected return, explicitly in equation (3). Factors here are constructed from such accounting information exhibited in the characteristic regressions.

4.1 SUMIS: An Income Statement Factor Based on Earnings, ROE, and Expensed Investment

Current earnings-to-price, E/P (implied by ROE x B/P in FF models) is denominated in the same price as in equation (2), so is a potential indicator of expected earnings-to-price and of *r*. Indeed, if current earnings were a sufficient statistic for expected future earnings in the numerator of equation (2), the matter would end there (and the E/P factor return in Panel C of Table 4 would say it all). However, while earnings are positively serially correlated, expected earnings typically differ from current earnings. Further, current earnings are reduced by expensed investment that produces the expected future earnings, and that investment conveys risk under conservative accounting. And, under double entry, the book value affected by the conservative accounting is also relevant, the book value in the denominator of ROE.

This suggests factor construction along the following lines. With E/P as a starting point, distinguish low versus high ROE for a given E/P. That incorporates effect of conservative accounting on the earnings in E/P (via the ROE numerator) but also the effect on book value (via the denominator); a high ROE indicates earnings realized on previously expensed investment that produced the low book value—risk resolved—and a low ROE indicates risky expensed investment with future earnings still at risk, yet to be realized. But a low ROE can also be due to poor ex post performance. So, within those ROE groups, identify the extent to which ROE is affected by the expensing of investment. Regressions 4 and 7 in Table 4 so inform.

So, each year at March 31, firms are sorted into four portfolios, one with all negative E/P and three from a ranking on positive E/P, 1/3 each. Then, within each of those portfolios, firms are sorted on ROE, with the top and bottom third designated high and low. Finally, within each of those E/P-ROE portfolios, firms are sorted on $ExpInv_t/P_t$ with the top and bottom third again designated high and low. Importantly, these are nested sorts to capture the joint accounting effects for the same set of firms, not the independent sorts of extant models.

Panel A of Table 6 reports mean annual excess returns over years for portfolios formed from the first two sorts, on E/P and then ROE. The returns are increasing in E/P (across rows), though those for negative E/P are higher than the low positive E/P. For a given E/P, returns are decreasing in ROE (down columns); ROE adds to information about expected returns to that in E/P, with the difference between high ROE in row 1 and low ROE in row 3 statistically significant for all positive E/P. Beginning-of-year book value, the denominator of ROE, is the added information here; a higher (lower) book value (due to conservative accounting) relative to the earnings in the first sort indicating higher (lower) returns.[14] The panel reports (at the far right) the mean return difference between the high positive E/P portfolio each year with low ROE (superscripted a) and the low positive E/P portfolio with high ROE (superscripted b), the return difference across the diagonal spread in the table. The mean return spread (with zero net investment) is a significant 7.6%, indicating payoffs to risk differences in the portfolios.

[14] Tests confirm that the ranking on ROE is not just a further ranking on E/P. Leverage adds to risk and return and to ROE, so the negative correlation between ROE and returns cannot be attributed to the leverage effect on ROE. But note that leverage (risk) is built in: Leverage increases both E/P and ROE. See Penman, Reggiani, Richardson, and Tuna (2018) for the proof.

Implicitly, the portfolio formation involves B/P: Because B/P = E/P x 1/ROE (approximately), a ranking on ROE is an inverse ranking on B/P for a given E/P. (The book value is beginning book value but B/P ratios are strongly serially correlated.) Thus B/P, so ubiquitous in asset pricing models, enters, but in a quite different way. Double entry (whereby earnings and book articulate) requires earnings and book value to enter jointly, as here.

Panel B adds the third sort on expensed investment, $ExpInv_t/P_t$, to that in Panel A. Returns in row 1 are for high ROE but with low $ExpInv_t/P_t$, deemed relatively low risk by accounting principle, and those in row 3 for low ROE but high $ExpInv_t/P_t$, deemed high risk. With realized earnings, book value, and expensed investment involved, both the realization principle and conservative investment principle are incorporated with both income statement and balance sheet effects. The return difference between the two is reported as rows 3 – 1. The t-statistics indicate that the differences are significant, enhanced from Panel A, with those for negative E/P now positive (but not large). The far right of the panel reports the mean return difference on the diagonal spread over positive E/P portfolios, a – b. As negative E/P cannot indicate a positive future earnings yield—the expected yield on a positive price there must be positive—the construction is limited to portfolios with positive E/P. The mean annual return spread is 13.3%, indicating significant risk differences between the long and short positions of a zero-net-investment factor formed on the accounting information.

We refer to this summary accounting factor as SUMIS as it is based on earnings from the income statement. Reservations about the crude estimate of expensed investment in SG&A prompted the same portfolio sort but with that component excluded from expensed investment, that is, with the $ExpInv_t/P_t$ (w/o SG&A) measure. Results are similar, with a mean return spread of 15.7% for the summary factor.[15]

4.2 SUMBS: A Balance Sheet Factor Based on Investment and Financing

SUMIS involves the accounting for current earnings that add to lagged booked value. But there is also a concurrent, time t, accounting in the balance sheet that has no effect on contemporaneous earnings: The accounting for investment activities and financing activities. As explained in Section 3.1, balance-sheet investment is the realization of uncertain investment opportunities that projects future realized earnings in equation (2). The accounting distinguishes that investment from more risky investment expensed against earnings. As a realization, it conveys lower risk and return, and the market prices it as such as Table 2 confirms. However, leverage via debt versus debt financing of investment adds to risk to future earnings in equation (2), so both must be incorporated. As double-entry accounting produces both a balance sheet and an income statement, the accounting for risk is not complete without this accounting in the balance sheet. This section forms a factor, SUMBS, that recognizes this accounting.

Most factor models such as Fama and French (2015 and 2018) and Hou, Xue, and Zhang (2015) measure investment as the change in total assets. However, as recognized in Cooper,

[15] Conservative accounting for investment does not affect earnings if the amount expensed is equal to the amortization charge that would be made if the investment were capitalized to the balance sheet and then amortized. This steady-state case occurs with no growth in the investment being expensed. So portfolios were also formed with the expensing variable in the third sort as $(ExpInv_t - ExpInv_{t-1})/P_t$, with little difference in results. Portfolios were also formed with the expensing variable as $\Delta CapExpInv_t/P_t$ which equals zero when expensed investment equals amortization, Again, results were similar. Table 5 indicated that the cash flow profitability measure, CP_t/P_t, performed relatively well, as did a CP_t/P_t factor, suggesting it might be a better indicator of the future earnings yield than E/P. So, the portfolio formation was repeated with CP_t/P_t, replacing E/P in the first sort. The return over the diagonal spread, a – b, was 17.4% (t = 2.24), higher than with E/P. However, this is just a finding in the data so, as we are unclear as to what CP_t/P_t captures (in the discussion of Table 5), we stick with the E/P sort. In any case, the finding does suggest incorporating a measure that conveys the expected yield better than E/P, a measure of "sustainable earnings."

Galen, and Ion (2017), investment so defined is not the investment in business assets such as inventory and plant. Indeed, while including that investment change in total assets also involves the change in cash and cash investments reduce cash with zero effect on total assets. So, Table 7 forms the balance-sheet factor based on investment in the net operating assets (in the business), ΔNOA.

Each March 31, firms are formed into three portfolios based on $\Delta NOA_t/P_{t-1}$ then within each $\Delta NOA_t/P_{t-1}$ portfolio, on financing leverage, $Debt_t/P_t$. $\Delta NOA_t/P_t$ ranks forward returns negatively in Table 7 while leverage is positively (though not strongly) related to returns. The mean return over years for SUMBS, calculated from the a – b diagonal, is 0.099 with a t-statistic of 3.06.[16]

4.3 A Size Factor?

Both the Fama and French and *q*-investment models include a factor based on firm size (equity market capitalization). That is supported by correlations with returns in the data but without a clear explanation. Size is not an accounting number: Should that factor be added to SUMIS and SUMBS? The answer would be No if size is picking up the risk conveyed by those numbers. Findings in Penman, Reggani, Richardson, and Tuna (2018) and Penman and Reggani (2018) suggest so, as do the results for the characteristic regressions in Table 4 for large versus small firms.

Table 8 investigates further. Panel A reports the median ln(Size) of firms in the portfolios underlying SUMIS in Table 6, showing that firm size varies over the accounting features underlying those portfolios. There is some negative correlation between size and E/P (across

[16] Similar results were observed with the investment calculated as (Change in gross property, plant, and equipment and inventory)$_t/P_{t-1}$ (yielding SUMBS = 0.072, t = 2.65). With investment as $\Delta TA_t/P_{t-1}$, SUMBS = 0.093, t = 2.21.

rows), but the significant differences are down columns: Small firms, prominent in row 3, are associated with lower realized ROE and higher expensed investment than large firms, on average. This accords with the mean correlations in Table 1 where the rank correlations of size with ROE is 0.41 and that with $ExpInv_t/P_t$ is -0.41.[17] This makes sense: Small firms typically are investing intensively in R&D, advertising, and other intangibles relative to the (low) earnings they are currently realizing, and thus are higher risk, while large firms are realizing earnings from those investments, thus lower risk. Indeed, large firms are typically once-small firms that became successful with realized earnings, whereas small firms are those where those investments have not yet paid off (and may not). Think again of the start-up biotech versus the mature pharmaceutical.

So, Panel B of Table 8 excludes microcaps (below the 20^{th} size percentile of NYSE stocks) from the analysis in Panel B of Table 6 that yields the return to SUMIS. Results are not as strong, emphasizing that the excluded small firms are those with the features that convey higher risk. These findings indicate that the size premium in returns is partially explained by exposure to risk conveyed by the accounting; it is not an opaque size factor than explains returns but rather that exposure.

However, while the risk conveyed by the accounting is more associated with small firms, it is also associated with large firms with the same accounting characteristics. The portfolios in Panel C of Table 8 are the same as those in Panel B but the returns in the portfolios formed on ROE and $ExpInv_t/P_t$ (down columns) are only for firms that report the accounting measures that meet the cut-offs for those portfolios. The high minus low returns in the 3 – 1 row are significant as is

[17] The correlations with expensed investment in Table 1 are for $ExpInv_t/P_{t-1}$. Those reported in this section are for $ExpInv_t/P_t$.

the diagonal mean return spread of 0.091. It is not just a matter of the accounting measures explaining returns for small versus large firms but for any firm, small or large, with the relevant accounting numbers; size is a crude indicator of the differing risk between small and large firms.

These findings, along with those for Table 4 characteristic regressions partitioned on size, indicate that firm size proxies for risk conveyed by SUMIS. Comparative factor tests below further confirm. We have not documented the extent to which size returns are explained by SUMBS but that also is observed below.

4.4 A Tentative Factor Model

Adding the income statement and balance sheet factors to the excess return on the market portfolio, MKT, yields the following factor model:

$$R_{it} - R_t^{RF} = \alpha_{it} + \beta_{MKT,i} MKT_t + \beta_{SUMIS,i} SUMIS_t + \beta_{SUMBS,i} SUMBS_t + \varepsilon_{it}. \qquad (4)$$

where $R_{it} - R_t^{RF}$ is the excess return over the risk-free rate for asset i for period t. The market factor appears in most extant factor models. The added accounting factors represent extra-market risk or capture part of market risk if the proxy market portfolio does not represent the market portfolio of all assets adequately.

5. A Comparison of Accounting Factors with Extant Factor Models

We now compare this model with standard factor models for explaining returns. If (4) is a valid model, $\alpha_{it} = 0$, all i and t, the focus of the tests.

5.1 Standard Factor Returns Do Not Explain the Accounting Factor Returns nor Component Portfolio Returns

Table 9 reports results from time-series regressions of monthly excess returns for SUMIS and its underlying portfolios on factors in the Fama and French (FF) six-factor model. Panel A is for the high-risk portfolios within each E/P portfolio, those with low ROE with high expensed investment on row numbered 3 in Panel B of Table 6. Panel B is for the relatively low-risk portfolios, high ROE with low expensed investment on row numbered 1. Panel C is for the long minus short returns, row 3 minus 1. In all panels in Table 9, there is a portfolio consisting of all positive E/P stocks in the respective rows, labelled ALL In the far-right column in Panels A and B, the return is the difference between the return on the high positive E/P and low positive E/P in the row. In Panel C, it is the return on SUMIS.

In these so-called spanning tests, one might interpret the results as FF models explaining the returns for the accounting portfolios. Indeed, they do (in part) as a matter of correlation. But, with these models identified from correlations in the data, with the underlying risk they capture unexplained, that is not very illuminating. Rather, we take the analysis to ask if those factors capture the (explained) risk conveyed by the accounting underlying the construction of the accounting portfolios.

For the high and low risk portfolios in Panels A and B, the FF factors explain a good part of the variation in returns, with R^2 around 85%. We leave the reader to examine the slope coefficients for the explanatory power of each FF factor, except to note that β_{SMB} on the SMB size factor is close to 1.0 for the high-risk portfolios in Panel A, indicating that the risk conveyed by size is similar that from ROE and expensed investment (though β_{SMB} is conditional on other factors in the regression). The coefficients are lower for lower-risk portfolios in Panel B. They remain positive in Panel C, with SMB positively correlated with the SUMIS return. This attributes the size premium to the risk associated with the three accounting measures, complementing the size analysis above.

The Intercepts (alphas)—returns unexplained by the factors—are significantly different from zero, more so for the high-risk portfolios in Panel A. The alphas are positive, typically about 1% per month in that panel, with the return for the summary portfolio with zero net investment (in the last column) also 1% per month on average with a t-statistic of 7.73. This indicates that the FF factors do not fully capture the risk exposure for E/P yields with high-risk (that is, low ROE and high expensed investment). That is also the case in the lower-risk portfolios in Panel B, though the intercepts are not as high, reflecting the lower risk.

In Panel C where the spread on ROE and expensed investment is introduced, the alphas are also significantly positive with the R^2 much reduced. The alpha for SUMIS is 0.7% per month (t = 4.42).[18]

Table 10 repeats the same analysis but now with the *q*-investment factors. Beside the market factor, those models have an investment factor, IA, a size factor, ME, an ROE factor, and, in a recent innovation, a factor based on predicted investment growth in the future that the authors call EG. Here again, the intercepts are significantly different from zero and about 1% per month

[18] Two issues arise in these regressions. First, the FF factors are constructed excluding microcaps whereas the portfolios based on accounting measures include them. This is by design, for it is the accounting treatment to convey risk and return that we wish to document and, if this is associated more with small firms, so be it. Indeed, Panels A and B of Table 8 and the findings in Table 9 here so indicate. However, Panel C of Table 8 demonstrates that the risk is also in a large firms with the relevant accounting measures, requiring a higher return which an investor would want to identify. We repeated the analysis in Table 9 rejecting microcaps. Results were similar for slope coefficients, and significant intercepts were estimated. However, those intercepts were lower, 0.003 and 0.004 for the long minus short portfolios in Panel C and 0.003 with a t-statistic of 2.24 for SUMIS. Second, FF portfolio returns are value weighted whereas those constructed with accounting measures are equally weighted. If the accounting in focus is more associated with smaller firms, value weighting weights away from the accounting feature on which the portfolio was formed so the portfolio return is not indicative of the risk and return associated with that feature. Indeed, value weighting can be viewed as protecting against some of the risk indicated by accounting numbers and associated with smaller firms, by investing in larger firms where, on average, that risk is lower. That said, the alphas in Table 9 show that the accounting still explains returns in these value-weighted FF factors.

for the higher-risk portfolios in Panel A. The alpha for SUMIS, in Panel C is 0.7% per month (t = 3.19).

We performed the same tests in Tables 9 and 10 with the balance-sheet factor, SUMBS. As both FF and *q*-investment models include an investment factor, albeit measured differently, we expected these models to perform relatively well in explaining return for SUMBS, but that is not so. For the FF factors, the alphas (t-statistics) for the long and short components of SUMBS (corresponding to those in Panel C of Table 9) are 0.008 (6.84) and 0.003 (4.11) per month, respectively, and that for the long minus SUMBS factor itself is 0.005 (4.37). The corresponding numbers for the *q*-investment model are 0.009 (6.38), 0.004 (4.85), and 0.005 (3.51).

5.2 The Accounting Factor Model Explains Returns on a Variety of Test Portfolios

Table 11 reports on time-series regressions like those in Tables 9 and 10 but with monthly excess returns for selected portfolios now regressed on factors returns for MKT, SUMIS, and SUMBS in model (4). The Panel A portfolios are long minus short portfolios constructed as potential factors earlier in this paper. The portfolios in Panel B are factors in standard factor models.

The intercepts are of primary interest. Alphas are not significantly different from zero for B/P, E/P, Size, and ROE (long – short portfolios) in Panel A, and that for TotB/P (that includes expensed investment) only 0.002 per month. Nor are alphas significant for HML and SMB in Panel B. The results for the Size and SMB are particularly compelling: The zero alpha indicates that the size premium reflects risk conveyed by the accounting factors, as indicated in the discussion in Section 4.3. The exceptions that stand out in Table 11 are the FF operating profitability factor, RMW, and the two factors from the *q* models, *q*-ROE and *q*-EG. Here alphas are relatively large, positive, and statistically significant.

The q-EG factor return is interpretable under accounting principles, indeed it validates them. In Hou, Mo, Xue, and Zhang (2020), this factor is calculated as the difference in returns between high forecasted growth in investment and low forecasted growth. Under Cochrane (1991) q-theory, higher expected investment growth (EG) requires a higher return as it indicates higher marginal q which equals marginal cost indicating the discount rate. Interpreted under accounting principles, future investment is investment not yet looked to the balance sheet, investment not yet realized, still at risk. That requires a higher return over that indicated by SUMIS and realized investment in SUMBS. Indeed, it complements SUMBS: While SUMBS books investment opportunities realized, EG captures investment growth opportunities still at risk.

So, both production-investment theory and accounting principles have the same prediction. Indeed, one can construe the accounting as consistent with production-investment theory. In any case, the finding that q-EG factor return is not explained by the accounting factor model (4) but is explained by accounting principles leads to an expected investment growth factor, EG, added to model (4):

$$R_{it} - R_t^{RF} = \alpha_{it} + \beta_{MKT,i} MKT_t + \beta_{SUMIS,i} SUMIS_t + \beta_{SUMBS,i} SUMBS_t + \beta_{EG,i} EG_t + \varepsilon_{it}. \quad (5)$$

As EG is predicted from time t accounting numbers in Hou, Mo, Xue, and Zhang (2020), it can be nominated as an accounting factor. Investment begets earnings so, with equation (2) defining the expected return, EG captures expected future earnings growth from the expected investment, earnings still at risk for that investment is still at risk. That redefines the q-EG factor.[19]

Panels C and D of Table 11 repeats the analysis in Panels A and B of the table with the added EG factor return. β_{EG} is significant and positive, except for the Size and SMB (size) portfolios. Intercept alphas are close to zero for all portfolios, but positive for the size portfolios. That

[19] We concede that the revised model (5) comes from rationalization after observing the data, not a priori. Nevertheless, EG is rationalized as an implication of accounting principles.

coincides with a negative β_{EG}: Conditional on the other factors, the relative returns for small firms are negatively correlated with the risk in anticipated investment.

6. Conclusions

This paper is offered with the aim of putting empirical asset pricing and the accounting on which it draws on the same platform. That is done by recognizing that accounting numbers are governed by accounting principles that connect the numbers to risk and expected return as a matter of pricing theory. This is so for book value, earnings, investment, return on equity (ROE), and other numbers that are the basis for factor construction in existing asset pricing models. In many of those models, these numbers enter from data mining, without much definition. The paper provides that definition. But, in doing so, it recognizes the property of the double entry accounting system under which these numbers are codetermined and convey information about risk jointly. That leads to a parsimonious packaging of the accounting numbers into a factor model that differs from the standard factor model construction and outperforms these models in explaining returns.

That packaging begins with the recognition that, in theory, the expected return is given by the expected future earnings yield on current price, for which the current earnings-to-price, E/P, is a predictor. However, it is an imperfect predictor, for current earnings are typically disturbed by accounting principles that convey the risk to expected earnings that discounts their pricing. By recognizing this accounting and combining it with the current earnings-to-price, the paper produces a factor that explains stock returns well relative to existing models. That is complemented with an investment factor based on complementary accounting in the balance sheet under double entry. Accounting measures in standard models, like book value, investment, and ROE enter, but in a different way —in a way that conveys risk and return under accounting principles.

Here are some specific observations arising from the accounting analysis and the confirming empirical tests:

- The current earnings yield, E/P, is important for indicating risk and return but does not explicitly appear in standards models
- As ROE and B/P appear in Fama and French models and E/P = ROE x B/P (approximately), those models capture E/P implicitly
- Consequently, book value enters into risk models, not as a stand-alone B/P factor but as the denominator of ROE, adding information about risk and return to E/P.
- While those models and *q*-investment models incorporate ROE with a positive relationship to returns, the accounting for ROE implies a negative relationship to returns
- The determining accounting is the expensing of particularly risky investment that reduces earnings and ROE (and other profitability measures). That leads to an accounting factor based on E/P, ROE, and expensed investment based on the income statement.
- Investment booked to the balance sheet, deemed less risky, is the realization of investment opportunities and thus conveys lower risk. That supplies an accounting justification for the investment factor in extant factor models, though requires a different measure of investment to change in total assets.
- The construction of factors that adds expensed investment to balance-sheet investment loses information about the differential risk in these two types of investment. So do factors that add book value from expensed investment to B/P.

- That balance-sheet investment combines with financing leverage to yield a balance-sheet factor additive to the income statement factor under double-entry accounting.
- Expected investment, unrealized and still at risk under accounting principles, indicates higher risk. This supplies an accounting justification for the expected investment growth factor in *q*-investment models.

The paper does not provide a definitive factor model. With the aim of making comparisons to existing models it has been limited to accounting measures appearing in those models. Accounting principles affect a wider set of accounting items that could be exploited to build a model. Those items add to the prediction of the future earnings yield over that in the current earnings in E/P. However, determined under accounting principles, that also convey the risk to that growth. Penman and Zhu (2022) pursues this, not only generating a factor model, but also casting it in terms of consumption-based asset pricing theory with accounting numbers connecting to consumption and the risk to consumption.

Appendix. Calculation of Variables

This appendix details the calculation of variables used in the empirical analysis. Compustat sometimes reports an item as missing. These are either in the financial reports as zero, aggregated with other items in the financial reports, or reported but not found by Compustat. We follow the practice of setting these to zero, relying on the finding in Casey, Gao, Kirschenheiter, Li, and Pandit (2016) that setting to zero satisfies certain articulating accounting equations 82.7% of the time. Nevertheless, the remaining 17.3% are a qualification to our analysis.

Ret_{t+1}	One-year, buy-and-hold return calculated as compounded CRSP monthly returns, starting at the beginning of the fourth month after the current fiscal year end. For firms that are delisted during the 12 months, the return for the remaining months is calculated by first applying the CRSP delisting return and then reinvesting any remaining proceeds at the risk-free rate. This mitigates concerns with potential survivorship biases. Firms that are delisted for poor performance (delisting codes 500 and 520-584) frequently have missing delisting returns (see Shumway 1997). We control for this potential bias by applying delisting returns of -100% in such cases. Results are qualitatively similar if we make no such adjustment. Accounting data for a fiscal year are presumed to have been published during the three months after fiscal-year end, as required by law (and before the beginning of the return period).
$Earnings_t$	Earnings before extraordinary items (Compustat item IB) and special items (Compustat item SPI), minus preferred dividends (Compustat item DVP), with a tax allocation to special items at the prevailing federal statutory corporate income tax rate for the year. The top statutory federal tax rate was 50% in 1964, 48% in 1965-1967, 52.8% in 1968-1969, 49.2% in 1970, 48% in 1971-1978, 46% in 1979-1986, 40% in 1987, 34% in 1988-1992, 35% in 1993-2017, and 21% in 2018.
B_t	Book value of common equity at the end of the fiscal-year, calculated as Compustat common equity (Compustat item CEQ) plus any preferred treasury stock (Compustat item TSTKP) less any preferred dividends in arrears (Compustat item DVPA).
P_t	Market value of equity three months after fiscal-year end. It is calculated as the number of shares outstanding multiplied by the price per share from CRSP at three months after fiscal-year end. It can also be calculated as the number of shares outstanding at the end of the fiscal year from Compustat (#CSHO) multiplied by the price per share from CRSP at three months after fiscal-year end, adjusted for any intervening stock splits and stock dividends. This excludes any change in the market price from net share issues over the three months when calculating E/P and B/P ratios based on earnings and book value at year end. Market price can also be calculated as per-share price at three months after fiscal-year end, adjusted for stock splits and stock dividends over the three months, in which case, earnings and book value are on a per-share basis. The first calculation is applied in this paper as that seems to be the practice for E/P and B/P calculations in asset pricing papers to which this paper compares.
TA_t	Total assets (Compustat item AT).

$Investment_t$	Change in Total Assets. This is the measure in Fama and French (2015 and 2018) five- and six-factor models and the Hou, Xue, and Zhang (2015) and Hou, Mo, Xue, and Zhang (2020).
$ExpInv_t$	$ExpInv_t = R\&D_t + Advertising_t + SG\&A_t \times 0.3$. R&D is Compustat item XRD, Advertising is Compustat item XAD and the SG&A variable is XSGA-XAD-XRD when XSGA is greater than the sum of XAD and XRD, otherwise SG&A equals XSGA. The modified measure, $ExpInv_t$ (w/o SG&A) excludes the SG&A component.
$TotInv_t$	Change in ($Total\ Assets_t + R\&D_t + Advertising_t + SG\&A_t \times 0.3$), with $SG\&A_t$ calculated as in $ExpInv_t$.
$CapExpInv_t$	Capitalized expensed investment, calculated as the sum of capitalized R&D, capitalized advertising expense, and capitalized SG&A x 0.3. R&D is capitalized with amortization rates estimated in Lev and Sougiannis (1996). Advertising expenses are capitalized using the sum-of-years digit amortization method with a two-year useful life, as indicated by the findings in Bublitz and Ettredge (1989) and Hall, Mansfield, and Jaffe (1993). SG&A is capitalized and amortized using the straight-line method with a five-year useful life.
$TotB_t$	Book value plus capitalized expensed investment, $B_t + CapExpInv_t$.
$Debt_t$	Financing debt at the end of the fiscal year calculated as Debt in current liabilities (#DLC) + Long-term debt (#DLTT) + Preferred stock (#PSTK) – Preferred treasury stock (#TSTKP) + Preferred dividends in arrears (#DVPA).
$Net\ Debt_t$	Net financing debt at the end of the fiscal year, the difference between financing debt (above) and financial assets. Financial assets are Cash and Short-term investments (#CHE) + Investments and advances-other (#IVAO).
NOA_t	Net operating assets at the end of the fiscal year, measured as $Net\ Debt_t$ (above) plus book value of common equity, B_t, with added non-controlling (minority) interest (Compustat item MIB).
ROE_t	$Earnings_t/B_{t-1}$, with $Earnings_t$ calculated as above.
ROE_t^{Adj}	$\frac{(Earnings + Expensed\ Investment - Amortized\ CapExpInv)_t}{TotB_{t-1}}$
GP_t	Gross profit, as in Novy-Marx (2013) where it is relative to total assets.
CP_t	Cash flow profit as in Ball, Gerakos, Linnainmaa, and Nikolaev (2018) where it is relative to lagged total assets. In Fama and French (2018) it is deflated by end-of-period book value.
OP_t	Operating profits as in Fama and French RMW factor. That is defined as revenue minus cost of goods sold minus SG&A minus interest on the Kenneth French web site, though it appears to include R&D in Fama and French (2018). On the web site OP_t is deflated by lagged book value, but in Fama and French (2018) it is deflated by end-of-period book value for the construction of RMW.
$Beta_t$	Stock market betas estimated from monthly stock returns during the previous five years ending in the third month after the current fiscal year end, or a lessor period if five years of data are not available.
$Size_t$	Logarithm of market value of $equity_t$.
MOM_t	Twelve month buy-and-hold return ending in the third month after the current fiscal year end.

References

Arnott, R., C. Harvey, V. Kalesnik, and J. Linnainmaa. 2020. Reports of value's death may be greatly exaggerated. *Financial Analysts Journal* 77 (1), 44-67.

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

Ball, R., J. Gerakos, J. Linnainmaa, and V. Nikolaev. 2015. Deflating profitability. *Journal of Financial Economics* 117, 225-248.

Ball, R., J. Gerakos, J. Linnainmaa, and V. Nikolaev. 2018. Profitability, cash flows and accruals in the cross section of stock returns. *Journal of Financial Economics* 121, 28-45.

Ball, R., J. Gerakos, J. Linnainmaa, and V. Nikolaev. 2020. Earnings, retained earnings, and book-to-market in the cross-section of stock returns. *Journal of Financial Economics* 135, 231-254.

Bansal, R., and A. Yaron. 2004. Risks for the Long-run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59, 1481-1509.

Barker, R., A. Lennard, S. Penman, and A. Teixeira. 2021. Accounting for intangible assets: Suggested solutions. *Accounting and Business Research*, forthcoming.

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

Basu, S. 1983. The relationship between earnings yield, market value, and return for NYSE stocks: further evidence. *Journal of Financial Economics* 12, 129-56.

Berk, J., R. Green, and V. Naik. 1999. Optimal investment, growth options, and security returns. *Journal of Finance* 54, 1153-1608.

Bongaerts, D., X. Kang, and M. van Dijk. 2021. Intangible assets and the cross-section of returns. Unpublished paper, Rotterdam School of Management, Erasmus University. At https://ssrn.com/abstract=3927990.

Bublitz, B., and M. Ettredge. 1989. The information in discretionary outlays: Advertising, research, and development. *The Accounting Review* 64 (1):108-124.

Casey, R. J., F. Gao, M. T. Kirschenheiter, S. Li, and S. Pandit. 2016. Do Compustat financial statement data articulate? *Journal of Financial Reporting* 1 (1): 37-59.

Cho, T., L. Kremens, D. Lee, and C. Polk. 2021. Scale or yield? A present-value identity. Unpublished paper, London School,of Economics, University of Washington, and UCLA Anderson. At https://ssrn.com/abstract=3857368.

Cochrane, J. 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance* 46¸ 209-237.

Cooper, M., H. Gulen, and M. Ion. 2017. The use of asset growth in empirical asset pricing models. At https://ssrn.com/abstract=3026534.

Duan, Z., Z. Gong, and Q. Qi. 2021. Factorization asset pricing. Unpublished paper, Peking University. At https://ssrn.com/abstract=3940074.

Enache, L., and A. Srivastava. 2018. Should intangible investments be reported separately or commingled with operating expenses? New evidence. *Management Science* 64 (7): 3446-3468.

Eisfeldt, A., and D. Papanikolaou. 2013. Organizational capital and the cross-section of expected returns. *Journal of Finance* 67 (4), 1365-1406.

Eisfeldt, A., E. Kim, and D. Papanikolaou. 2020. Intangible value. Unpublished paper at https://ssrn.com/abstract=3720983.

Epstein, L., and S. Zin. 1989. Substitution, Risk aversion, and the temporal behavior of consumption and asset returns. *Econometrica* 57, 937-969.

Dechow, P., R. Sloan, and M. Soliman. 2004. Implied equity duration: A new measure of equity risk. *Review of Accounting Studies* 9, 197-228.

Fama, E., and K. French. 1993. Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E., and K. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116, 1-22.

Fama, E., and K. French. 2018. Choosing factors. *Journal of Financial Economics* 128, 234-252.

Fama, E., and K. French. 2020. Comparing cross-section and time-series factor models. *Review of Financial Studies* 33, 1891-1926.

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

Hall, B. H., E. Mansfield, and A. B. Jaffe. 1993. Industrial research during the 1980s: Did the rate of return fall? *Brookings Papers on Economic Activity. Microeconomics* 1993 (2):289-343.

Hou, K., H. Mo, C. Xue, and L. Zhang. 2021. An augmented *q*-factor model with expected growth. *Review of Finance* 25 (1), 1-41.

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

Kapons, M. 2020. R&D expenses, R&D capitalization, the book-to-price ratio and the cross-section of returns. Working paper, Tilburg School of Economics and Management, Tilburg University. Presented at the Accounting Design Project Forum. At https://www8.gsb.columbia.edu/ceasa/sites/ceasa/files/20200918_ADP_Kapons.pdf.

Lev, B., and Sougiannis, T. (1996). 'The capitalization, amortization, and value relevance of R&D'. *Journal of Accounting and Economics*, *2* (February): 107-138.

Lev, B., and A. Srivastava. 2020. Explaining the recent failure of value investing. Working paper, New York University and University of Calgary.

Li, F. 2020. Intangibles: The missing ingredient in book value. Paper, Research Affiliates, LLC. At https://ssrn.com/abstract=3686595.

Merton, R. 1973. An intertemporal asset pricing model. *Econometrica* 41, 867-887.

Mullins, G. 2020. Equity duration. Unpublished paper, St. Cross College, University of Oxford.

Novy-Marx, R. 2013. The other side of value: The gross profitability premium. *Journal of Financial Economics* 108, 1-28.

Oh, H., and S. Penman. 2020. Income statement mismatching has not reduced the information conveyed by accounting over time. Unpublished paper, Columbia University and University of Washington, Bothell. At https://ssrn.com/abstract=3778173.

Ohlson, J. 1995. Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research* 12, 661-687.

Oswald, D., A. Simpson, and P. Zarowin. 2020. The information benefits of R&D capitalization. Woking paper, University of Michigan, London School of Economics, and New York University. At https://ssrn.com/abstract=

Penman, S., and F. Reggiani. 2018. Fundamentals of value vs. growth investing and an explanation for the value trap. *Financial Analysts Journal* 74, 102-119.

Penman, S., F. Reggiani, S. Richardson, and İ Tuna. 2018. A framework for identifying accounting characteristics for asset pricing models, with an evaluation of book-to-price. *European Financial Management* 24, 488-520.

Penman, S., and N. Yehuda. 2019. A matter of principle: The identification of cash-flow news and discount-rate news in financial statements. *Management Science* 65, 5584-5602.

Penman, S. and X. Zhang. 2020. A theoretical analysis connecting conservative accounting to the cost of capital. *Journal of Accounting and Economics* 69, 1-25.

Penman, S. and X. Zhang. 2021. Connecting book rate of return to risk and return: The information conveyed by conservative accounting. *Review of Accounting Studies* 26 (1), 391-423.

Penman, S., and J. Zhu. 2022. An accounting-based asset pricing model and a fundamental factor. Forthcoming, *Journal of Accounting and Economics*. At https://ssrn.com/abstract=3393068.

Peters, R., and L. Taylor. 2017. Intangible capital and the investment-*q* relation. *Journal of Financial Economics* 123, 251-272.

Rizova, S., and N. Saito. 2020. Intangibles and expected stock returns. Unpublished paper, Dimensional Fund Advisors.

Shumway, T. (1997). The delisting bias in CRSP data. *Journal of Finance* 52, 327-340.

Weber, M. 2018. Cash Flow Duration and the term structure of equity returns. *Journal of Financial Economics* 128 (3), 486-503.

Zhang, X. (2000). Conservative accounting and equity valuation. *Journal of Accounting and Economics* 29, 125-149.

Table 1: Distribution of Variables and Correlation Matrix

Panel A summarizes average of annual cross-sectional distributions from 1964 to 2018 for variables appearing in the paper. For the calculation of means and standard deviations, variables are trimmed at the top and bottom one percent, each year, except for returns. N is the average number of observations per year. Panel B reports means of annual cross-sectional correlation coefficients across the sample years. Pearson correlations are presented in the upper off-diagonal and Spearman correlations in the lower off-diagonal. For the Pearson correlations, all variables are trimmed at the top and bottom 1 percent, except for returns. R^{RF} is the risk-free rate. Variables are defined in the appendix.

Panel A: Distributions of Variables

Variable	N	Mean	Median	St. Dev.	P5	P25	P75	P95
$Investment_t/TA_{t-1}$	3,163	0.213	0.078	0.579	-0.233	-0.019	0.227	1.120
$Investment_t/P_{t-1}$	2,907	0.083	0.075	0.499	-0.566	-0.032	0.208	0.759
$ExpInv_t/TA_{t-1}$	3,078	0.159	0.111	0.168	0.012	0.054	0.203	0.474
$ExpInv_t$ (w/o SG&A)/TA_{t-1}	2,345	0.091	0.043	0.138	0.000	0.013	0.108	0.349
$ExpInv_t/P_{t-1}$	2,830	0.202	0.123	0.246	0.014	0.059	0.243	0.669
$TotInv_t/P_{t-1}$	2,907	0.287	0.205	0.548	-0.300	0.079	0.414	1.176
B_t/P_t	3,106	0.794	0.618	0.688	0.129	0.356	1.004	2.059
$CapExpInv_t/P_t$	3,054	0.316	0.139	0.621	0.014	0.062	0.304	1.153
$CapExpInv_t$ (w/o SG&A)/P_t	2,494	0.089	0.034	0.159	0.000	0.010	0.094	0.367
$TotB_t/P_t$	3,106	1.120	0.804	1.159	0.194	0.472	1.332	3.089
ROE_t	3,035	-0.020	0.090	0.536	-0.778	-0.058	0.178	0.408
OP_t/B_{t-1}	3,035	0.200	0.237	0.563	-0.575	0.062	0.394	0.845
CP_t/TA_{t-1}	2,900	0.109	0.131	0.211	-0.234	0.048	0.210	0.377
E_t/P_t	3,216	-0.058	0.049	0.402	-0.603	-0.036	0.085	0.156
$\Delta B_t/B_{t-1}$	3,035	0.251	0.079	0.992	-0.439	-0.037	0.213	1.405
$CapExpInv_t/NOA_t$	2,943	1.117	0.357	3.112	0.035	0.163	0.764	3.964
$Beta_t$	3,017	1.242	1.161	0.726	0.191	0.769	1.639	2.578
$Size_t$	3,216	4.632	4.504	1.926	1.665	3.227	5.906	8.074
$Momentum_t$	3,026	0.165	0.053	0.685	-0.543	-0.204	0.364	1.216
R_{t+1}	3,216	0.161	0.056	0.696	-0.575	-0.210	0.368	1.213

Panel B: Average Pearson Correlations (above the diagonal) and Spearman Correlations (below the diagonal) between Variables

	Investment$_t$/P$_{t-1}$	ExpInv$_t$/P$_{t-1}$	B$_t$/P$_t$	E$_t$/Pt	CapExpInv$_t$/P$_t$	ROE$_t$	OP$_t$/B$_{t-1}$	CP$_t$/TA$_{t-1}$	Beta$_t$	Size$_t$	Momentum$_t$	R$_{t+1}$-R$^{RF}_{t+1}$
Investment$_t$/P$_{t-1}$		-0.083	-0.075	0.371	-0.171	0.201	0.188	-0.113	-0.015	0.121	0.144	-0.023
		(-4.09)	(-5.06)	(23.00)	(-10.35)	(14.84)	(17.47)	(-5.14)	(-1.44)	(9.10)	(12.80)	(-2.82)
ExpInv$_t$/P$_{t-1}$	-0.048		0.365	-0.172	0.615	-0.147	-0.106	-0.102	-0.015	-0.308	0.098	0.033
	(-2.50)		(41.75)	(-11.18)	(26.32)	(-12.67)	(-10.22)	(-8.13)	(-1.14)	(-33.22)	(6.88)	(3.32)
B$_t$/P$_t$	-0.055	0.333		-0.137	0.374	-0.130	-0.182	-0.139	-0.086	-0.331	-0.267	0.070
	(-4.21)	(26.01)		(-5.31)	(19.66)	(-4.54)	(-7.28)	(-7.94)	(-6.64)	(-32.23)	(-27.96)	(5.48)
E$_t$/P$_t$	0.348	-0.035	0.194		-0.328	0.480	0.346	0.225	-0.084	0.269	0.181	0.025
	(28.62)	(-1.45)	(7.82)		(-12.01)	(32.07)	(27.98)	(19.73)	(-7.29)	(15.91)	(10.21)	(2.07)
CapExpInv$_t$/P$_t$	-0.182	0.828	0.458	-0.083		-0.170	-0.159	-0.128	-0.002	-0.358	-0.188	0.044
	(-8.95)	(121.11)	(38.44)	(-3.05)		(-15.34)	(-15.97)	(-15.16)	(-0.16)	(-32.02)	(-15.66)	(3.79)
ROE$_t$	0.339	-0.231	-0.326	0.652	-0.348		0.825	0.481	-0.058	0.272	0.186	0.042
	(22.76)	(-15.20)	(-12.52)	(19.49)	(-30.77)		(144.30)	(33.04)	(-3.96)	(30.26)	(12.22)	(3.58)
OP$_t$/B$_{t-1}$	0.314	-0.183	-0.319	0.533	-0.308	0.858		0.536	-0.024	0.287	0.203	0.041
	(24.99)	(-11.72)	(-14.26)	(18.65)	(-26.85)	(273.45)		(32.88)	(-1.49)	(28.54)	(14.47)	(3.80)
CP$_t$/TA$_{t-1}$	-0.038	-0.157	-0.256	0.332	-0.216	0.527	0.584		-0.049	0.303	0.126	0.069
	(-1.48)	(-9.93)	(-16.03)	(16.03)	(-18.23)	(52.01)	(55.35)		(-4.18)	(32.58)	(9.59)	(5.39)
Beta$_t$	-0.00	0.006	-0.101	-0.138	0.003	-0.054	-0.046	-0.044		0.013	0.004	-0.038
	(-0.03)	(0.39)	(-6.44)	(-8.47)	(0.16)	(-2.75)	(-2.32)	(-3.55)		(0.69)	(0.21)	(-2.71)
Size$_t$	0.164	-0.344	-0.327	0.239	-0.463	0.407	0.411	0.352	0.038		0.171	-0.028
	(10.52)	(-23.24)	(-23.33)	(8.84)	(-39.91)	(32.72)	(29.06)	(38.06)	(1.85)		(9.63)	(-1.63)
Momentum$_t$	0.213	0.074	-0.333	0.173	-0.289	0.325	0.298	0.211	-0.044	0.271		-0.008
	(21.71)	(4.36)	(-23.55)	(6.84)	(-17.10)	(20.68)	(19.70)	(16.53)	(-1.83)	(13.52)		(-0.68)
$R_{t+1} - R^{RF}_{t+1}$	-0.004	0.033	0.087	0.147	0.042	0.092	0.099	0.123	-0.065	0.059	0.039	
	(-0.40)	(2.26)	(5.59)	(9.45)	(2.79)	(5.72)	(6.41)	(9.01)	(-3.95)	(2.93)	(2.25)	

TABLE 2: Investment and Forward Stock Returns

Panel A reports mean coefficient estimates from annual cross-sectional decile-rank regressions of excess annual forward stock returns, Ret_{t+1} over the risk-free rate, on accounting numbers involving investment, 1964-2018. t-statistics (in parentheses) are calculated as the mean coefficients relative to their standard errors estimated from the time-series of 55 coefficient estimates. Panel B reports mean excess annual returns for portfolios formed from ranking on selected investment measures each March 31st, along with mean return differences and t-statistics based on standard errors calculated from the time series of return differences between the high and low portfolios. Variables are defined in the appendix.

Panel A: Decile Rank Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.173	0.102	0.113	0.101	0.070	0.028	0.040
	(4.21)	(3.35)	(3.68)	(2.64)	(1.39)	(0.61)	(0.73)
$Investment_t/TA_{t-1}$	-0.013						
	(-4.51)						
$ExpInv_t/TA_{t-1}$		0.003					
		(0.78)					
$ExpInv_t/TA_{t-1}$ (w/o SG&A)			0.002				
			(0.40)				
$Investment_t/P_{t-1}$				-0.009	-0.011	-0.008	-0.004
				(-2.95)	(-5.50)	(-3.00)	(-1.93)
$ExpInv_t/P_{t-1}$				0.012	0.014	0.014	0.008
				(4.95)	(5.72)	(5.77)	(3.37)
ROE_t					0.008		
					(2.03)		
CP_t/TA_{t-1}						0.013	0.018
						(5.14)	(10.22)
$Beta_t$							0.000
							(0.06)
B_t/P_t							0.011
							(3.78)
$Size_t$							-0.010
							(-2.67)
MOM_t							-0.006
							(-1.22)
Adjusted R^2	0.007	0.008	0.008	0.014	0.024	0.024	0.060

Panel B: Mean Returns for Portfolios Formed from Ranking on Investment Measures

The High (Low) portfolio consists of firms in the top (bottom) 1/3 of the ranked number each year.

	Portfolios formed by ranking on					
	$\frac{Investment_t}{TA_{t-1}}$	$\frac{ExpInv_t}{TA_{t-1}}$	$\frac{TotInv_t}{TA_{t-1}}$	$\frac{Investment_t}{P_{t-1}}$	$\frac{ExpInv_t}{P_{t-1}}$	$\frac{TotInv_t}{P_{t-1}}$
Low	0.158	0.103	0.145	0.154	0.067	0.122
Medium	0.124	0.117	0.128	0.114	0.121	0.121
High	0.065	0.128	0.076	0.090	0.171	0.115
High-Low	-0.093	0.025	-0.069	-0.064	0.105	-0.007
	(-4.16)	(0.86)	(-3.90)	(-2.44)	(4.31)	(-0.37)

TABLE 3: Book Value and Forward Stock Returns

Panel A reports mean coefficient estimates from annual cross-sectional decile-rank regressions of forward excess stock returns, Ret_{t+1} over the risk-free rate, on accounting numbers involving book value measures, 1964-2018. t-statistics (in parentheses) are calculated as the mean coefficients relative to their standard errors estimated from the time-series of 55 coefficient estimates. Panel B reports mean excess returns for portfolios from ranking on selected measures of book value-to-price each March 31st, along with mean return differences and t-statistics based on standard errors calculated from the time series of return differences between the high and low portfolios. Variables are defined in the appendix.

Panel A: Decile Rank Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.033	0.023	0.048	-0.072	-0.096	0.005	-0.014	0.035
	(1.03)	(0.93)	(1.73)	(-1.69)	(-2.66)	(0.10)	(-0.26)	(0.66)
B_t/P_t	0.018			0.013	0.013	0.010	0.010	-0.005
	(4.49)			(3.40)	(3.44)	(3.13)	(3.13)	(-1.00)
$CapExpInv_t/P_t$		0.020		0.017	0.017	0.014	0.013	
		(4.98)		(5.32)	(4.66)	(5.27)	(4.60)	
$CapExpInv_t/P_t$ (w/o SG&A)			0.016					
			(3.35)					
$TotB_t/P_t$								0.023
								(4.43)
ROE_t				0.011		0.013		0.013
				(3.11)		(4.49)		(4.57)
CP_t					0.016		0.017	
					(6.92)		(9.36)	
$Beta_t$						-0.000	0.000	-0.000
						(-0.11)	(0.07)	(-0.11)
$Size_t$						-0.008	-0.009	-0.009
						(-2.03)	(-2.31)	(-2.32)
MOM_t						-0.003	-0.002	-0.003
						(-0.80)	(-0.57)	(-0.75)
Adjusted R^2	0.015	0.013	0.008	0.033	0.036	0.056	0.058	0.053

Panel B: Mean Returns for Portfolios Formed from Ranking on Book Value Measures

The High (Low) portfolio consists firms in the top (bottom) 1/3 of the ranked number each year.

	Portfolios formed by ranking on		
	$\frac{B_t}{P_t}$	$\frac{CapExpIn}{P_t}$	$\frac{TotB_t}{P_t}$
Low	0.070	0.059	0.064
Medium	0.112	0.110	0.108
High	0.162	0.175	0.170
High-Low	0.091	0.116	0.106
	(3.17)	(3.66)	(3.48)

TABLE 4: ROE and Forward Stock Returns

The table reports mean coefficient estimates from annual cross-sectional decile-rank regressions of forward excess stock returns, Ret_{t+1} over the risk-free rate, on accounting numbers involving ROE and its components, 1964-2018. t-statistics (in parentheses) are calculated as the mean coefficients relative to their standard errors estimated from the time-series of 55 coefficient estimates. Panel A reports results for firms with positive ROE, Panel B for negative ROE. Panel C reports mean excess annual returns on portfolios formed each March 31on E_t/P_t, ROE_t, and $ExpInv_t/P_t$, along with the mean return differences (with t-statistics) between high and low portfolios. Variables are defined in the appendix.

Panel A: Decile Rank Regression Results for Positive Earnings Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	0.015	0.002	0.064	0.124	0.059	0.136	0.052	0.133	0.222	0.131
	(4.44)	(0.06)	(2.20)	(3.49)	(1.90)	(4.00)	(1.74)	(3.37)	(5.42)	(3.37)
ROE_t	-0.005	0.007		-0.011	-0.002	-0.007	-0.004	-0.001	0.005	0.002
	(-2.26)	(2.47)		(-3.33)	(-0.92)	(-1.85)	(-1.53)	(-0.37)	(2.52)	(0.80)
B_t/P_t		0.017			0.009					
		(4.47)			(2.81)					
E_t/P_t			0.012	0.014	0.009	0.013	0.010	0.007		0.007
			(4.38)	(4.48)	(3.31)	(4.22)	(3.19)	(2.54)		(2.22)
$\Delta BV_t/BV_{t-1}$						-0.006				
						(-2.61)				
$ExpInv_t/P_t$							0.013	0.009		0.009
							(4.46)	(3.72)		(3.67)
$Investment_t/TA_{t-1}$									-0.012	-0.008
									(-6.04)	(-4.82)
$Beta_t$								-0.001		
								(-0.40)		
$Size_t$								-0.010	-0.013	-0.010
								(-3.04)	(-3.96)	(-2.97)
MOM_t								-0.002		
								(-0.65)		
Adjusted R^2	0.004	0.019	0.013	0.020	0.022	0.024	0.028	0.056	0.028	0.044

Panel B: Decile Rank Regression Results for Negative Earnings Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	0.076	0.009	0.144	0.143	0.098	0.156	-0.018	-0.020	0.143	0.001
	(1.32)	(0.16)	(2.18)	(2.06)	(1.37)	(2.20)	(-0.29)	(-0.27)	(2.13)	(0.02)
ROE_t	0.037	0.012		0.085	0.055	0.081	0.056	0.052	0.042	0.049
	(3.57)	(0.99)		(6.60)	(5.64)	(6.20)	(4.52)	(4.92)	(4.38)	(3.87)
B_t/P_t		0.019			0.009					
		(3.28)			(2.29)					
E_t/P_t			-0.008	-0.027	-0.020	-0.026	-0.011	-0.010		-0.009
			(-1.25)	(-3.12)	(-2.69)	(-2.81)	(-1.43)	(-1.75)		(-1.06)
$\Delta BV_t/BV_{t-1}$						-0.024				
						(-1.70)				
$ExpInv_t/P_t$							0.022	0.025		0.022
							(4.71)	(4.89)		(4.93)
$Investment_t/TA_{t-1}$									-0.011	-0.005
									(-3.83)	(-1.54)
$Beta_t$								-0.000		
								(-0.11)		
$Size_t$								-0.010	-0.015	-0.005
								(-1.07)	(-2.09)	(-0.85)
MOM_t								-0.003		
								(-0.46)		
Adjusted R^2	0.007	0.020	0.005	0.028	0.033	0.034	0.032	0.023	0.020	0.036

Panel C: Mean Returns for Portfolios formed on E_t/P_t, ROE_t, and $ExpInv_t/P_t$

The High (Low) portfolio consists of the top (bottom) 1/3 of the ranked number each year. NA is not available.

	Negative E/P	Positive E/P	ROE	$ExpInv_t/P_t$
Low	0.143	0.097	0.118	0.057
Medium	0.093	0.112	0.124	0.112
High	0.077	0.141	0.110	0.174
High - Low	-0.066	0.044	-0.008	0.117
	(-1.66)	(2.33)	(-0.24)	(3.53)

TABLE 5: Alternative Profitability Measures and Forward Stock Returns

Panels A and B report mean coefficient estimates from annual cross-sectional decile-rank regressions of forward excess stock returns, Ret_{t+1} over the risk-free rate, on accounting numbers that include operating profitability, OP_t/B_{t-1}, cash flow profitability, CP_t/TA_{t-1}, and a profitability measure, ROE_t^{Adj}, that corrects for the effect of conservative accounting on ROE, 1964-2018. Panel A is for firms with profits, Panel B for loss firms. t-statistics (in parentheses) are calculated as the mean coefficients relative to their standard errors estimated from the time-series of 55 coefficient estimates. Panel C reports mean excess annual returns for portfolios formed each March 31 on CP_t/P_t and OP_t/P_t, along with mean return differences (with t-statistics) between high and low portfolios. Variables are defined in the appendix.

Panel A: Decile Rank Regression Results for Positive Earnings Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Intercept	0.128	0.079	0.155	0.121	0.097	0.132	0.050	0.008	0.059	0.132	0.095	0.126	-0.005	-0.028
	(3.98)	(2.27)	(4.62)	(3.32)	(2.54)	(3.65)	(1.65)	(0.23)	(1.96)	(3.33)	(2.28)	(3.38)	(-0.17)	(-0.83)
OP_t/B_{t-1}	-0.002			0.006			0.005			0.005				
	(-0.71)			(1.70)			(1.39)			(1.65)				
CP_t/TA_{t-1}		0.007			0.011			0.012			0.013	0.010	0.011	
		(3.06)			(4.46)			(4.95)			(5.55)	(3.62)	(4.67)	
CP_t/P_t														0.015
														(5.99)
ROE_t^{Adj}			-0.006			-0.006			-0.005					
			(-3.14)			(-2.58)			(-1.95)					
ROE_t				-0.016	-0.017	-0.007	-0.009	-0.009	-0.001	-0.005	-0.006	-0.010	-0.003	0.003
				(-3.35)	(-4.79)	(-1.86)	(-1.82)	(-3.02)	(-0.32)	(-1.26)	(-1.94)	(-3.88)	(-1.10)	(1.44)
E_t/P_t				0.014	0.015	0.014	0.010	0.010	0.010	0.006	0.006			
				(4.32)	(4.84)	(4.46)	(3.03)	(3.29)	(3.16)	(2.40)	(2.56)			
$ExpInv_t/P_t$							0.012	0.015	0.013	0.008	0.010		0.019	0.014
							(4.39)	(5.07)	(4.50)	(3.62)	(4.10)		(6.13)	(4.60)
$Beta_t$										-0.001	-0.001			
										(-0.36)	(-0.31)			
$Size_t$										-0.011	-0.012			
										(-3.17)	(-3.46)			
MOM_t										-0.003	-0.003			
										(-0.71)	(-0.73)			
Adjusted R^2	0.003	0.008	0.004	0.022	0.029	0.021	0.030	0.037	0.030	0.058	0.063	0.013	0.027	0.030

Panel B: Decile Rank Regression Results for Negative Earnings Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Intercept	0.079	0.057	0.080	0.132	0.099	0.136	-0.028	-0.054	-0.014	-0.125	-0.045	0.040	-0.120	-0.108
	(1.56)	(1.14)	(1.36)	(1.95)	(1.50)	(1.95)	(-0.47)	(-0.93)	(-0.21)	(-0.81)	(-0.66)	(0.70)	(-2.51)	(-2.30)
OP_t/B_{t-1}	0.020			0.016			0.016			0.010				
	(3.43)			(2.94)			(2.82)			(1.23)				
CP_t/TA_{t-1}		0.021			0.018			0.014			0.012	0.017	0.013	
		(5.50)			(4.50)			(3.43)			(2.13)	(4.12)	(2.93)	
CP_t/P_t														0.009
														(3.12)
ROE_t^{Adj}			0.025			0.041			0.013					
			(1.18)			(2.07)			(0.55)					
ROE_t				0.078	0.083	0.056	0.050	0.059	0.047	0.045	0.055	0.030	0.036	0.037
				(6.98)	(6.23)	(4.76)	(4.99)	(4.86)	(4.19)	(5.17)	(5.24)	(2.73)	(3.71)	(3.97)
E_t/P_t				-0.027	-0.028	-0.031	-0.012	-0.013	-0.014	-0.017	-0.006			
				(-3.25)	(-3.19)	(-3.22)	(-1.55)	(-1.76)	(-1.56)	(-1.77)	(-0.93)			
$ExpInv_t/P_t$							0.022	0.021	0.021	0.032	0.024		0.028	0.025
							(4.50)	(4.58)	(4.09)	(2.66)	(4.26)		(4.59)	(4.41)
$Beta_t$										0.008	-0.004			
										(0.75)	(-0.72)			
$Size_t$										-0.038	-0.002			
										(-1.14)	(-0.21)			
MOM_t										0.014	-0.008			
										(0.61)	(-1.79)			
Adjusted R^2	0.001	0.006	0.008	0.026	0.023	0.035	0.031	0.020	0.046	0.033	0.041	0.011	0.025	0.016

Panel C: Mean Annual Returns for Portfolios formed on CP_t/P_t, and OP_t/P_t

The High (Low) portfolio consists of the top (bottom) 1/3 of the ranked number each year. NA is not available.

	Negative CP	Positive CP	Negative OP	Positive OP
Low	0.082	0.081	0.102	0.094
Medium	0.047	0.125	0.068	0.118
High	0.040	0.181	0.046	0.163
High - Low	-0.042	0.100	-0.057	0.069
	(-1.48)	(3.64)	(-1.68)	(2.41)

Table 6: Mean Annual Returns for the SUMIS Factor and Component Portfolios

Panel A reports mean excess annual returns over years for portfolios formed by sorting firms each year on E_t/P_t and ROE_t. Panel B reports returns from sorting on E_t/P_t, ROE_t, and $ExpInv_t/P_t$ and for the SUMIS factor. High (Low) portfolios consist of the top (bottom) 1/3 of the ranked number each year. Portfolios are formed at March 31 each year, 1964-2018. Return differences between indicated portfolios are also reported, with t-statistics calculated as the mean return differences relative to their standard errors calculated from the time series of return differences. Variables are defined in the appendix.

Panel A: Mean Returns for Portfolios formed on E_t/P_t and ROE_t

		Positive Earnings			Diagonal Spread Return
	Negative Earnings	Low E/P	Medium E/P	High E/P	a – b
1: High ROE_t	0.138	0.080[b]	0.100	0.128	
2: Medium	0.131	0.087	0.101	0.139	
3: Low ROE_t	0.067	0.133	0.138	0.156[a]	
3 − 1	-0.071	0.053	0.038	0.028	0.076
	(-2.58)	(2.54)	(3.20)	(2.16)	(2.57)

Panel B: Mean Returns for Portfolios formed on E_t/P_t, ROE_t, and Expensed Investment ($ExpInv_t/P_t$)

		Positive Earnings			Diagonal Spread Return
	Negative Earnings	Low E/P	Medium E/P	High E/P	a – b (SUMIS)
1: High ROE_t and Low $ExpInv_t/P_t$	0.064	0.054[b]	0.064	0.098	
2: Medium	0.107	0.095	0.110	0.140	
3: Low ROE_t and High $ExpInv_t/P_t$	0.146	0.183	0.182	0.187[a]	
3 − 1	0.082	0.129	0.118	0.089	0.133
	(1.52)	(4.02)	(4.60)	(3.01)	(3.23)

Table 7: Mean Annual Returns for the SUMBS Factor and Component Portfolios

The table reports mean excess annual returns from sorting firms each year on the change in net operating assets relative to price ($\Delta NOA_t/P_{t-1}$) and financing leverage ($Debt_t/P_t$) and for the SUMBS factor. High (Low) portfolios consist of the top (bottom) 1/3 of the ranked number each year. Portfolios are formed at March 31 each year, 1964-2018. Return differences between indicated portfolios are also reported, with t-statistics calculated as the mean return differences relative to their standard errors calculated from the time series of return differences. Variables are defined in the appendix.

Mean Returns for Portfolios formed on $\Delta NOA_t/P_{t-1}$ and $Debt_t/P_t$

	$\Delta NOA_t/P_{t-1}$			Diagonal Spread Return
	Low	Medium	High	a – b (SUMBS)
1: Low $Debt_t/P_t$	0.141	0.103	0.080[b]	
2: Medium $Debt_t/P_t$	0.159	0.101	0.088	
3: High $Debt_t/P_t$	0.179[a]	0.129	0.098	
3 − 1	0.038	0.025	0.018	0.099
	(1.08)	(1.30)	(0.97)	(3.06)

Table 8: Mean Annual Returns for the SUMIS Factor and Component Portfolios Differentiated by Firm Size

Panel A reports median ln(Size) for the portfolios for forming SUMIS in Table 6. Panel B reports mean annual excess returns for the SUMIS portfolios when microcap stocks are excluded. Panel C also excludes microcaps but reports returns for the remaining (larger) firms that have the accounting characteristics that determine the SUMIS portfolios.

Panel A: Median ln(Market Capitalization) for Portfolios Underlying the SUMIS Factor

		Positive Earnings			Diagonal Spread
	Negative Earnings	Low E/P	Medium E/P	High E/P	a – b
1: High ROE_t and Low $ExpInv_t/P_t$	4.375	6.116[b]	5.961	5.238	
2: Medium	3.356	5.207	5.487	4.596	
3: Low ROE_t and High $ExpInv_t/P_t$	2.360	3.450	3.889	3.550[a]	
3 − 1	-2.015	-2.667	-2.071	-1.688	-2.566
	(-17.58)	(-32.05)	(-24.80)	(-16.17)	

Panel B: Mean Returns for Portfolios Underlying the SUMIS Factor, Excluding Microcaps

Microcap stocks are those below the 20th percentile of NYSE stocks by size.

		Positive Earnings			Diagonal Spread Return
	Negative Earnings	Low E/P	Medium E/P	High E/P	a – b
1: High ROE_t and Low $ExpInv_t/P_t$	0.064	0.056[b]	0.073	0.095	
2: Medium	0.055	0.083	0.097	0.116	
3: Low ROE_t and High $ExpInv_t/P_t$	0.157	0.149	0.127	0.143[a]	
3 − 1	0.093	0.093	0.054	0.048	0.086
	(1.73)	(3.17)	(3.03)	(1.71)	(2.14)

Panel C: Mean Returns for Portfolios Underlying the SUMIS Factor, Excluding Microcaps but with Returns in Rows 1 and 3 only for those Non-microcaps Satisfying the Cut-offs for Portfolios in Table 6

		Positive Earnings			Diagonal Spread Return
	Negative Earnings	Low E/P	Medium E/P	High E/P	a – b
1: High ROE_t and Low $ExpInv_t/P_t$	0.055	0.055[b]	0.073	0.094	
2: Medium	0.060	0.083	0.097	0.116	
3: Low ROE_t and High $ExpInv_t/P_t$	0.151	0.149	0.127	0.145[a]	
3 − 1	0.096	0.094	0.054	0.052	0.091
	(1.62)	(3.17)	(3.03)	(1.84)	(2.23)

Table 9: Tests of Accounting Factors Against Fama and French Factors

The table reports estimated intercepts, α, and slope coefficients, β, from time-series regressions of monthly excess returns for portfolios underlying SUMIS (in Panel B of Table 6) on Fama and French factor returns. In Panel A, the returns are for the high-risk, long portfolios for each E/P portfolio in the SUMIS factor (on row 3 in Panel B of Table 6) and in Panel B they are for the low-risk, short portfolios (on row 1). These are differentiated on ROE and expensed investment. Panel C is the return difference between the long and short portfolios (rows 3 – 1). The All portfolio in each panel is all positive E/P stocks pooled into one portfolio. The return in the last column in Panels A and B is that for the high positive E/P portfolio minus that low positive E/P portfolio in row 3 and row 1 respectively. In Panel C, it is that for SUMIS capturing the diagonal spread in Panel B of Table 6, a – b. Portfolios are reformed each March 31, 1964 to March 31, 2018, with the monthly returns running from April, 1964 to March, 2019. The Fama and French factor returns are from the Kenneth French web library at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Panel A: Long (High Risk) Portfolios (Low ROE_t, High $ExpInv_t/P_t$) for E/P Portfolios

	Negative Earnings	Positive Earnings				High E/P – Low E/P Portfolio
		Low E/P	Medium E/P	High E/P	All	
α	0.013	0.011	0.009	0.010	0.010	0.010
	(3.79)	(8.45)	(9.06)	(7.73)	(10.77)	(7.73)
β_{MKT}	0.987	0.959	0.967	1.005	0.976	1.005
	(11.56)	(29.32)	(38.96)	(32.88)	(42.77)	(32.88)
β_{HML}	0.114	0.308	0.360	0.605	0.419	0.605
	(0.67)	(4.76)	(7.33)	(10.00)	(9.27)	(10.00)
β_{SMB}	1.683	1.345	1.047	1.122	1.169	1.122
	(13.97)	(29.15)	(29.30)	(26.04)	(36.31)	(26.04)
β_{UMD}	-0.371	-0.235	-0.131	-0.319	-0.226	-0.319
	(-4.54)	(-7.50)	(-5.50)	(-10.91)	(-10.32)	(-10.91)
β_{RMW}	-1.197	0.076	0.166	0.296	0.177	0.296
	(-7.07)	(1.17)	(3.37)	(4.87)	(3.92)	(4.87)
β_{CMA}	-0.002	0.071	0.194	0.110	0.130	0.110
	(-0.01)	(0.76)	(2.71)	(1.25)	(1.98)	(1.25)
Adjusted R^2	0.55	0.81	0.85	0.81	0.88	0.81

Panel B: Short (Low Risk) Portfolios (High ROE_t, Low $ExpInv_t/P_t$) for E/P Portfolios

	Negative Earnings	Positive Earnings				High E/P – Low E/P Portfolio
		Low E/P	Medium E/P	High E/P	All	
α	0.005	0.002	0.002	0.003	0.002	0.002
	(2.80)	(2.30)	(1.99)	(2.50)	(3.12)	(2.30)
β_{MKT}	1.131	1.099	1.053	1.059	1.069	1.099
	(25.58)	(46.42)	(49.26)	(41.34)	(63.72)	(46.42)
β_{HML}	0.421	-0.339	0.111	0.444	0.077	-0.339
	(4.80)	(-7.23)	(2.62)	(8.75)	(2.30)	(-7.23)
β_{SMB}	1.115	0.603	0.682	0.836	0.705	0.603
	(17.87)	(18.06)	(22.63)	(23.14)	(29.80)	(18.06)
β_{UMD}	-0.113	0.004	-0.088	-0.240	-0.110	0.004
	(-2.67)	(0.16)	(-4.30)	(-9.77)	(-6.86)	(0.16)
β_{RMW}	-0.361	0.062	0.358	0.458	0.281	0.062
	(-4.11)	(1.32)	(8.44)	(9.01)	(8.43)	(1.32)
β_{CMA}	-0.208	-0.390	-0.118	-0.178	-0.234	-0.390
	(-1.64)	(-5.73)	(-1.93)	(-2.43)	(-4.86)	(-5.73)
Adjusted R^2	0.73	0.88	0.87	0.84	0.92	0.88

Panel C: Long minus Short Portfolios within E/P Portfolios

	Negative Earnings	Positive Earnings				SUMIS Portfolio: Diagonal Spread a-b
		Low E/P	Medium E/P	High E/P	All	
α	0.008	0.009	0.007	0.007	0.008	0.007
	(2.27)	(5.44)	(6.14)	(4.49)	(7.27)	(4.42)
β_{MKT}	-0.142	-0.139	-0.086	-0.054	-0.093	-0.094
	(-1.66)	(-3.42)	(-2.89)	(-1.41)	(-3.48)	(-2.28)
β_{HML}	-0.306	0.647	0.249	0.161	0.343	0.944
	(-1.79)	(8.01)	(4.24)	(2.12)	(6.50)	(11.58)
β_{SMB}	0.567	0.742	0.365	0.286	0.464	0.520
	(4.68)	(12.90)	(8.71)	(5.29)	(12.35)	(8.95)
β_{UMD}	-0.256	-0.239	-0.043	-0.080	-0.115	-0.323
	(-3.12)	(-6.11)	(-1.50)	(-2.16)	(-4.52)	(-8.19)
β_{RMW}	-0.835	0.014	-0.192	-0.163	-0.103	0.234
	(-4.90)	(0.17)	(-3.26)	(-2.13)	(-1.95)	(2.86)
β_{CMA}	0.207	0.461	0.312	0.288	0.364	0.500
	(0.84)	(3.93)	(3.66)	(2.62)	(4.76)	(4.23)
Adjusted R^2	0.12	0.41	0.25	0.11	0.39	0.49

Table 10: Accounting Factors and *q*-Investment Factors

The table reports estimated intercepts, α, and slope coefficients, β, from time-series regressions of monthly excess returns for portfolios underlying SUMIS (in Panel B of Table 6) on factor returns for *q*-investment models. As is Table 9, Panel A is for the high-risk, long portfolios for each E/P portfolio (on row 3 in Panel B of Table 6) and Panel B is for the low-risk, short portfolios (on row 1). These are differentiated on ROE and expensed investment. Panel C is the return difference between the long and short portfolios (rows 3 – 1). The All portfolio in each panel is all positive E/P stocks pooled into one portfolio. The return for the summary portfolio in Panels A and B is that for the high positive E/P portfolio minus that low positive E/P portfolio in row 3 and row 1 respectively. In Panel C, it is that for SUMIS capturing the diagonal spread in Panel B of Table 6, a – b. Portfolios are reformed each March 31, 1964 to March 31, 2018, with the monthly returns running from April, 1964 to March, 2019. The *q*-investment factor returns are at http://global-q.org/index.html.

Panel A: Long (High Risk) Portfolios (Low ROE_t, High $ExpInv_t/P_t$) for E/P Portfolios

	Negative Earnings	Positive Earnings				
		Low E/P	Medium E/P	High E/P	All	High E/P – Low E/P Portfolio
α	0.012	0.011	0.009	0.010	0.010	0.010
	(3.22)	(7.15)	(7.87)	(6.02)	(8.56)	(6.02)
β_{MKT}	1.085	0.970	0.945	1.002	0.971	1.002
	(13.41)	(27.93)	(35.27)	(27.84)	(36.98)	(27.84)
β_{IA}	-0.210	0.100	0.322	0.558	0.325	0.558
	(-1.14)	(1.26)	(5.26)	(6.80)	(5.42)	(6.80)
β_{ME}	0.015	0.012	0.010	0.010	0.010	0.010
	(13.66)	(25.00)	(26.31)	(19.70)	(28.96)	(19.70)
β_{ROE}	-1.514	-0.537	-0.285	-0.510	-0.440	-0.510
	(-10.29)	(-8.50)	(-5.85)	(-7.79)	(-9.22)	(-7.79)
β_{EG}	0.306	0.187	0.098	0.152	0.146	0.152
	(1.40)	(2.00)	(1.36)	(1.57)	(2.06)	(1.57)
Adjusted R^2	0.59	0.80	0.83	0.76	0.86	0.76

Panel B: Short (Low Risk) Portfolios (High ROE_t, Low $ExpInv_t/P_t$) for E/P Portfolios

	Negative Earnings	Positive Earnings				
		Low E/P	Medium E/P	High E/P	All	High E/P – Low E/P Portfolio
α	0.009	0.003	0.003	0.005	0.003	0.003
	(4.91)	(2.65)	(2.42)	(3.48)	(4.19)	(2.65)
β_{MKT}	1.100	1.106	1.024	1.010	1.045	1.106
	(26.56)	(44.10)	(44.54)	(32.65)	(57.10)	(44.10)
β_{IA}	0.064	-0.763	-0.126	0.101	-0.266	-0.763
	(0.68)	(-13.34)	(-2.40)	(1.43)	(-6.36)	(-13.34)
β_{ME}	0.010	0.006	0.006	0.007	0.006	0.006
	(16.83)	(16.18)	(19.16)	(16.11)	(24.38)	(16.18)
β_{ROE}	-0.542	0.269	0.142	-0.067	0.103	0.269
	(-7.20)	(5.89)	(3.40)	(-1.18)	(3.10)	(5.89)
β_{EG}	-0.252	-0.230	-0.152	-0.200	-0.198	-0.230
	(-2.26)	(-3.40)	(-2.45)	(-2.40)	(-4.01)	(-3.40)
Adjusted R^2	0.77	0.88	0.86	0.79	0.92	0.88

Panel C: Long minus Short Portfolios within E/P Portfolios

	Negative Earnings	Positive Earnings				SUMIS Portfolio: Diagonal Spread a-b
		Low E/P	Medium E/P	High E/P	All	
α	0.003	0.008	0.007	0.005	0.007	0.007
	(0.69)	(4.22)	(5.18)	(2.81)	(5.49)	(3.19)
β_{MKT}	-0.015	-0.136	-0.079	-0.008	-0.074	-0.105
	(-0.18)	(-3.15)	(-2.64)	(-0.21)	(-2.74)	(-2.23)
β_{IA}	-0.274	0.863	0.447	0.457	0.590	1.322
	(-1.45)	(8.78)	(6.54)	(5.18)	(9.59)	(12.33)
β_{ME}	0.006	0.006	0.004	0.003	0.004	0.004
	(4.91)	(10.73)	(8.82)	(5.45)	(11.65)	(6.45)
β_{ROE}	-0.971	-0.806	-0.427	-0.443	-0.543	-0.778
	(-6.42)	(-10.27)	(-7.84)	(-6.30)	(-11.07)	(-9.10)
β_{EG}	0.557	0.417	0.250	0.352	0.343	0.382
	(2.49)	(3.59)	(3.10)	(3.38)	(4.72)	(3.02)
Adjusted R^2	0.12	0.39	0.28	0.16	0.43	0.37

Table 11: Explaining Factor Returns with Accounting Factors

Panels A and B report estimated intercepts, α, and slope coefficients, β, from time-series regressions of monthly excess returns for selected zero-net-investment portfolios on the excess return on the market factor, MKT, and the summary accounting factor from the income statement and balance sheet, SUMIS and SUMBS. The market factor is the return on the CRSP value-weighted return index. The SUMIS factor captures the diagonal spread of returns in Panel B in Table 6, a – b, based on E/P, ROE, and expensed investment. The SUMBS factor is constructed from the diagonal spread of returns in Table 7 based on investment and financing leverage. Panel A and C portfolios involve the full range of observations, including loss firms. HML, SMB, and RMW in Panel B are Fama and French factors returns tabulated at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. *q*-ROE and *q*-EG are factors in *q*-investment models tabulated at http://global-q.org/index.html. Portfolios in Panels A and C, along with SUMIS and SUMBS, are reformed each March 31, 1964 to March 31, 2018, with the monthly returns running from April, 1964 to March, 2019. Panels C and D add the EG factor return.

Panel A: Dependent Variables are Returns for Portfolios Formed from Ranking on Selected Variables. High (low) portfolios consist of the top (bottom) 1/3 of stocks.

	High B/P– Low B/P	High TotB/P – Low TotB/P	High E/P – Low E/P	Small Size– Large Size	High ROE – Low ROE
α	0.001	0.002	-0.000	0.000	0.003
	(1.18)	(2.68)	(-0.30)	(0.08)	(1.78)
β_{MKT}	-0.104	-0.054	-0.101	0.063	-0.218
	(-6.98)	(-3.62)	(-7.16)	(1.78)	(-6.26)
β_{SUMIS}	0.276	0.296	0.261	0.219	-0.140
	(18.24)	(19.69)	(18.31)	(6.08)	(-4.00)
β_{SUMBS}	0.446	0.433	0.126	0.443	-0.198
	(18.86)	(18.37)	(5.62)	(7.87)	(-3.61)
Adjusted R^2	0.76	0.76	0.60	0.29	0.14

Panel B: Dependent Variables are Factor Returns from Extant Factor Models

	HML	SMB	RMW	*q*-ROE	*q*-EG
α	-0.001	0.000	0.003	0.008	0.009
	(-0.61)	(0.00)	(3.83)	(8.57)	(13.31)
β_{MKT}	-0.120	0.224	-0.121	-0.146	-0.193
	(-6.46)	(8.58)	(-6.36)	(-7.02)	(-12.47)
β_{SUMIS}	0.182	0.080	-0.034	-0.140	-0.003
	(9.70)	(3.05)	(-1.78)	(-6.42)	(-0.21)
β_{SUMBS}	0.265	0.016	0.037	-0.047	-0.010
	(9.03)	(0.39)	(1.25)	(-1.39)	(-0.39)
Adjusted R^2	0.46	0.11	0.05	0.17	0.20

Panel C: Panel A with the Expected Growth Factor, EG, Added

	High B/P– Low B/P	High TotB/P – Low TotB/P	High E/P – Low E/P	Small Size– Large Size	High ROE – Low ROE
α	0.000	0.001	-0.001	0.002	-0.004
	(0.24)	(1.73)	(-1.70)	(2.73)	(-2.08)
β_{MKT}	-0.088	-0.043	-0.075	-0.042	-0.089
	(-5.28)	(-2.55)	(-4.80)	(-1.05)	(-2.31)
β_{SUMIS}	0.279	0.306	0.268	0.235	-0.148
	(17.97)	(19.51)	(18.28)	(6.31)	(-4.11)
β_{SUMBS}	0.442	0.422	0.116	0.418	-0.176
	(18.34)	(17.33)	(5.09)	(7.25)	(-3.16)
β_{EG}	0.102	0.067	0.139	-0.536	0.695
	(2.64)	(1.73)	(3.83)	(-5.81)	(7.81)
Adjusted R^2	0.77	0.77	0.62	0.33	0.21

Panel D: Panel B with the Expected Growth Factor, EG, Added

	HML	SMB	RMW	*q*-ROE
α	-0.002	0.005	-0.001	0.002
	(-1.78)	(4.16)	(-0.97)	(1.93)
β_{MKT}	-0.099	0.108	-0.033	-0.016
	(-4.66)	(3.89)	(-1.62)	(-0.78)
β_{SUMIS}	0.188	0.093	-0.033	-0.137
	(9.50)	(3.59)	(-1.76)	(-7.29)
β_{SUMBS}	0.261	-0.005	0.047	-0.010
	(8.51)	(-0.11)	(1.60)	(-1.38)
β_{EG}	0.131	-0.605	0.466	0.675
	(2.68)	(-9.41)	(9.89)	(14.47)
Adjusted R^2	0.47	0.22	0.18	0.38